

Estadística

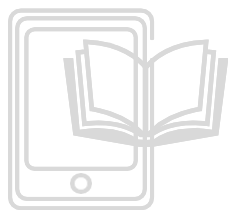
para ingenieros
y científicos



William Navidi

**Mc
Graw
Hill**





BASE LEGAL PARA DIGITALIZACIÓN DE LIBROS CON DERECHOS DE AUTOR

El Sistema Integrado de Bibliotecas de la Universidad Central del Ecuador, digitalizará su acervo bibliográfico siempre y cuando sea para fines educativos y de investigación. No se permite la reproducción y distribución para la comercialización directa e indirecta del mismo.

La digitalización del material bibliográfico se lo realiza de acuerdo al Código Orgánico De La Economía Social de los Conocimientos, Creatividad e Innovación Art. 212 numeral 9, Literal ii “(...) Una biblioteca o archivo podrá, además, realizar los siguientes actos (...) La reproducción electrónica y comunicación pública de obras de su colección para ser consultadas gratuita y simultáneamente hasta por un número razonable de usuarios, sólo en terminales de redes de la respectiva institución o para usuarios de esa institución bajo su control, en condiciones que garanticen que no se puedan hacer copias electrónicas de esas reproducciones” y literal vii “La reproducción, adaptación, traducción, transformación, arreglo, distribución y comunicación de una obra protegida por derechos de autor o una prestación protegida por derechos conexos, en uno o más formatos accesibles para el uso exclusivo de personas con discapacidad”.

Este material se considera un producto intelectual a favor de su autor; por tanto, la titularidad de sus derechos se encuentra protegida por el Código Orgánico De La Economía Social de los Conocimientos, Creatividad e Innovación. La violación a dichos derechos constituye un delito que será responsabilidad del usuario.

Este libro está disponible físicamente en:

BIBLIOTECA	Biblioteca Facultad de Filosofía, Letras y Ciencias de la Educación	PISO	01
		ESTANTE	04
		BANDEJA	01
CONVENIO INTERINSTITUCIONAL		UNIVERSIDAD	

Estadística para ingenieros





Estadística para ingenieros

William Navidi

Colorado School of Mines

Traducción

Ana Elizabeth García Hernández

Profesora de Estadística

Universidad Iberoamericana, Ciudad de México

Revisión técnica

Nicolás Gómez Castillo

Profesor de Matemáticas Aplicadas

(Probabilidad y Estadística, Series de tiempo)

Universidad Iberoamericana, Ciudad de México



MÉXICO • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA • LISBOA
MADRID • NUEVA YORK • SAN JUAN • SANTIAGO
AUCKLAND • LONDRES • MILÁN • MONTREAL • NUEVA DELHI
SAN FRANCISCO • SINGAPUR • SAN LOUIS • SIDNEY • TORONTO

Director Higher Education: Miguel Ángel Toledo Castellanos
Director editorial: Ricardo A. del Bosque Alayón
Editor sponsor: Pablo Eduardo Roig Vázquez
Editora de desarrollo: Paula Montaña González
Supervisor de producción: Zeferino García García

ESTADÍSTICA PARA INGENIEROS

Prohibida la reproducción total o parcial de esta obra,
por cualquier medio, sin la autorización escrita del editor.

 **McGraw-Hill
Interamericana**

DERECHOS RESERVADOS © 2006 respecto a la primera edición en español por
McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.

A subsidiary of The McGraw-Hill Companies, Inc.

Prolongación Paseo de la Reforma 1015, Torre A

Piso 17, Colonia Desarrollo Santa Fe,

Delegación Álvaro Obregón

C. P. 01376, México, D. F.

Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736

ISBN 970-10-5629-9

Traducido de la primera edición de: STATISTICS FOR ENGINEERS AND SCIENTISTS

Copyright © 2006 by The McGraw-Hill Companies, Inc. All rights reserved.

ISBN 0-07-255160-7

1234567890

09875432106

Impreso en México

Printed in Mexico

Para Catherine, Sarah y Thomas



ACERCA DEL AUTOR

William Navidi es profesor de Matemáticas y Ciencias de la Computación en la Escuela de Minas de Colorado. Obtuvo la licenciatura en Matemáticas por el New College, la maestría en Matemáticas por la Universidad Estatal de Michigan y el doctorado en Estadística por la Universidad de California, en Berkeley. El profesor Navidi ha escrito más de 50 artículos de investigación, tanto en teoría estadística como en una vasta gama de aplicaciones, entre éstas se incluyen Redes Computacionales, Epidemiología, Biología Molecular, Ingeniería Química y Geofísica.



CONTENIDO BREVE

Prefacio	xv
Reconocimientos a revisores y colaboradores	xix
Características clave	xxiii
1 Muestreo y estadística descriptiva	1
2 Probabilidad	50
3 Propagación de errores	157
4 Distribuciones comúnmente usadas	192
5 Intervalos de confianza	300
6 Pruebas de hipótesis	368
7 Correlación y regresión lineal simple	475
8 Regresión múltiple	556
9 Experimentos factoriales	623
10 Control estadístico de calidad	723
Apéndice A: Tablas	763
Apéndice B: Derivadas parciales	787
Apéndice C: Bibliografía	789
Respuestas a los ejercicios impares	792
Índice	859



CONTENIDO

Prefacio xv

**Reconocimientos a revisores
y colaboradores** xix

Características clave xxiii

Capítulo 1

Muestreo y estadística descriptiva 1

Introducción 1

1.1 Muestreo 3

1.2 Resumen estadístico 13

1.3 Resúmenes gráficos 25

Capítulo 2

Probabilidad 50

Introducción 50

2.1 Ideas básicas 50

2.2 Métodos de conteo (opcional) 62

2.3 Probabilidad condicional
e independencia 69

2.4 Variables aleatorias 88

2.5 Funciones lineales de variables
aleatorias 111

2.6 Variables aleatorias con distribución
conjunta (opcional) 120

Capítulo 3

Propagación de errores 157

Introducción 157

3.1 Error de medición 157

3.2 Combinaciones lineales
de las mediciones 163

3.3 Incertidumbres para funciones
de una medición 173

3.4 Incertidumbres para funciones
de varias mediciones 179

Capítulo 4

Distribuciones comúnmente usadas 192

Introducción 192

4.1 Distribución de Bernoulli 192

4.2 La distribución binomial 195

4.3 La distribución de Poisson 206

4.4 Algunas otras distribuciones
discretas 220

4.5 Distribución normal 231

4.6 Distribución lognormal 244

4.7 La distribución exponencial 250

4.8 Las distribuciones gamma
y de Weibull 259

4.9 Gráficas de probabilidad 265

4.10 El teorema del límite central 270

4.11 Simulación 281

Capítulo 5

Intervalos de confianza 300

Introducción 300

5.1 Intervalos de confianza para la
media poblacional con muestras
grandes 301

5.2 Intervalos de confianza para
proporciones 315

5.3 Intervalos de confianza para la
media poblacional con muestras
pequeñas 321

5.4 Intervalos de confianza para la
diferencia entre dos medias 331

- 5.5** Intervalos de confianza para la diferencia entre dos proporciones 335
- 5.6** Intervalos de confianza para la diferencia entre dos medias con pequeñas muestras 339
- 5.7** Intervalos de confianza con datos apareados 346
- 5.8** Uso de simulación para construir intervalos de confianza 351

Capítulo 6

Pruebas de hipótesis 368

- Introducción 368
- 6.1** Pruebas de hipótesis para la media poblacional con muestras grandes 368
- 6.2** Concluir a partir de las pruebas de hipótesis 377
- 6.3** Pruebas de hipótesis para la proporción poblacional 385
- 6.4** Pruebas de hipótesis para la media poblacional con muestras pequeñas 390
- 6.5** Pruebas de hipótesis para la diferencia entre dos medias con muestras grandes 395
- 6.6** Pruebas de hipótesis para la diferencia entre dos proporciones 401
- 6.7** Pruebas de hipótesis para la diferencia entre dos medias con muestras pequeñas 407
- 6.8** Pruebas de hipótesis con datos apareados 415
- 6.9** Pruebas de hipótesis con distribuciones libres 420
- 6.10** La prueba χ^2 cuadrada 429
- 6.11** La prueba F para la igualdad de varianzas 439
- 6.12** Prueba de nivel fijo 443

- 6.13** Potencia 449
- 6.14** Pruebas múltiples 458
- 6.15** Uso de la simulación para realizar pruebas de hipótesis 462

Capítulo 7

Correlación y regresión lineal simple 475

- Introducción 475
- 7.1** Correlación 475
- 7.2** La recta de mínimos cuadrados 492
- 7.3** Incertidumbres en los coeficientes de mínimos cuadrados 508
- 7.4** Comprobación de supuestos y transformación de datos 527

Capítulo 8

Regresión múltiple 556

- Introducción 556
- 8.1** El modelo de regresión múltiple 556
- 8.2** Confusión y colinealidad 574
- 8.3** Selección de modelos 583

Capítulo 9

Experimentos factoriales 623

- Introducción 623
- 9.1** Experimentos de un solo factor 623
- 9.2** Comparaciones apareadas en los experimentos de un solo factor 646
- 9.3** Experimentos de dos factores 659
- 9.4** Diseños de bloque completamente aleatorios 683
- 9.5** Experimentos factoriales 2^p 693

Capítulo 10

Control estadístico de calidad 723

- Introducción 723
- 10.1** Ideas básicas 723

10.2	Diagramas de control para variables	726
10.3	Diagramas de control para atributos	746
10.4	El diagrama CUSUM	751
10.5	Capacidad del proceso	755

Apéndice A: Tablas 763

Apéndice B: Derivadas parciales 787

Apéndice C: Bibliografía 789

Respuestas a los ejercicios impares 792

Índice 859





PREFACIO

MOTIVACIÓN

La idea de realizar este libro surgió de pláticas entre los profesores de estadística e ingeniería de la Escuela de Minas de Colorado respecto de nuestro curso de “Introducción a la estadística para ingenieros”. Nuestros profesores de ingeniería pensaban que los estudiantes necesitaban cubrir ampliamente el tema de propagación del error, así como un mayor énfasis en las habilidades en el ajuste de modelos. Los profesores de estadística creían que los estudiantes necesitaban estar más conscientes de algunos puntos importantes en la práctica de la estadística, como la comprobación de los supuestos del modelo y del uso de la simulación.

Mi punto de vista es que un libro introductorio a la estadística para ingenieros y científicos debe ofrecer todos estos temas con cierta profundidad. Además, debe ser lo suficientemente flexible para permitir diferentes elecciones del material que debe cubrirse, ya que hay muchas formas para diseñar un curso exitoso de introducción a la estadística. Finalmente, éste debe proporcionar ejemplos que presenten ideas importantes en contextos reales. De acuerdo con lo anterior, el libro tiene las siguientes características:

- El libro es flexible en su presentación de probabilidad, ello permite a los profesores elegir la profundidad y extensión de la cobertura de este tema.
- El libro contiene muchos ejemplos en contexto real y con conjuntos de datos actuales; lo anterior motiva a los estudiantes y muestra la interrelación entre la industria y la investigación científica.
- El libro contiene muchos ejemplos con resultados obtenidos mediante computadora y ejercicios adecuados para resolverse con algún software estadístico.
- El libro proporciona una extensa cobertura de la propagación del error.
- El libro presenta una espléndida introducción a los métodos de simulación y a la estimación bootstrap, incluyendo aplicaciones para comprobar supuestos de normalidad, cálculo de probabilidades, estimación del sesgo, cálculo de intervalos de confianza y pruebas de hipótesis.
- El libro proporciona una mayor cobertura en los procedimientos de diagnóstico del modelo lineal que la que se encuentra en la mayoría de los textos introductorios. Ésta incluye material acerca del diagnóstico de la gráfica de los residuales, transformación de variables y principios de selección de variables en los modelos multivariados.
- El libro cubre los temas introductorios usuales, tales como estadística descriptiva, probabilidad, intervalos de confianza, pruebas de hipótesis, regresión lineal, experimentos factoriales y control estadístico de calidad.

NIVEL MATEMÁTICO

La mayor parte del libro será matemáticamente accesible a todas las personas que hayan estudiado un semestre de cálculo. Las excepciones son la propagación multivariada del error, que

requiere derivadas parciales y las distribuciones de probabilidad conjunta, que requieren de integración múltiple. Estos temas se los puede saltar en una primera lectura, si así lo desea.

USO DE LA COMPUTADORA

En los últimos 20 años el desarrollo de computadoras rápidas y baratas ha revolucionado la práctica de la estadística; efectivamente, ésta es una de las razones principales del porqué los métodos estadísticos han estado penetrando cada vez más en el trabajo científico. Los científicos y los ingenieros actuales no sólo deben ser expertos en el manejo de paquetes de software, sino que también deben contar con la habilidad para concluir a partir de los resultados computacionales y expresar estas conclusiones en palabras. De acuerdo con esto, el libro contiene ejercicios y ejemplos que requieren la interpretación, así como la generación de resultados por computadora, especialmente en los capítulos de modelos lineales y experimentos factoriales.

La disponibilidad actual de computadoras y de paquetes computacionales estadísticos también ha producido un importante beneficio en la educación, al hacer accesibles los métodos de simulación a los estudiantes de los cursos de introducción. La simulación hace que los principios fundamentales de la estadística revivan. El material de simulación que aquí se presenta está diseñado para reforzar algunas ideas estadísticas básicas e introducir a los estudiantes en algunos de los usos de esta poderosa herramienta.

CONTENIDO

El capítulo 1 cubre el muestreo y la estadística descriptiva. La razón por la que los métodos estadísticos funcionan es que las muestras, cuando se toman en forma adecuada, semejan a la población. Por consiguiente, el capítulo 1 empieza con la descripción de algunas formas de tomar muestras válidas. En la segunda parte del capítulo se analiza la estadística descriptiva.

El capítulo 2 trata la probabilidad. Existe una gran discrepancia en las preferencias de los profesores acerca de qué tanto y tan profundamente se debe cubrir este tema. Por lo tanto, se ha tratado de hacer este capítulo lo más flexible posible. Los resultados principales se deducen de axiomas, demostrando la mayoría de ellos. Esto último permitirá a los profesores establecer un enfoque matemático riguroso. Por otra parte, he intentado mostrar cada resultado con uno o dos ejemplos, en donde sea posible un contexto científico que esté diseñado para presentar la intuición que se encuentra detrás del resultado. Por tanto, los profesores que prefieran un enfoque más informal se pueden dedicar a los ejemplos más que a las demostraciones.

En el capítulo 3 se presenta el tema de la propagación del error, que algunas veces se llama “análisis del error” o, por los estadísticos, “el método delta”. La cobertura es más amplia que en la mayoría de los textos, pero el tema es tan importante que pensé que era útil. La presentación está diseñada para permitir que los profesores ajusten la cantidad de temas que debe cubrir de acuerdo con las necesidades del curso.

En el capítulo 4 se presentan muchas de las funciones de distribución de probabilidad comúnmente usadas en la práctica. También se tratan las gráficas de probabilidad y el teorema del límite central. En la última sección se presentan los métodos de simulación para evaluar los supuestos de normalidad, cálculo de probabilidades y estimación de sesgo.

Los capítulos 5 y 6 tratan los intervalos de confianza y las pruebas de hipótesis, respectivamente. Se hace énfasis en el enfoque del P -valor para las pruebas de hipótesis, pero tam-

bién se presentan pruebas de nivel-fijo y el cálculo de la potencia. El problema de pruebas múltiples se trata con cierta profundidad. También se presentan métodos de simulación para calcular intervalos de confianza y realizar pruebas de hipótesis.

En el capítulo 7 se trata la correlación y la regresión lineal simple. He trabajado arduamente para enfatizar que los modelos lineales sólo son apropiados cuando la relación entre las variables es lineal. Este punto es muy importante ya que con frecuencia ingenieros y científicos lo ignoran (sin mencionar a los estadísticos). No es difícil encontrar en la bibliografía científica ajustes lineales y coeficientes de correlación resumidos en gráficas que presentan una curvatura evidente o en las cuales la pendiente de la recta se ve afectada mediante algunos puntos influyentes. Por tanto, en este capítulo se incluye una larga sección para comprobar los supuestos del modelo y la transformación de variables.

En el capítulo 8 se trata el tema de la regresión múltiple. Se hace un énfasis especial en los métodos de selección de modelo, ya que la selección de variables que se incluirán en el modelo constituye un paso esencial en muchos análisis de la vida real. También el tema de la confusión se trata cuidadosamente.

En el capítulo 9 se analizan algunos diseños experimentales y los métodos que comúnmente se aplican para analizar sus datos. Los métodos de análisis de varianza en uno y dos sentidos junto con el diseño de bloques completamente aleatorios y los diseños factoriales 2^p se tratan con amplitud.

En el capítulo 10 se presenta el tema del control de calidad estadístico, se analizan los diagramas CUSUM, la capacidad del proceso y se concluye con una breve descripción de la calidad con six-sigma.

MATERIAL RECOMENDADO

El libro contiene suficiente material para un curso de un año. Si se requiriera un curso de un semestre, hay varias opciones. En nuestro curso de tres horas en la Escuela de Minas de Colorado cubrimos el total de los primeros cuatro capítulos, excepto las distribuciones conjuntas, la exponencial, la gamma y de Weibull. Después se cubren los temas de intervalos de confianza y las pruebas de hipótesis en los capítulos 5 y 6, tocando rápidamente los métodos de dos muestras y los cálculos de potencia y se omiten los métodos de distribución libre y las pruebas de Ji cuadrada y F . Terminamos cubriendo todo el material posible, que el tiempo restante permita, sobre correlación y regresión lineal simple del capítulo 7.

Se puede planear un curso con un sentido diferente que incluya más material de probabilidad, invirtiendo más tiempo en los métodos de dos muestras y de potencia y reduciendo la cobertura de la propagación del error, simulación o regresión. Hay muchas otras opciones; por ejemplo, se puede elegir incluir material de experimentos factoriales en lugar de algunos de los temas anteriores. En el manual del profesor que está disponible en el centro de aprendizaje en línea (Online Learning Center) en www.mhhe.com/navidi se puede encontrar una variedad de enfoques y duración de cursos.

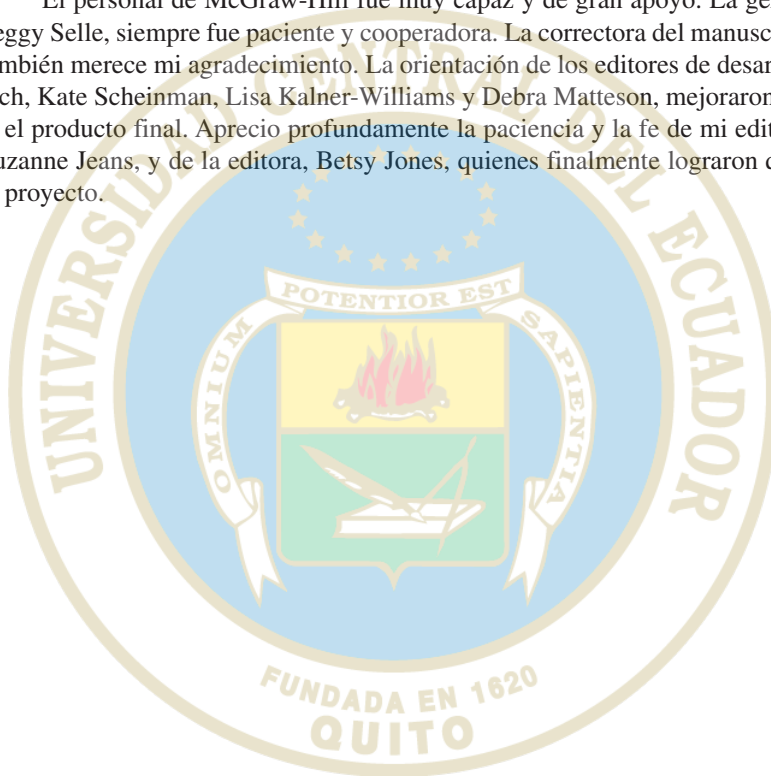
AGRADECIMIENTOS

Estoy en deuda con muchas de las personas que contribuyeron en cada etapa del desarrollo. Mis colegas de la división de ingeniería en la Escuela de Minas de Colorado fueron pacientes y generosos al ayudarme a evaluar los caminos en los que las ideas estadísticas interactúan con la práctica de la ingeniería; Terry Parker merece mención especial por esto. Otros

colegas, que impartieron clases y estudiantes que estudiaron en borradores del manuscrito encontraron muchos errores y me hicieron sugerencias muy valiosas. En particular, Barbara Moskal y Gus Greivel enseñaron varias veces con el manuscrito en evolución, lo que fue de ayuda y apoyo durante todo el tiempo; asimismo, Melissa Laeser encontró muchos conjuntos de datos interesantes en fuentes publicadas. Mike Colagrosso, de la Escuela de Minas de Colorado, desarrolló algunos applets excelentes, así como Chris Boisclair y el equipo en sistemas de enlace. Jessica Kohlschmidt, de la Universidad Estatal de Ohio, desarrolló diapositivas en PowerPoint para complementar el texto y Jackie Miller, también de la Universidad Estatal de Ohio, encontró muchos errores en todo el manuscrito e hizo valiosas sugerencias para mejorarlo.

El personal de McGraw-Hill fue muy capaz y de gran apoyo. La gerente del proyecto, Peggy Selle, siempre fue paciente y cooperadora. La correctora del manuscrito, Lucy Mullins, también merece mi agradecimiento. La orientación de los editores de desarrollo, Maja Lorkovich, Kate Scheinman, Lisa Kalner-Williams y Debra Matteson, mejoraron considerablemente el producto final. Aprecio profundamente la paciencia y la fe de mi editora-patrocinadora, Suzanne Jeans, y de la editora, Betsy Jones, quienes finalmente lograron que se realizara este proyecto.

William Navidi



RECONOCIMIENTOS A REVISORES Y COLABORADORES

Este texto refleja las contribuciones generosas de más de un centenar de profesores de estadística y de sus estudiantes, quienes, a través de numerosas evaluaciones, encuestas y pruebas en clase, nos ayudaron a comprender cómo cubrir sus necesidades y cómo mejorar cuando no lo lográbamos. Las ideas de estos profesores y sus estudiantes están inmersas en todo el libro, desde su contenido y organización hasta sus complementos.

El autor y el equipo del área de ingeniería de McGraw-Hill estamos agradecidos con aquéllos por sus atentos comentarios y contribuciones durante el desarrollo del texto, de sus complementos y de los recursos multimedia.

Sabri Abou-Ward

University of Toronto, Canada

Jeff Arthur

Oregon State University

Georgiana Baker

University of South Carolina

Barb Barnett

University of Wisconsin, Platteville

Samantha Bates

Virginia Polytechnic Institute

Beno Benhabib

University of Toronto, Canada

L. Bernaards

University of California, Los Angeles

Mary Besterfield-Saore

University of Pittsburgh

J. D. Bhattacharyya

Iowa State University

Amy Biesterfeld

University of Colorado, Boulder

Rick Bilonick

Carnegie Mellon University

Peter Bloomfield

North Carolina State University

Saul Blumenthal

Ohio State University

Robert J. Boik

Montana State University

Tamara Bowcutt

California State University at Chico

Donald Boyd

Montana State University

Raymond Brown

Drexel University

Ronald B. Bucinell

Union College

Karen M. Bursic

University of Pittsburgh

John Butler

*California Polytechnic State University,
San Luis Obispo*

M. Jaya Chandra

Pennsylvania State University

Huann-Shen Chen

Michigan Technological University

Nicolas Christou

University of California, Los Angeles

Dave Clark

Kettering University

Peyton Cook

University of Tulsa

William Cooper

University of Minnesota

Casey Cremins

University of Maryland

Dan Dalthorp

Cornell University

Valentin Deaconu

University of Nevada, Reno

Darinka Dentcheva

Stevens Institute of Technology

Art Diaz

San Jose State University

Don Edwards

University of Toronto, Canada

Judith Ekstrand

San Francisco State University

Timothy C. Elston

North Carolina State University

Randy Eubank

Texas A&M University

Thomas Z. Fahidy

University of Waterloo, Canada

Nasser Fard

Northeastern University

Steve From

University of Nebraska at Omaha

Gary Gadbury

University of Missouri, Rolla

Pierre Gharghouri

Ryerson University, Canada

Sampson Gholston

University of Alabama

Robert Gilbert

The University of Texas at Austin

Dave Goldman

Georgia Institute of Technology

Pierre Goovaerts

University of Michigan

Canan Bilen Green

North Dakota State University

Trevor Hale

Ohio University

Jim Handley

*Montana Tech of the
University of Montana*

James Higgins

Kansas State University

David Hoeppener

University of Utah

Carol O'Connor Holloman

University of Louisville

Joseph Horowitz

University of Massachusetts, Amherst

Wei-Min Huang

Lehigh University

Norma Hubele

Arizona State University

Floyd Hummel

Pennsylvania State University

Wafik Iskander

West Virginia University

Roger Johnson

*South Dakota School of Mines
and Technology*

Scott Jordan

Arkansas Tech University

Kailash Kapur

University of Washington

David Kender

Wright State University

Kim Wee Kong

Multimedia University, Malaysia

Ravindra Khattree

Oakland University

Vadim Khayms

Stanford University

Claire Komives

San Jose State University

Thomas Koon

Binghamton University

Milo Koretsky

Oregon State University

Roger Korus

University of Idaho

Tomasz Kozubowski

University of Nevada, Reno

Gary Kretchik

University of Alaska, Anchorage

Hillel Kumin

University of Oklahoma

Samuel Labi

Purdue University

Robert Lacher

South Dakota State University

John Lawson

Brigham Young University

Stephen Lee

University of Idaho

Marvin Lentner

Virginia Polytechnic Institute

Liza Levina

University of Michigan

Quingchong John Liu

Oakland University

Graham Lord

Princeton University

Arthur Lubin

Illinois Institute of Technology

Zhen Luo

Pennsylvania State University

James Lynch

University of South Carolina

Peter MacDonald

McMaster University, Canada

James Maneval

Bucknell University

Munir Mahmood

Rochester Institute of Technology

Glen Marotz

University of Kansas

Timothy Matis

New Mexico State University

Laura McSweeney

Fairfield University

Megan Meece

University of Florida

Vivien Miller

Mississippi State University

Bradley Mingels

University of Massachusetts, Lowell

Satya Mishra

University of South Alabama

N. N. Murulidhar

National Institute of Technology, India

Steve Ng Hoi-Kow

Hong Kong Institute of Vocational

Education, Hong Kong

Zulkifli Mohd Nopiah

Universiti Kebangsaan Malaysia, Malaysia

Bob O'Donnell

Oregon State University

Nancy Olmstead

Milwaukee School of Engineering

K. P. Patil

Veermeeta Jijabai Technological

Institute, India

Arunkumar Pennathur

University of Texas at El Paso

Joseph Petrucelli

Worcester Polytechnic University

Elizabeth Podlaha

Louisiana State University

Michael Pore

The University of Texas at Austin

Jeffrey Proehl*Dartmouth College***Jorge Prozzi***The University of Texas at Austin***D. Ramachandran***California State University at Sacramento***Amelia Regan***University of California, Irvine***Larry Ringer***Texas A&M University***Iris V. Rivero***Texas Tech University***Timothy Robinson***University of Wyoming***Derrick Rollins***Iowa State University***Andrew Ross***Lehigh University***Manual Rossetti***University of Arkansas***V. A. Samaranayake***University of Missouri, Rolla***Doug Schmucker***Valparaiso University***David Schradly***Naval Postgraduate School***Neil Schwertman***California State University, Chico***Nong Shang***Rensselaer Polytechnic Institute***Galit Shmueli***Carnegie Mellon University***Ruey-Lin Sheu***National Cheng Kung University, Taiwan***Tony Smith***University of Pennsylvania***Jery Stedinger***Cornell University***Frank E. Stratton***San Diego State University***John Ting***University of Massachusetts, Lowell***Gilberto Urroz***Utah State University***Margot Vigeant***Bucknell University***Joseph G. Voelkel***Rochester Institute of Technology***Natascha Vukasinovic***Utah State University***Amy Wagler***Oklahoma State University***Tse-Wei Wang***University of Tennessee, Knoxville***Bill Warde***Oklahoma State University***Simon Washington***University of Arizona***Daniel Weiner***Boston University***Alison Weir***University of Toronto at**Mississauga, Canada***Bruce Westermo***San Diego State University***Grant Willson***The University of Texas at Austin***Jae Yoon***Old Dominion University***Ali Zargar***San Jose State University*

Características clave

Resumen del contenido

Este libro permite cubrir el material de una forma flexible ya que hay muchas maneras de diseñar un curso exitoso de introducción a la estadística.

- **Cobertura flexible de probabilidad**
dirigida a las diferentes necesidades de los cursos. Permite el enfoque matemático riguroso, los principales resultados son deducidos de axiomas al proporcionar demostraciones en la mayoría de ellos. Por otra parte, cada resultado se muestra con uno o dos ejemplos para motivar intuitivamente la comprensión. Los profesores que prefieran un enfoque más informal pueden, por tanto, dedicarse a los ejemplos en lugar de las demostraciones y omitir las secciones opcionales.
- **Cobertura extensa de propagación del error**, algunas veces se llama “análisis de errores” o “el método delta”, se trata en un capítulo exclusivo del tema. La cobertura es más minuciosa que en la mayoría de los textos. El formato es flexible con el propósito de que la cantidad del material se adapte a las necesidades del curso.
- **Una sólida introducción a los métodos de simulación y a la estimación bootstrap**
se presenta en las secciones finales de los capítulos 4, 5 y 6.
- **Cobertura extensa de los procedimientos de diagnóstico del modelo lineal**
en el capítulo 7, se incluye una gran sección de la comprobación de supuestos del modelo y de la transformación de variables. El capítulo enfatiza que los modelos lineales sólo son apropiados cuando la relación entre las variables es lineal. Este punto es el más importante, ya que con frecuencia se ignora en la práctica de ingenieros y científicos (sin mencionar a los estadísticos).

Capítulo 1

Muestreo y estadística descriptiva 1

Capítulo 2

Probabilidad 50

2.1 Ideas básicas 50

2.2 Métodos de conteo (opcional) 62

2.3 Probabilidad condicional e independencia 69

2.4 Variables aleatorias 88

2.5 Funciones lineales de variables aleatorias 111

2.6 Variables aleatorias con distribución conjunta (opcional) 120

Capítulo 3

Propagación de errores 157

3.1 Error de medición 157

3.2 Combinaciones lineales de las mediciones 163

3.3 Incertidumbres para funciones de una medición 173

3.4 Incertidumbres para funciones de varias mediciones 179

Capítulo 4

Distribuciones comúnmente usadas 192

4.11 Simulación 281

- Uso de la simulación para calcular una probabilidad
- Cálculo de medias y varianzas
- Usos de la simulación en análisis de confiabilidad
- Uso de la simulación para calcular sesgamiento
- La estimación autosuficiente

Capítulo 5

Intervalos de confianza 300

5.8 Uso de simulación para construir intervalos de confianza 351

- Intervalos de confianza usando estimación bootstrap
- Uso de simulación para evaluar intervalos de confianza

Capítulo 6

Pruebas de hipótesis 368

6.15 Uso de la simulación para realizar pruebas de hipótesis 462

- Pruebas de hipótesis con intervalos de confianza de estimación bootstrap
- Pruebas aleatorias
- Uso de simulación para calcular la potencia

Capítulo 7

Correlación y regresión lineal simple 475

7.4 Comprobación de supuestos y transformación de datos 527

- La gráfica de los residuos contra valores ajustados
- Transformación de variables
- Determinación de la transformación que se aplicará
- Las transformaciones no siempre funcionan
- Las gráficas de los residuales con sólo pocos puntos pueden ser difíciles de interpretar
- Puntos atípicos e influyentes
- Otros métodos para transformar variables
- Pruebas de independencia y normalidad
- Modelos empíricos y leyes físicas

Capítulo 8

Regresión múltiple 556

Capítulo 9

Experimentos factoriales 623

Capítulo 10

Control estadístico de calidad 723

Características clave

El artículo "Virgin Versus Recycled Wafers for Furnace Qualification: Is the Expense Justified?" (V. Czitrom y J. Reece, en *Statistical Case Studies for Industrial Process Improvement*, ASA y SIAM, 1997:87-104) describe un proceso para el crecimiento de una capa delgada de dióxido de silicio sobre placas de silicio que se usan en la fabricación de semiconductores. La tabla 1.6 presenta las mediciones del espesor, en angstroms (Å), de la capa de óxido para 24 placas. Se hicieron nueve mediciones en cada placa. Las placas se fabricaron en dos corridas distintas, con 12 placas por cada corrida.

TABLA 1.6 Espesor de las capas de óxido de silicio en placas de silicio

Placa		Espesor (Å)								
Corrida 1	1	90.0	92.2	94.9	92.7	91.6	88.2	82.0	98.2	96.0
	2	91.8	94.5	93.9	77.3	92.0	89.9	87.9	92.8	93.3
	3	90.3	91.1	93.3	93.5	87.2	88.1	90.1	91.9	94.5
	4	92.6	90.3	92.8	91.6	92.7	91.7	89.3	95.5	93.6
	5	91.1	89.8	91.5	91.5	90.6	93.1	88.9	92.5	92.4
	6	76.1	90.2	96.8	84.6	93.3	95.7	90.9	100.3	95.2
	7	92.4	91.7	91.6	91.1	88.0	92.4	88.7	92.9	92.6
	8	91.3	90.1	95.4	89.6	90.7	95.8	91.7	97.9	95.7
	9	96.7	93.7	93.9	87.9	90.4	92.0	90.5	95.2	94.3
	10	92.0	94.6	93.7	94.0	89.3	90.1	91.3	92.7	94.5
	11	94.1	91.5	95.3	92.8	93.4	92.2	89.4		
	12	91.7	97.4	95.1	96.7	77.5	91.4	90.1		
Corrida 2	1	93.0	99.9	93.6	89.0	93.6	90.9	89.8		
	2	91.4	90.6	92.2	91.9	92.4	87.6	88.1		
	3	91.9	91.8	92.8	96.4	93.8	86.5	92.1		
	4	90.6	91.3	94.9	88.3	87.9	92.2	90.1		
	5	93.1	91.8	94.6	88.9	90.0	97.9	92.1		
	6	90.8	91.5	91.5	91.5	94.0	91.0	92.1		
	7	88.0	91.8	90.5	90.4	90.3	91.5	89.4		
	8	88.3	96.0	92.8	93.7	89.6	89.6	90.1		
	9	94.2	92.2	95.8	92.5	91.0	91.4	92.1		
	10	101.5	103.1	103.2	103.5	96.1	102.5	102.1		
	11	92.8	90.8	92.2	91.7	89.0	88.5	87.1		
	12	92.1	93.4	94.0	94.7	90.8	92.1	91.1		

Las 12 placas en cada corrida eran de varios tipos y se procesaron en el horno. El propósito en la recopilación de datos fue determinar si el espesor de la capa de óxido se afectaba ya sea por el tipo de placa o por la posición en el horno. Este fue un experimento factorial, con los factores, tipo de placa y posición. El resultado del experimento fue que no hubo una diferencia sistemática entre las placas de una corrida y las de la otra. El análisis de los datos de la corrida 1 mostró que el espesor de la capa de óxido era mayor en las placas de tipo 1 y 2 que en las de tipo 3 y 4. En el análisis de los datos de la corrida 2 se supuso ninguna diferencia sistemática entre las placas de una corrida y las de la otra. El análisis de los datos de la corrida 2 mostró que el espesor de la capa de óxido era mayor en las placas de tipo 1 y 2 que en las de tipo 3 y 4. El experimento se diseñó para ayudar a determinar si esta condición se satisfacía realmente. Las observaciones se debían eliminar. Los resultados se presentan en la figura 1.6.

Conjuntos de datos del mundo real

Con un enfoque fresco del tema, el autor usa datos del mundo real actuales para motivar a los estudiantes mostrando una conexión con la industria y la investigación.

22. El artículo "Seismic Hazard in Greece Based on Different Strong Ground Motion Parameters" (S. Koutrakis, G. Karakaisis y cols., en *Journal of Earthquake Engineering*, 2002:75-109) presenta un estudio de episodios sísmicos en Grecia durante 1978-1997. Es de interés la duración de los "fuertes movimientos de tierra", que es el tiempo en que la aceleración de la tierra excede de un valor específico. En cada episodio las mediciones de la duración de temblores fuertes de tierra se hicieron en una o más ubicaciones. La tabla SE22 de la página 618 presenta cada uno de 121 temblores medidos, los datos con el tiempo de duración m (en segundos) durante los cuales la aceleración de la tierra excedió el doble de la aceleración de la gravedad. La magnitud m del sismo, la distancia d (en km) de la medición desde el epicentro, y los dos indicadores del tipo de suelo s_1 y s_2 , definidos de la siguiente manera: $s_1 = 1$ si el suelo consta de depósitos aluviales blandos, $s_1 = 0$ de otra manera, y $s_2 = 1$ si el suelo consta de rocas terciarias o más viejas, $s_2 = 0$ de otra manera. Los casos donde tanto $s_1 = 0$ como $s_2 = 0$ corresponden a condiciones intermedias del suelo. El artículo presenta mediciones repetidas en algunas ubicaciones que no se incluyen aquí.

TABLA SE22 Datos para el ejercicio 22

y	m	d	s_1	s_2	y	m	d	s_1	s_2	y	m	d	s_1	s_2
8.82	6.4	30	1	0	4.31	5.3	6	0	0	5.74	5.6	15	0	0
4.08	5.2	7	0	0	28.27	6.6	31	1	0	5.13	6.9	128	1	0
15.90	6.9	105	1	0	17.94	6.9	33	0	0	3.20	5.1	13	0	0
6.04	5.8	15	0	0	3.60	5.4	6	0	0	7.29	5.2	19	1	0
0.15	4.9	16	1	0	7.98	5.3	12	1	0	0.02	6.2	68	1	0
5.06	6.2	75	1	0	16.23	6.2	13	0	0	7.03	5.4	10	0	0
0.01	6.6	119	0	1	3.67	6.6	85	1	0	2.17	5.1	45	0	1
4.13	5.1	10	1	0	6.44	5.2	21	0	0	4.27	5.2	18	1	0
0.02	5.3	22	0	1	10.45	5.3	11	0	1	2.25	4.8	14	0	1
2.14	4.5	12	0	1	8.32	5.5	22	1	0	3.10	5.5	15	0	0
4.41	5.2	17	0	0	5.43	5.2	49	0	1	6.18	5.2	13	0	0
17.19	5.9	9	0	0	4.78	5.5	1	0	0	4.56	5.5	1	0	0
5.14	5.5	10	1	0	2.82	5.5	20	0	1	0.94	5.0	6	0	1
0.05	4.9	14	1	0	3.51	5.7	22	0	0	2.85	4.6	21	1	0
20.00	5.8	16	1	0	13.92	5.8	34	1	0	4.21	4.7	20	1	0
12.04	6.1	31	0	0	3.96	6.1	44	0	0	1.93	5.7	39	1	0
0.87	5.0	65	1	0	6.91	5.4	16	0	0	1.56	5.0	44	1	0
0.62	4.8	11	1	0	5.63	5.3	6	1	0	5.03	5.1	2	1	0
8.10	5.4	12	1	0	0.10	5.2	21	1	0	0.51	4.9	14	1	0
1.30	5.8	34	1	0	5.10	4.8	16	1	0	13.14	5.6	5	1	0
11.92	5.6	5	0	0	16.52	5.5	15	1	0	8.16	5.5	12	1	0
3.93	5.7	65	1	0	19.84	5.7	50	1	0	10.04	5.1	28	1	0
2.00	5.4	27	0	1	1.65	5.4	27	1	0	0.79	5.4	35	0	0
0.43	5.4	31	0	1	1.75	5.4	30	0	1	0.02	5.4	32	1	0
14.22	6.5	20	0	1	6.37	6.5	90	1	0	0.10	6.5	61	0	1
0.06	6.5	72	0	1	2.78	4.9	8	0	0	5.43	5.2	9	0	0
1.48	5.2	27	0	0	2.14	5.2	22	0	0	0.81	4.6	9	0	0
3.27	5.1	12	0	0	0.92	5.2	29	0	0	0.73	5.2	22	0	0
6.36	5.2	14	0	0	3.18	4.8	15	0	0	11.18	5.0	8	0	0
0.18	5.0	19	0	0	1.20	5.0	19	0	0	2.54	4.5	6	0	0
0.31	4.5	12	0	0	4.37	4.7	5	0	0	1.55	4.7	13	0	1
1.90	4.7	12	0	0	1.02	5.0	14	0	0	0.01	4.5	17	0	0
0.29	4.7	5	1	0	0.71	4.8	4	1	0	0.21	4.8	5	0	1
6.26	6.3	9	1	0	4.27	6.3	9	0	1	0.04	4.5	3	1	0
3.44	5.4	4	1	0	3.25	5.4	4	0	1	0.01	4.5	1	1	0
2.32	5.4	5	1	0	0.90	4.7	4	1	0	1.19	4.7	3	1	0
1.49	5.0	4	1	0	0.37	5.0	4	0	1	2.66	5.4	1	1	0
2.85	5.4	1	0	1	21.07	6.4	78	0	1	7.47	6.4	104	0	0
0.01	6.4	86	0	1	0.04	6.4	105	0	1	30.45	6.6	51	1	0
9.34	6.6	116	0	1	15.30	6.6	82	0	1	12.78	6.6	65	1	0
10.47	6.6	117	0	0										

Con los datos de la tabla SE22 construya un modelo lineal para pronosticar la duración y a partir de alguna o de todas las variables m , d , s_1 y s_2 . Asegúrese de considerar las transformaciones de las variables, así como las potencias de y y las interacciones entre las variables independientes. Describa sus pasos para construir su modelo. Realice una gráfica de residuos contra valores ajustados para comprobar que su modelo satisface los supuestos necesarios. Además, observe que los datos se presentan en orden cronológico, al leer hacia abajo en las columnas. Realice una gráfica para determinar si se debe incluir al tiempo como una variable independiente.

Características clave

Resultados de computadora

El libro contiene ejercicios y ejemplos que requieren la interpretación y la generación de resultados por medio de computadora.

12. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para una media poblacional μ .

One-Sample Z: X

Test of $\mu = 73.5$ vs not = 73.5
The assumed standard deviation = 2.3634

Variable	N	Mean	StDev	SE Mean	95% CI	Z	P
X	145	73.2461	2.3634	0.1963	(72.8614, 73.6308)	-1.29	0.196

- ¿Es ésta una prueba de una cola o de dos colas?
- ¿Cuál es la hipótesis nula?
- ¿Cuál es el P -valor?
- Utilice el resultado y una tabla adecuada para calcular el P -valor para la prueba de $H_0: \mu \geq 73.6$ contra $H_1: \mu < 73.6$
- Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 99% para μ .

7. En un estudio de la función pulmonar de niños, el volumen de aire exhalado por la fuerza en un segundo se llama FEV₁, (FEV₁ es el volumen de espiración forzada en un segundo.) Se hicieron mediciones en un grupo de niños cada año durante dos años. Se ajustó a un modelo lineal para pronosticar los FEV₁ de estos años como una función del FEV₁ (en litros) del último año, el sexo del niño (0 = masculino, 1 = femenino), la estatura del niño (en m), y la presión atmosférica ambiental (en mm). El siguiente resultado de MINITAB presenta los resultados de ajuste del modelo

$$\text{FEV}_1 = \beta_0 + \beta_1 \text{ Último FEV}_1 + \beta_2 \text{ Sexo} + \beta_3 \text{ Estatura} + \beta_4 \text{ Presión} + \epsilon$$

The regression equation is

FEV1 = -0.219 + 0.779 Last FEV - 0.108 Gender + 1.354 Height - 0.00134 Pressure

Predictor	Coef	SE Coef	T	P
Constant	-0.21947	0.4503	-0.49	0.627
Last FEV	0.779	0.04909	15.87	0.000
Gender	-0.10827	0.0352	-3.08	0.002
Height	1.3536	0.2880	4.70	0.000
Pressure	-0.0013431	0.0004722	-2.84	0.005

S = 0.22039

R-Sq = 93.5%

R-Sq(adj) = 93.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	111.31	27.826	572.89	0.000
Residual Error	160	7.7716	0.048572		
Total	164	119.08			

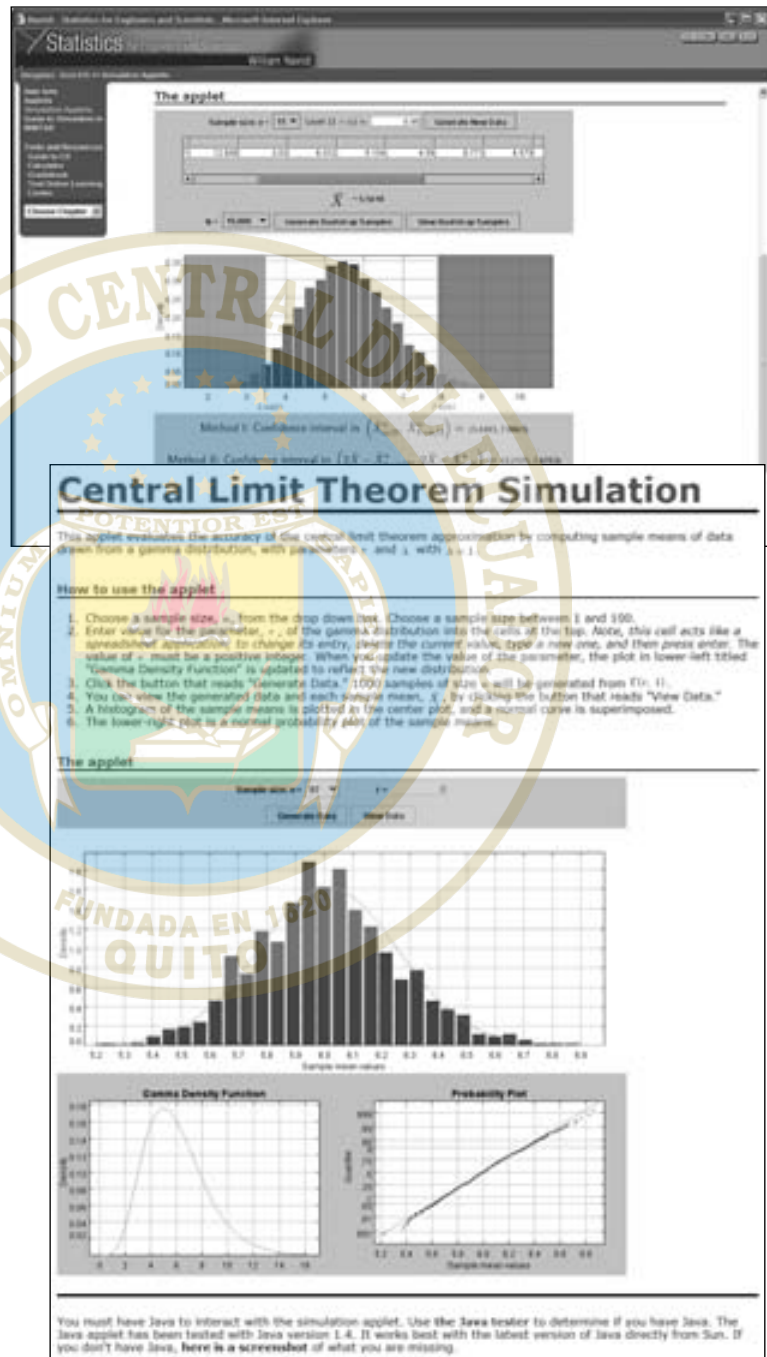
- Pronostique el FEV₁ para un niño con estatura de 1.4 m, si la medida se tomó a presión de 730 mm y la medición del último año fue 2.113 L.
- Si dos niñas difieren en estatura por 5 cm, ¿qué tanto esperaría que sus mediciones de FEV₁ difieran; los otros conceptos siguen igual?
- Se estima que el término constante β_0 es negativo, pero el FEV₁ debe ser siempre positivo. ¿Algo está erróneo? Explique.
- El responsable de este experimento quiere rediseñar el algoritmo que registra las mediciones electrónicamente con el fin de ajustar la presión atmosférica automáticamente. Se fija un barómetro al dispositivo para registrar la presión. Utilice el resultado anterior de MINITAB para determinar cómo calcular un valor FEV₁ ajustado como función del valor FEV₁ medido y de la presión.

Complementos de aprendizaje para los estudiantes

CD-ROM con recursos para los estudiantes

Empaquetado gratis con cada libro nuevo, este CD proporciona todos los conjuntos de datos del texto, así como applets basados en el contenido del texto para reforzar un entendimiento visual de la estadística.

- **Todos los conjuntos de datos** se pueden descargar en diferentes formatos:
 - ASCII delimitado con comas
 - ASCII delimitado con tabuladores
 - MINITAB
 - Excel
 - SAS
 - SPSS
 - TI-89
- **Applets de Java**, creados específicamente para los cálculos de este curso, proporcionan ejercicios interactivos basados en el contenido del texto, lo que permite a los estudiantes modificar las variables y explorar escenarios de "¿Qué sucede si?". También se incluyen en la suite de applet los **applets de simulación**, que refuerzan la excelente cobertura de texto de los métodos de simulación. Los applets permiten que los estudiantes vean los ejemplos de simulación del texto en acción y que modifiquen los parámetros para una mayor exploración.
- **Una guía a la simulación con MINITAB** preparada por el autor donde se describe cómo se pueden implementar en MINITAB los ejemplos de simulación en el texto.
- **Herramientas y recursos**, que incluye un vínculo al centro de aprendizaje del libro, ofrece en línea recursos para el profesor y el estudiante en www.mhhe.com/navidi.



Material de apoyo

Esta obra cuenta con interesantes complementos que fortalecen los procesos de enseñanza-aprendizaje, así como la evaluación de éstos. Mismos que se otorgan a profesores que adopten este texto para sus cursos.

Para obtener más información y conocer la política de entrega de estos materiales, contacte a su representante de McGraw-Hill o envíe un correo electrónico a marketinghe@mcgraw-hill.com

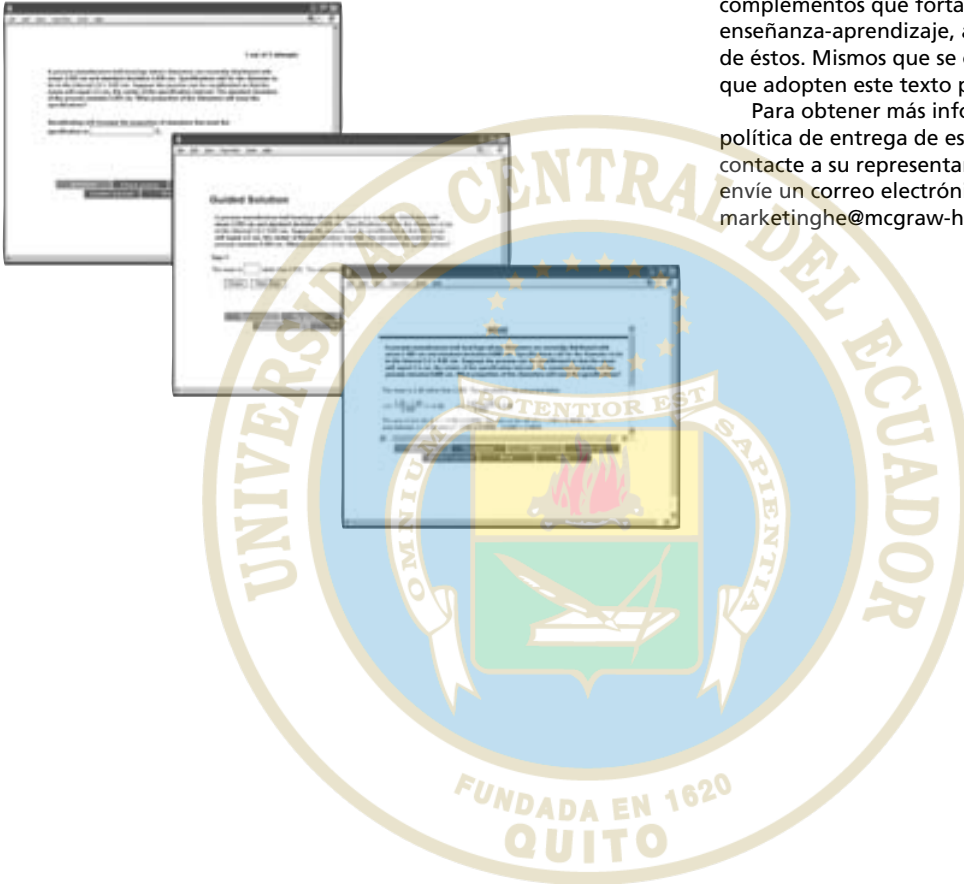
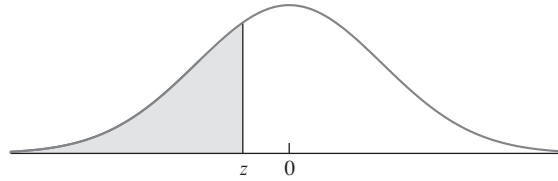


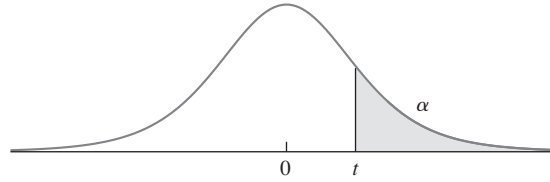


TABLA A.2 Distribución normal acumulativa (tabla z)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

TABLA A.3 Puntos porcentuales superiores para la distribución t de Student



ν	α								
	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
35	0.255	0.682	1.306	1.690	2.030	2.438	2.724	3.340	3.591
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Capítulo 1

Muestreo y estadística descriptiva

Introducción

La recopilación y el análisis de datos son fundamentales en la ciencia e ingeniería. Al analizar los datos recopilados en experimentos, los científicos descubren los principios que gobiernan el mundo físico y los ingenieros aprenden cómo diseñar nuevos productos y procesos importantes. Una dificultad muy importante que se presenta con los datos científicos es que éstos se encuentran sujetos a variaciones aleatorias o incertidumbre. Es decir, cuando se repiten las mediciones científicas cada vez salen un poco diferentes. Lo anterior plantea un problema: ¿cómo se pueden obtener conclusiones de los resultados de un experimento cuando éstos pueden ser diferentes? Para analizar esta pregunta, es esencial contar con cierto conocimiento estadístico. La estadística se dedica a la recopilación, el análisis y la interpretación de datos con incertidumbre. Los métodos de la estadística permiten que los científicos e ingenieros diseñen experimentos válidos y obtengan conclusiones confiables a partir de datos obtenidos.

Aunque nuestro interés en este libro es tratar con las aplicaciones de la estadística en la ciencia y en la ingeniería, cabe mencionar que el análisis y la interpretación de datos son cada vez más importantes en todos los aspectos de la vida moderna. Para bien o para mal, se están recopilando enormes cantidades de datos con nuestras opiniones y estilos de vida, con fines que van desde la creación de campañas de mercadotecnia más eficaces hasta el desarrollo de políticas sociales diseñadas para mejorar nuestro estilo de vida. Casi a diario, los artículos que se publican en los periódicos pretenden explicar las tendencias sociales o económicas a través del análisis de datos. Por tanto, un conocimiento básico de estadística es necesario no sólo para ser un científico o ingeniero eficiente, sino también para ser un miembro bien informado de la sociedad.

La idea básica

La idea básica que yace en todos los métodos estadísticos de análisis de datos es inferir respecto de una población por medio del estudio de una muestra relativamente pequeña elegida de

ésta. Como ejemplo, considere una máquina que hace varillas de acero para su uso en dispositivos ópticos de almacenamiento. La especificación del diámetro de las varillas es 0.45 ± 0.02 cm. En la última hora, la máquina ha hecho mil varillas. El ingeniero que supervisa la calidad quiere saber cuántas de estas varillas satisfacen la especificación. No tiene tiempo para medir todas. En este contexto, toma una muestra aleatoria de 50 varillas, las mide y encuentra que 46 de éstas (92%) satisfacen la especificación del diámetro. De acuerdo con lo anterior, no es probable que la muestra de 50 varillas represente perfectamente a la población de mil. La proporción de buenas varillas en la población probablemente es un poco diferente que la proporción de la muestra de 92%. En este sentido, lo que el ingeniero debe conocer es la probabilidad de que esa diferencia sea grande. Por ejemplo, ¿es admisible que los porcentajes de población sean superiores a 95%? ¿y de 98%? ¿O menores de 90%?, ¿o de 85%?

He aquí algunas preguntas específicas que el ingeniero podría responder con base en los datos de la muestra:

1. El ingeniero necesita calcular la magnitud de la diferencia probable entre las proporciones de la muestra y de la población. ¿Qué tan grande es una diferencia típica para esta clase de muestra?
2. Asimismo, necesita llevar una bitácora con los porcentajes de varillas aceptables fabricadas en la última hora. Después de que ha observado que 92% de las varillas de la muestra estaba bien, indicará los porcentajes de las varillas aceptables en la población como un intervalo de la forma $92\% \pm x\%$, donde x es un número calculado para tener una confianza razonable de que los porcentajes reales de la población están en este intervalo. ¿Cómo se debe calcular x ?
3. Por último, quiere estar muy seguro de que el porcentaje de varillas buenas es de al menos 90%; en otro caso detendrá el proceso para recalibrarlo. ¿Qué seguridad puede tener de que al menos 90% de las mil varillas está bien?

Gran parte de este libro está dedicada a solucionar preguntas semejantes. La primera de éstas requiere del cálculo de una **desviación estándar**, que se analizará en los capítulos 2 y 4. La segunda pregunta requiere de la construcción de un **intervalo de confianza**, ello se aprenderá en el capítulo 5. La tercera invoca una **prueba de hipótesis**, que se estudiará en el capítulo 6.

Los capítulos restantes del libro cubren otros temas importantes. Por ejemplo, el ingeniero de nuestro ejemplo querrá saber cómo está relacionada la fuerza de tensión con la cantidad de carbono en las varillas de acero. Esta clase de problemas se puede tratar con los métodos de **correlación y regresión**, que se presentan en los capítulos 7 y 8. Podría también ser importante determinar cómo ajustar el proceso de fabricación respecto de algunos factores, con el fin de producir resultados óptimos. Esto último requiere del diseño de **experimentos factoriales**, que se analizarán en el capítulo 9. Definitivamente, el ingeniero necesitará desarrollar un plan para controlar la calidad del producto que se fabrica en el proceso. En el capítulo 10 se presenta el tema **control de la calidad**, donde los métodos estadísticos se usan para mantener la calidad en un contexto industrial.

Los temas que se han mencionado son métodos que se dedican a obtener conclusiones a partir de datos. Estos métodos constituyen el campo de la **estadística inferencial**. Antes de que se analicen estos temas, se aprenderá más acerca de los métodos de recopilación de datos y a resumir claramente la información básica que contienen. Éstos son los temas de **muestreo y estadística descriptiva**, que se tratan en lo que resta de este capítulo.

1.1 Muestreo

Como se ha mencionado, los métodos estadísticos están basados en la idea de analizar una **muestra** tomada de una **población**. Para trabajar con esta idea, la muestra se debe elegir de manera adecuada. Por ejemplo, digamos que se quiere conocer la estatura de los estudiantes de la Escuela de Minas, de Colorado, al considerar una muestra de 100 estudiantes. ¿Cómo se deben elegir los 100 estudiantes que se medirían? Algunos métodos son malos. Por ejemplo, elegir a los estudiantes de las listas del fútbol americano y de los equipos de basquetbol daría como resultado una muestra que indudablemente no representaría la distribución de estatura de la población de estudiantes. Usted podría pensar que sería razonable usar alguna muestra convenientemente obtenida; por ejemplo, todos los estudiantes que viven en cierta área o todos aquellos que se inscribieron en el curso de estadística para la ingeniería. Después de todo, no hay razón para pensar que la estatura de estos estudiantes debiera ser diferente de la estatura, en general, de los estudiantes. Sin embargo, muestras así no son ideales, porque pueden volverse engañosas en formas no previstas. Los mejores métodos del muestreo implican el **muestreo aleatorio**. Hay muchos métodos diferentes del muestreo aleatorio, el básico es el **muestreo aleatorio simple**.

Para entender la naturaleza de una muestra aleatoria simple, piense en una lotería. Imagine que se han vendido diez mil billetes y que se eligen cinco ganadores. ¿Cuál es la manera más justa de elegir a los ganadores? Es colocar todos los boletos en un recipiente, mezclarlos y extraer cinco de ellos uno tras otro. Los boletos premiados constituyen una muestra aleatoria simple de la población de diez mil billetes de la lotería. Cada boleto es igualmente probable de ser uno de los cinco boletos extraídos. Es importante indicar que cada conjunto de cinco boletos que se puede formar del total tiene la misma probabilidad de ser el grupo que se extrae. Esta idea constituye la base de la definición de una muestra aleatoria simple.

Resumen

- Una **población** representa la colección completa de elementos o resultados de la información buscada.
- Una **muestra** constituye un subconjunto de una población, que contiene elementos o resultados que realmente se observan.
- Una **muestra aleatoria simple** de tamaño n es una muestra elegida por un método en el que cada colección de n elementos de la población tiene la misma probabilidad de formar la muestra, de la misma manera que en una lotería.

Debido a que una muestra aleatoria simple es similar a una lotería, con frecuencia se puede tomar la muestra con el mismo método que el que se usa en muchas loterías: con un generador de números aleatorios de una computadora. Suponga que hay N elementos en la población y que se le asigna a cada elemento de la población un entero entre 1 y N . Después se genera una lista de enteros aleatorios entre 1 y N y se eligen los elementos correspondientes de la población para que formen la muestra aleatoria simple, precisamente como en la lotería.

Ejemplo

1.1

Una maestra de educación física quiere estudiar los niveles de condición física de los estudiantes en su universidad. Hay 20 000 estudiantes inscritos y desea tomar una muestra de tamaño 100 para hacerles una prueba de sus condiciones físicas. Obtiene una lista de todos los estudiantes, numerada del 1 al 20 000. Usa un generador de números aleatorios de la computadora que genera 100 enteros aleatorios entre el total de números y después invita a los 100 estudiantes, a quienes corresponden dichos números, a que participen en el estudio. ¿Ésta es una muestra aleatoria simple?

Solución

Sí, ésta es una muestra aleatoria simple. Observe que es similar a una lotería en la que cada estudiante tiene un boleto y se sacan 100 de éstos.

Ejemplo

1.2

Una ingeniero que supervisa la calidad quiere inspeccionar rollos de papel tapiz para obtener información acerca de la tasa de fallas que tiene la imprenta. Decide tomar una muestra de 50 rollos de la producción de un día. Cada hora durante cinco horas, toma los diez últimos rollos producidos y cuenta el número de fallas de cada uno. ¿Ésta es una muestra aleatoria simple?

Solución

No. No todo subconjunto de 50 rollos de papel tapiz tiene la misma probabilidad de pertenecer a la muestra. Para formar una muestra aleatoria simple, la ingeniero necesitaría asignar un número a cada rollo producido durante el día y después generar números aleatorios para determinar con qué rollos se forma la muestra.

En algunos casos, es difícil o imposible extraer una muestra de una manera realmente aleatoria. En esta situación, lo mejor que se puede hacer es seleccionar los elementos de la muestra por algún método conveniente. Por ejemplo, imagine que un ingeniero civil acaba de recibir una remesa de mil bloques de hormigón, que pesan aproximadamente 50 libras cada uno. Los bloques se han entregado en una gran pila. El ingeniero quiere investigar la fuerza de compresión de los bloques midiendo las fuerzas en una muestra de diez bloques. Para tomar una muestra aleatoria simple se requeriría sacar bloques del centro y de la parte inferior de la pila, lo que puede ser muy difícil. Por esta razón, el ingeniero puede tomar una muestra simplemente tomando diez bloques de la parte superior de la pila. Una muestra así se llama **muestra de conveniencia**.

Definición

Una **muestra de conveniencia** es una muestra que no se extrae por un método aleatorio bien definido.

El problema con las muestras de conveniencia es que podrían diferir sistemáticamente de la población en alguna forma. Por esta razón, tales muestras no se deben usar, excepto en situaciones donde no es viable tomar una muestra aleatoria. Cuando se necesita tomar una

muestra de conveniencia, es importante pensar en todas las formas en las que aquélla podría diferir sistemáticamente de la población. Si es razonable pensar que no existe una diferencia sistémica importante, entonces puede ser aceptable tratar la muestra de conveniencia como si fuera una muestra aleatoria simple. Respecto de los bloques de hormigón, si el ingeniero está seguro de que los bloques superiores en la pila no difieren sistemáticamente en alguna forma importante del resto, entonces puede tratar la muestra de conveniencia como una muestra aleatoria simple. Sin embargo, si es posible que los bloques en diferentes lugares de la pila hayan sido hechos con diferentes cantidades de mezclas o que puedan tener diferentes tiempos de cocido o diferentes temperaturas, entonces una muestra de conveniencia podría dar resultados falsos.

Algunas personas piensan que una muestra aleatoria simple es garantía de que refleja perfectamente a su población. Esto no es cierto. Las muestras aleatorias simples siempre son diferentes de sus poblaciones en algunos aspectos y en ocasiones podrían ser considerablemente diferentes. Dos muestras diferentes de la misma población también serán diferentes entre sí. Este fenómeno se conoce como **variación del muestreo**. Esta última constituye una de las razones por la que los experimentos científicos tienen resultados diferentes cuando se repiten, aun cuando las condiciones parecen ser idénticas.

Ejemplo 1.3

Un inspector de calidad prueba 40 pernos de una gran remesa y mide la longitud de cada uno. Descubre que 34 de ellos (85%) cubre la especificación de longitud. Llega entonces a la conclusión de que exactamente 85% de los pernos de la remesa satisfacen la especificación. Por otra parte, el supervisor del inspector concluye que la proporción de pernos buenos está cerca de 85% con cierta probabilidad, pero que no es exactamente igual. ¿Cuál es la conclusión correcta?

Solución

Debido a la variación del muestreo, las muestras aleatorias simples no reflejan a la población perfectamente. Sin embargo, con frecuencia están bastante cerca. Por tanto, resulta adecuado inferir que la proporción de pernos buenos en la remesa esté cerca de la proporción de muestra, que es de 85%, con cierta probabilidad. Sin embargo, no es probable que la proporción de población sea igual a 85 por ciento.

Ejemplo 1.4

Continuando con el ejemplo 1.3, otra inspectora repite el estudio con una muestra aleatoria simple diferente de 40 pernos. Descubre que 36 de ellos, 90%, son buenos. El primer inspector afirma que ella debió haber cometido algún error, ya que sus resultados mostraban que 85% y no 90% de los pernos son buenos. ¿Tiene razón?

Solución

No, él no tiene razón. Es la variación del muestreo en acción. Dos muestras diferentes de la misma población serán diferentes entre sí y de la población.

Ya que las muestras aleatorias simples no reflejan a sus poblaciones perfectamente, ¿por qué es importante que el muestreo sea aleatorio? La ventaja de una muestra aleatoria simple es que no hay ningún mecanismo sistémico que la haga poco representativa. Las dife-

rencias entre la muestra y su población son atribuibles completamente a la variación aleatoria. Debido a que la teoría matemática sobre la variación aleatoria se comprende bien, se pueden usar modelos matemáticos para estudiar la relación entre muestras aleatorias simples y sus poblaciones. En general, para una muestra que no fue seleccionada de forma aleatoria, no existe una teoría disponible que describa los mecanismos que causaron que la muestra difiera de su población. Por tanto, con frecuencia las muestras que no fueron obtenidas aleatoriamente son difíciles de analizar de manera confiable.

En los ejemplos 1.1 a 1.4, las poblaciones constaban de elementos físicos reales: estudiantes de una universidad, bloques de concreto de una pila, pernos de una remesa. Estas poblaciones se denominan **poblaciones tangibles**. Este tipo de poblaciones son siempre finitas. Después de que se muestrea un elemento, el tamaño de población disminuye en 1. En principio, uno podría en algunos casos regresar el elemento muestreado a la población, con oportunidad de muestrearlo nuevamente, pero esto rara vez se hace en la práctica.

En ingeniería es frecuente que los datos sean producto de mediciones realizadas durante un experimento científico, más que por muestreo de una población tangible. Tomando un ejemplo simple, imagine que un ingeniero mide la longitud de una varilla cinco veces, haciendo las mediciones en la forma más cuidadosa posible con condiciones idénticas. No importa qué tan cuidadosamente se hayan hecho las mediciones, diferirán un poco una de otra, debido a la variación en el proceso de medición que no se puede controlar o predecir. Esto último da como resultado que con frecuencia sea adecuado considerar estos datos como una muestra aleatoria simple de una población. En estos casos, la población consta de todos los valores que posiblemente pueden haber sido observados. Esta población se denomina **población conceptual**, ya que no consta de elementos reales.

Una muestra aleatoria simple puede consistir de valores obtenidos en un proceso en condiciones experimentales idénticas. En este caso, la muestra proviene de una población que consta de todos los valores posibles que se han observado. A este tipo de población se le denomina **población conceptual**.

El ejemplo 1.5 implica una población conceptual.

Ejemplo

1.5

Un geólogo pesa una roca varias veces en una balanza analítica. Cada vez, la balanza da una lectura ligeramente diferente. ¿Bajo qué condiciones se pueden considerar estas lecturas como una muestra aleatoria simple? ¿Cuál es la población?

Solución

Si las características físicas de la balanza permanecen iguales cada vez que se pesa, se puede considerar que las mediciones se hacen bajo condiciones idénticas, entonces las lecturas se pueden considerar como una muestra aleatoria simple. La población es conceptual. Consta de todas las lecturas que la balanza en principio podría producir.

Observe que en el ejemplo 1.5, son las características físicas del proceso de medición las que determinan si los datos constituyen una muestra aleatoria simple. En general, cuando se decide si un conjunto de datos se puede considerar una muestra aleatoria simple, es muy

útil tener una comprensión del proceso que generó los datos. Algunas veces los métodos estadísticos pueden ayudar, especialmente cuando la muestra es grande, pero el conocimiento del mecanismo que produjo los datos es más importante.

Ejemplo

1.6

Se ha diseñado un nuevo proceso químico que se supone tendrá una producción más alta de cierta sustancia química que durante el proceso anterior. Para investigar los resultados de este proceso, lo realizamos 50 veces y registramos los 50 resultados. ¿Bajo qué condiciones sería razonable considerar lo anterior como una muestra aleatoria simple? Describa algunas condiciones bajo las cuales puede no resultar adecuado considerar esto último como una muestra aleatoria simple.

Solución

Para responder a esto, primero debemos especificar la población. La población es conceptual y consta del conjunto de todos los resultados que se obtienen de este proceso, así como de las veces que se realizó. Lo que hemos llevado a cabo es un muestreo de los 50 primeros resultados del proceso. Si y sólo si estamos seguros de que los primeros 50 resultados se han generado en condiciones idénticas y que no difieren en ninguna forma sistémica de los resultados de futuras realizaciones, podremos tratarlos como una muestra aleatoria simple.

Sin embargo, sea cauteloso. Hay muchas condiciones por las que 50 resultados podrían dejar de ser una muestra aleatoria simple. Por ejemplo, con procesos químicos, algunas veces se da el caso de que realizaciones con resultados altos son seguidas de realizaciones con resultados bajos y viceversa. A veces los resultados tienden a aumentar con el tiempo, conforme los ingenieros de proceso aprenden por la experiencia cómo hacer funcionar el proceso de manera más eficiente. En estos casos, los resultados no se han generado bajo las mismas condiciones y no constituyen una muestra aleatoria simple.

El ejemplo 1.6 muestra nuevamente que un buen conocimiento de la naturaleza del proceso en estudio es importante para determinar si los datos se pueden considerar como muestra aleatoria simple. Los métodos estadísticos algunas veces se usan para mostrar que un conjunto de datos dado no representa necesariamente una muestra aleatoria simple. Por ejemplo, a veces las condiciones experimentales cambian gradualmente con el tiempo. Un método simple, pero efectivo para detectar esta condición, es realizar una gráfica con las observaciones en el orden en que se tomaron. Una muestra aleatoria simple no debe mostrar ningún patrón o tendencia obvia.

La figura 1.1 presenta las gráficas de tres muestras en el orden en que se tomaron. La gráfica de la figura 1.1a muestra un patrón oscilatorio. La gráfica en la figura 1.1b muestra una tendencia creciente. Ninguna de estas muestras se debe tratar como muestra aleatoria simple. La gráfica en la figura 1.1c no parece mostrar ningún patrón o tendencia obvia. Podría ser apropiado tratar estos datos como una muestra aleatoria simple. Sin embargo, antes de tomar esa decisión, es aún importante pensar acerca del proceso que produjo estos datos, ya que puede haber cuestiones que no son evidentes en la gráfica (véase el ejemplo 1.7).

A veces la pregunta respecto de si un conjunto de datos es una muestra aleatoria simple, depende de la población en estudio. Se puede dar el caso para el cual una gráfica pueda parecer buena, aun cuando los datos no sean una muestra aleatoria simple. En el ejemplo 1.7 se da un caso.

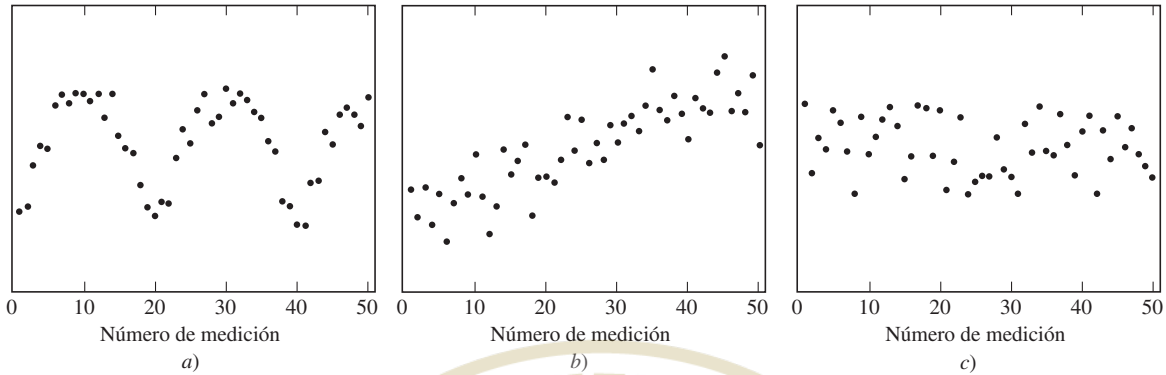


FIGURA 1.1 Tres gráficas de valores observados contra el orden en que se hicieron. *a)* Los valores indican un patrón definido en el tiempo. Esta no es una muestra aleatoria simple. *b)* Los valores muestran una tendencia en el tiempo. Esta no es una muestra aleatoria simple. *c)* Los valores no muestran un patrón o tendencia. Puede ser adecuado tratar estos datos como una muestra aleatoria simple.

Ejemplo

1.7

Un nuevo proceso químico se realiza diez veces cada mañana durante cinco días consecutivos. Una gráfica de los resultados en el orden en que aparecieron no presenta ningún patrón o tendencia obvia. Si el nuevo proceso se pone en producción, haciéndolo funcionar diez horas todos los días, desde las 7 a.m. hasta las 5 p.m. ¿Es razonable considerar que los 50 resultados sean una muestra aleatoria simple? ¿Qué ocurre si el proceso está siempre funcionando por la mañana?

Solución

Debido a que se intenta poner en funcionamiento el nuevo proceso tanto durante la mañana como en la tarde, la población consta de todos los resultados que alguna vez se observarán, incluyendo tanto las realizaciones por la mañana como por la tarde. La muestra se toma sólo de la parte de la población de los resultados matutinos; por tanto, no es una muestra aleatoria simple. Hay muchas cosas que podrían estar equivocadas si esto se usa como una muestra aleatoria simple. Por ejemplo, las temperaturas ambientales pueden ser diferentes entre la mañana y la tarde, ello podría afectar los resultados.

Si el proceso funcionara sólo por la mañana, entonces la población constaría sólo de resultados matutinos. Debido a que la muestra no presenta ningún patrón o tendencia obvia, bien podría ser apropiado considerarla como muestra aleatoria simple.

Independencia

Se dice que los elementos en una muestra son **independientes** si al conocer los valores de algunos de ellos no ayuda a predecir los valores de los otros. Con una población finita y tangible, los elementos en una muestra aleatoria simple no son estrictamente independientes, ya que cuando se extrae cada elemento, la población cambia. Este cambio puede ser importante cuando la población es pequeña. Sin embargo, cuando la población es muy grande, este cambio resulta insignificante y los elementos se pueden tratar como si fueran independientes.

Para ilustrar esta idea, imagine que se toma una muestra aleatoria simple de dos elementos de la población



Para la primera extracción, los números 0 y 1 son igualmente probables. Pero el valor del segundo elemento está evidentemente influido por el primero; si el primero es 0, es más probable que el segundo sea 1 y viceversa. Por tanto, los elementos de la muestra son dependientes. Ahora suponga que sacamos una muestra de tamaño 2 de esta población:



Nuevamente en la primera extracción, los números 0 y 1 son igualmente probables. Pero a diferencia del ejemplo anterior, también estos dos valores permanecen casi de la misma manera en la segunda extracción, sin que importe lo que sucede en la primera extracción. Con poblaciones grandes, los elementos de la muestra son para todos los propósitos prácticos independientes.

Es razonable preguntarse qué tan grande debe ser una población para que los elementos en una muestra aleatoria simple se traten como independientes. Una regla general señala que cuando se toma una muestra de una población finita, los elementos se pueden tratar como independientes en tanto la muestra consista de 5% o menos de la población.

Curiosamente, es posible hacer que una población se comporte como si fuera infinitamente grande, reemplazando cada elemento después de que se ha muestreado. Este método se denomina **muestreo con reemplazo**. Con este método la población es exactamente la misma en cada extracción y los elementos muestreados son realmente independientes.

Con una población conceptual, se requiere que los elementos de la muestra se produzcan en condiciones experimentales idénticas. En particular, ningún valor de muestra puede influir en las condiciones bajo las cuales se producen los otros. Por tanto, los elementos en una muestra aleatoria simple de una población conceptual se pueden tratar como independientes. Podemos pensar que una población conceptual es infinita, o de manera equivalente que los elementos se muestrean con reemplazo.

Resumen

- Los elementos en una muestra son **independientes** si el conocimiento de algunos de los valores de los elementos no ayuda a predecir los valores de los otros.
- Los elementos en una muestra aleatoria simple se pueden tratar como independientes en muchos casos que se encuentran en la práctica. Ocurre una excepción cuando la población es finita y la muestra consiste de una parte importante (más de 5%) de la población.

Otros métodos de muestreo

Además del muestreo aleatorio simple, existen otros métodos de muestreo que son útiles en diversas situaciones. En el **muestreo ponderado** a algunos elementos se les da una mayor oportunidad que a los otros para ser seleccionados, de la misma manera que en una lotería en la que algunas personas tienen más boletos que otros. En el **muestreo aleatorio estratificado**, la población se divide en subpoblaciones, llamadas **estratos** y se extrae una muestra aleatoria simple de cada estrato. En el **muestreo agrupado**, los elementos se extraen de la población en grupos o conglomerados. El muestreo agrupado es útil cuando la población es demasiado grande y se encuentra extendida de tal forma que es posible tomar una muestra aleatoria simple. Por ejemplo, muchos de los organismos del gobierno estadounidense usan muestreo agrupado para muestrear a la población de los Estados Unidos para medir factores sociológicos, como ingresos y número de desempleados. Una buena fuente de información acerca de métodos de muestreo es Cochran (1977).

El muestreo aleatorio simple no es el único método válido de muestreo aleatorio. Pero es el más importante y se le prestará la mayor parte de la atención. Por el momento, a menos que se indique otra cosa, se considerará que los términos “muestra” y “muestreo aleatorio” significan “muestra aleatoria simple”.

Tipos de experimentos

Hay muchas clases de experimentos que se pueden usar para generar datos. Describiremos brevemente algunos de ellos. En un experimento de **una-muestra**, hay sólo una población de interés y se extrae únicamente una muestra de ésta. Por ejemplo, imagine que se ha diseñado un proceso para producir polietileno que se usará para hacer tubos. En este contexto, un experimento mediante el cual se producen algunas muestras de polietileno y se mide la fuerza de tensión de cada una constituye un experimento de una-muestra. Se considera que las fuerzas medidas representan una muestra aleatoria simple de una población conceptual de todas las fuerzas posibles que se pueden observar en las muestras fabricadas por este proceso. Los experimentos del tipo una-muestra se pueden usar para determinar si un proceso satisface cierta norma; por ejemplo, si tienen la fuerza suficiente para una aplicación dada.

En un experimento de **muestras-múltiples**, hay dos o más poblaciones de interés y se toma una muestra de cada población. Por ejemplo, si están compitiendo algunos procesos para ser considerados en la fabricación de polietileno y se miden las fuerzas de tensión en una muestra de los elementos de cada proceso, se entiende que éste es un experimento de muestras-múltiples. A cada proceso le corresponde una población distinta y a las mediciones hechas sobre los elementos de un proceso dado se les considera una muestra aleatoria simple de esa población. El propósito habitual de los experimentos de muestras-múltiples es hacer comparaciones entre las poblaciones. En este ejemplo, el propósito podría ser que se determine el proceso que produce la mayor fuerza o que se determine si hay alguna diferencia en las fuerzas en el polietileno que se produjo mediante los diferentes procesos.

En muchos experimentos de muestras-múltiples, las poblaciones se distinguen entre sí al cambiar uno o más **factores** que pueden afectar el resultado. A estos experimentos se les llama **experimentos factoriales**. Por ejemplo, G. Fredrickson, en su tesis de maestría en la Escuela de Minas, de Colorado, midió la dureza ante el impacto de la muesca Charpy V para un importante número de soldaduras. Cada soldadura estaba hecha de uno de dos tipos de me-

tales base y se había medido su dureza a diferentes temperaturas. Éste fue un experimento factorial con dos factores: el metal base y la temperatura. Los datos consistían de varias mediciones de la dureza hechas con combinaciones del metal base y la temperatura. En un experimento factorial, cada combinación de los factores para los cuales se recopilan datos define una población y se extrae una muestra aleatoria simple de cada población. El propósito de un experimento factorial es determinar cómo afecta el resultado al cambiar los niveles de los factores. En su experimento, Fredrickson encontró que para cada tipo de metal base, la dureza no es afectada por la temperatura a menos que esta última estuviese en un nivel muy bajo, debajo de -100°C . Conforme la temperatura se reducía de -100°C a -200°C , la dureza bajaba uniformemente.

Tipos de datos

Cuando se asigna una cantidad numérica a cada elemento de una muestra, al conjunto de valores resultante se le llama **numérico** o **cuantitativo**. En algunos casos, los elementos de la muestra son puestos en categorías. Entonces los datos son **categoricos** o **cualitativos**. En el ejemplo 1.8 se presenta un caso.

Ejemplo

1.8

En el artículo “Hysteresis Behavior of CFT Column to H-Beam Connections with External T-Stiffeners and Penetrated Elements” (C. Kang, K. Shin y colaboradores, *Engineering Structures*, 2001:1194-1201) se reportaron los resultados de las pruebas de carga cíclicas en una columna tubular llena de concreto (CFT) para conexiones soldadas de vigas-H. Se cargaron algunos especímenes de prueba hasta que fallaron. Algunas fallas ocurrieron en la unión soldada; otras ocurrieron al doblarse la viga misma. Para cada muestra se registró la posición de la falla, junto con el par de torsión aplicado en la falla [en kilonewton-metros ($\text{kN} \cdot \text{m}$)]. Los resultados para las primeras cinco muestras fueron los siguientes:

Muestra	Par de torsión ($\text{kN} \cdot \text{m}$)	Posición de la falla
1	165	Soldadura
2	237	Viga
3	222	Viga
4	255	Viga
5	194	Soldadura

¿Qué datos son numéricos y cuáles categoricos?

Solución

Los pares de torsión, en la columna de en medio, son datos numéricos. Las posiciones de la falla, en la columna de la derecha, son datos categoricos.

Ejercicios para la sección 1.1

1. Cada uno de los siguientes procesos implica el muestreo de una población. Defina la población y diga si es tangible o conceptual.
 - a) Se recibe una remesa de pernos de un distribuidor. Para verificar si la remesa es aceptable respecto de la fuerza de corte, un ingeniero selecciona diez pernos, uno tras otro, del recipiente para probarlos.
 - b) La resistencia de cierto resistor se mide cinco veces con el mismo óhmetro.
 - c) Un estudiante de posgrado que se especializa en ciencia ambiental forma parte de un equipo de estudio que está evaluando el riesgo para la salud humana de cierto contaminante presente en el agua de la llave en su pueblo. Una parte del proceso de evaluación implica calcular la cantidad de tiempo que las personas que viven en ese pueblo están en contacto con el agua de la llave. El estudiante convence a los residentes del pueblo para que lleven una agenda mensual, detallando la cantidad de tiempo que están en contacto con el agua de la llave día con día.
 - d) Se hacen ocho soldaduras con el mismo proceso y se mide la fuerza en cada una.
 - e) Un ingeniero responsable del control de calidad tiene que calcular el porcentaje de piezas fabricadas defectuosas en determinado día. A las 2:30 de la tarde muestrea las últimas 100 piezas fabricadas.
2. Si usted quisiera calcular la altura media de todos los estudiantes en una universidad, ¿cuál de las siguientes estrategias de muestreo sería la mejor? ¿Por qué? Observe que ninguno de los métodos son realmente muestras aleatorias simples.
 - i) Medir la estatura de 50 estudiantes que se encuentran en el gimnasio durante el juego de basquetbol en la escuela.
 - ii) Medir la estatura de todos los especialistas en ingeniería.
 - iii) Medir la estatura de los estudiantes, eligiendo el primer nombre de cada página de la guía telefónica del campus universitario.
3. Verdadero o falso:
 - a) Una muestra aleatoria simple garantiza que refleja exactamente a la población de la que se extrajo.
 - b) Una muestra aleatoria simple está libre de cualquier tendencia sistémica en diferir de la población de la que se extrajo.
4. Una ingeniera de control de calidad extrae una muestra aleatoria simple de 50 anillos-O de un lote de varios miles. Mide el espesor de cada uno y descubre que 45 de ellos, 90%, cumple con cierta especificación. ¿Cuál de los siguientes enunciados es correcto?
 - i) La proporción de anillos-O en el lote completo que cumple con la especificación probablemente es igual a 90 por ciento.
 - ii) La proporción de anillos-O en el lote completo que cumple con la especificación probablemente está cerca de representar 90%, pero probablemente no es igual al total.
5. Se ha usado durante mucho tiempo un proceso para la fabricación de botellas de plástico y se sabe que 10% de éstas se encuentra defectuoso. Se está probando un nuevo proceso que, se supone, reduce la proporción de defectos. En una muestra aleatoria simple de 100 botellas producidas con el nuevo proceso, diez estaban defectuosas.
 - a) Uno de los ingenieros sugiere que la prueba demuestra que el nuevo proceso no es mejor que el proceso anterior, ya que la proporción de defectos es la misma. ¿Es ésta una conclusión justificada? Explique.
 - b) Suponga que hubieran sido solamente nueve las botellas defectuosas de la muestra de 100. ¿Esto habría probado que el nuevo proceso es mejor? Explique.
 - c) ¿Qué resultado presenta pruebas más evidentes de que el nuevo proceso es mejor: encontrar nueve botellas defectuosas en la muestra o encontrar dos botellas defectuosas en la muestra?
6. Con referencia al ejercicio 5. Verdadero o falso:
 - a) Si la proporción de defectos en la muestra es menor a 10%, es confiable concluir que el nuevo proceso es mejor.
 - b) Si la proporción de defectos en la muestra es sólo ligeramente menor a 10%, la diferencia bien podría ser completamente atribuible a la variación del muestreo y no es confiable concluir que el nuevo proceso es mejor.
 - c) Si la proporción de defectos en la muestra es mucho menor a 10%, es muy poco probable que la diferencia sea atribuible completamente a la variación del muestreo, por lo que es confiable llegar a la conclusión de que el nuevo proceso es mejor.
 - d) No importa qué tan pocos defectos aparezcan en la muestra, el resultado bien podría ser completamente atribuible a la variación del muestreo, por lo que no es confiable concluir que el nuevo proceso es mejor.
7. Para determinar si una muestra se debe tratar como una muestra aleatoria simple, ¿qué es más importante: un buen conocimiento de la estadística o un buen conocimiento del proceso que produce los datos?

1.2 Resumen estadístico

Con frecuencia una muestra constituye una larga lista de números. Para ayudar a que las características de una muestra sean evidentes, se calcula el resumen estadístico. Las dos cantidades más usadas en el resumen estadístico son la **media de la muestra** y la **desviación estándar de la muestra**. La primera indica el centro de los datos y la segunda señala cómo están distribuidos los datos.

Media muestral

La media muestral también se llama “media aritmética”, o, simplemente, “promedio”. Representa la suma de los números en la muestra, dividido entre la cantidad total de números que hay.

Definición

Sea $X_1, \dots, X_n$ una muestra. La **media muestral** es

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1.1)$$

Observe que se acostumbra usar una letra con una barra encima de ésta (por ejemplo $\bar{X}$) para denotar la media de una muestra. También observe que la media muestral tiene las mismas unidades que los valores de la muestra $X_1, \dots, X_n$.

Ejemplo

1.9

Una muestra aleatoria simple de cinco hombres se elige de entre una gran población de hombres y se mide su estatura. Las cinco cifras de estatura (en pulgadas) son 65.51, 72.30, 68.31, 67.05 y 70.68. Encuentre la media muestral.

Solución

Usamos la ecuación (1.1). La media muestral es

$$\bar{X} = \frac{1}{5}(65.51 + 72.30 + 68.31 + 67.05 + 70.68) = 68.77 \text{ pulgadas.}$$

Desviación estándar

He aquí dos series de datos: 28, 29, 30, 31, 32 y 10, 20, 30, 40, 50. Ambas tienen la misma media de 30. Pero obviamente difieren en una manera importante que no es captada por la media: la segunda serie es mucho más dispersa que la primera. La **desviación estándar** es una cantidad que mide el grado de dispersión en una muestra.

Sea $X_1, \dots, X_n$ una muestra. La idea básica detrás de la desviación estándar es que cuando la dispersión es grande, los valores de la muestra tenderán a alejarse de su media, pero

cuando la dispersión es pequeña, los valores tenderán a acercarse a su media. En este contexto, el primer paso en el cálculo de la desviación estándar es calcular las distancias (también llamadas desviaciones) de cada valor de la muestra a la media de la muestra. Las desviaciones son $(X_1 - \bar{X}), \dots, (X_n - \bar{X})$. Ahora algunas de estas desviaciones son positivas y otras negativas. Las desviaciones grandes, tanto negativas como positivas, son indicadores de la dispersión. Para hacer todas las desviaciones positivas se elevan al cuadrado, con lo que se obtienen las desviaciones al cuadrado $(X_1 - \bar{X})^2, \dots, (X_n - \bar{X})^2$. A partir de las desviaciones al cuadrado se puede calcular una medida de la dispersión llamada la **varianza muestral**. Ésta constituye el promedio de las desviaciones al cuadrado, excepto que lo dividimos entre $n - 1$ en lugar de n . Se acostumbra denotar a la varianza muestral con s^2 .

Definición

Sea $X_1, \dots, X_n$ una muestra. La **varianza muestral** es la cantidad

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (1.2)$$

Una fórmula equivalente, que puede ser más fácil de calcular, es

$$s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) \quad (1.3)$$

Mientras que la varianza muestral es una cantidad importante, tiene una seria desventaja como una medida de la dispersión. Sus unidades no son las mismas que las unidades de los valores de la muestra; éstas tienen unidades al cuadrado. Para obtener una medida de la dispersión cuyas unidades sean las mismas que las de los valores de la muestra, simplemente se toma la raíz cuadrada de la varianza. Esta cantidad se denomina **desviación estándar muestral**. Se acostumbra denotar a la desviación estándar muestral por la letra s (la raíz cuadrada de s^2).

Definición

Sea $X_1, \dots, X_n$ una muestra. La **desviación estándar muestral** es la cantidad

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (1.4)$$

Una fórmula equivalente, que puede ser más fácil de calcular, es

$$s = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)} \quad (1.5)$$

La desviación estándar muestral constituye la raíz cuadrada de la varianza muestral.

Es natural preguntarse por qué la suma de las desviaciones al cuadrado se divide entre $n - 1$ en lugar de n . El propósito de calcular la desviación estándar muestral es calcular la cantidad de la dispersión en la población de la cual se extrajo aquélla. Por tanto, idealmente se calcularían las desviaciones de la media de todos los elementos de la población, en vez de las desviaciones de la media de la muestra. Sin embargo, la media de la población generalmente no se conoce, por lo que en su lugar se usa la media de la muestra. Es un hecho matemático que las desviaciones alrededor de la media muestral tienden a ser un poco más pequeñas que las desviaciones alrededor de la media poblacional y que al dividir entre $n - 1$ en vez de n proporciona la rectificación correcta.

Ejemplo

1.10

Encuentre la varianza muestral y la desviación estándar muestral para los datos de las estaturas del ejemplo 1.9.

Solución

Primero se calculará la varianza muestral usando la ecuación (1.2). La media muestral es $\bar{X} = 68.77$ (véase el ejemplo 1.9). La varianza muestral es, por tanto,

$$s^2 = \frac{1}{4}[(65.51 - 68.77)^2 + (72.30 - 68.77)^2 + (68.31 - 68.77)^2 + (67.05 - 68.77)^2 + (70.68 - 68.77)^2] = 7.47665$$

Por otra parte, se puede usar la ecuación (1.3):

$$s^2 = \frac{1}{4}[65.51^2 + 72.30^2 + 68.31^2 + 67.05^2 + 70.68^2 - 5(68.77^2)] = 7.47665$$

La desviación estándar muestral es la raíz cuadrada de la varianza muestral:

$$s = \sqrt{7.47665} = 2.73$$

¿Qué le pasaría a la media, a la varianza y a la desviación estándar muestrales si la estatura de los estudiantes en el ejemplo 1.9 fuera medida en centímetros en vez de pulgadas? Denotemos la estatura en pulgadas por X_1, X_2, X_3, X_4, X_5 y la estatura en centímetros por Y_1, Y_2, Y_3, Y_4, Y_5 . Entonces, la relación entre las X_i y las Y_i está dada por $Y_i = 2.54 X_i$. Si regresa al ejemplo 1.9, convierte a centímetros y calcula la media muestral, encontrará que las medias de la muestra en centímetros y en pulgadas están relacionadas por la ecuación $\bar{Y} = 2.54 \bar{X}$. Por consiguiente, si multiplicamos cada elemento de la muestra por una constante, la media de la muestra se multiplica por la misma constante. En cuanto a la varianza de la muestra, se encontrará que las desviaciones están relacionadas por la ecuación $(Y_i - \bar{Y}) = 2.54 (X_i - \bar{X})$. Por tanto, $s_Y^2 = 2.54^2 s_X^2$ y $s_Y = 2.54 s_X$.

¿Qué sucede si cada hombre en la muestra utilizó zapatos que elevaron su estatura dos pulgadas? Entonces cada estatura de la muestra aumenta dos pulgadas y la media de la muestra también en esa proporción. En general, si se agrega una constante a cada elemento de la muestra, la media de la muestra aumenta (o disminuye) en la misma constante. Sin embargo, las desviaciones no cambian, por lo que la varianza y la desviación estándar de la muestra no se ven afectadas.

Resumen

- Si $X_1, \dots, X_n$ es una muestra y $Y_i = a + bX_i$, donde a y b son constantes, entonces $\bar{Y} = a + b\bar{X}$.
- Si $X_1, \dots, X_n$ es una muestra y $Y_i = a + bX_i$, donde a y b son constantes, entonces $s_Y^2 = b^2 s_X^2$, y $s_Y = |b|s_X$.

Datos atípicos

A veces una muestra puede contener algunos puntos que son mucho más grandes o pequeños que el resto. Estos puntos se llaman **datos atípicos**. Véase la figura 1.2 como ejemplo. A veces los datos atípicos son resultado de errores al ingresar datos; por ejemplo, un punto decimal perdido puede dar como resultado un valor que es de un orden de magnitud diferente del resto. Los datos atípicos se deben **examinar siempre** y cualquiera de ellos que se encuentre es resultado de un error que **se debe corregir o eliminar**. Aunque no todos los datos atípicos son errores. A veces una población podría contener algunos valores que son muy diferentes del resto y los datos atípicos en la muestra reflejan este hecho.



FIGURA 1.2 Conjunto de datos que contiene un dato atípico.

Los datos atípicos representan un verdadero problema para los analistas de datos. Como consecuencia de lo anterior, **cuando las personas ven datos atípicos en sus datos**, tratan de encontrar una razón o un pretexto para eliminarlos. Sin embargo, un dato atípico no se debe eliminar, a menos que **se tenga la seguridad de que es resultado de un error**. Cabe señalar que, si una población realmente **contiene datos atípicos** y son eliminados de la muestra, esta última no caracterizará correctamente a la población.

Mediana muestral

La **mediana**, al igual que la media, representa una medida de tendencia central de los datos. Para calcular la mediana de **una muestra**, **ordene los valores del más pequeño al más grande**. La mediana es el número de en medio. Si el tamaño de la muestra es un número par, se acostumbra tomar a la mediana muestral como el promedio de los dos números de en medio.

Definición

Si n números están ordenados del más pequeño al más grande:

- Si n es impar, la mediana muestral es el número en la posición $\frac{n + 1}{2}$.
- Si n es par, la mediana muestral representa el promedio de los números en las posiciones $\frac{n}{2}$ y $\frac{n}{2} + 1$.

Ejemplo

1.11

Encuentre la mediana muestral para los datos de la estatura en el ejemplo 1.9.

Solución

Las cifras de los cinco casos de estatura, en orden creciente, son 65.51, 67.05, 68.31, 70.68, 72.30. La mediana muestral es el número de en medio, que es 68.31.

La mediana se usa con frecuencia como una medida de tendencia central para muestras que contienen datos atípicos. Con el propósito de saber por qué, considere que la muestra consta de los valores 1, 2, 3, 4 y 20. La media es 6 y la mediana es 3. Es razonable pensar que la mediana es más representativa de la muestra que la media. Véase la figura 1.3.



FIGURA 1.3 Cuando una muestra contiene datos atípicos, la mediana podría ser más representativa de la muestra que la media.

La media recortada

De la misma manera que la mediana, la **media recortada** es una medida de tendencia central que se diseñó para que no esté afectada por datos atípicos. La media recortada se calcula al arreglar los valores de la muestra en orden, “recortar” un número igual a partir de cada extremo y calcular la media de los restantes. Si se “recorta” el $p\%$ de los datos de cada extremo, la media recortada resultante se denomina “media recortada un $p\%$ ”. No existe fórmula ni fácil ni difícil para saber cuántos valores se deben recortar. Las más comunes son las medias recortadas al 5, 10 y 20%. Observe que la mediana se puede pensar como una forma extrema de la media recortada, obtenida de recortar todo, excepto uno o dos valores de en medio de la muestra.

Debido a que el número de puntos de datos recortados debe ser un número entero, en muchos casos es imposible recortar los porcentajes exactos que se piden de los datos. Si el tamaño muestral se denota por n y se desea recortar un $p\%$, el número de datos a ser recortados es $np/100$. Si éste no es un número entero, lo más sencillo que se debe hacer cuando se calcula manualmente es redondear al número entero más cercano y recortar esa cantidad.

Ejemplo

1.12

En el artículo “Evaluation of Low-Temperature Properties of HMA Mixtures” (P. Sebaaly, A. Lake y J. Epps, en *Journal of Transportation Engineering*, 2002:578-583) se midieron los siguientes valores de la tensión de fractura (en megapascasles) para una muestra de 24 mezclas de asfalto mezclado caliente (HMA).

30	75	79	80	80	105	126	138	149	179	179	191
223	232	232	236	240	242	245	247	254	274	384	470

Calcule la media, mediana y las medias recortadas a 5, 10 y 20 por ciento.

Solución

La media se encuentra promediando los 24 números, con base en ello se obtiene un valor de 195.42. La mediana es el promedio del 12avo. y 13avo. números, que es $(191 + 223)/2 = 207.00$. Para calcular la media recortada a 5%, debemos eliminar 5% de los datos de cada extremo. Se obtienen $(0.05)(24) = 1.2$ observaciones. Redondeamos 1.2 a 1 y recortamos una observación de cada extremo. La media recortada a 5% constituye el promedio de los 22 números restantes:

$$\frac{75 + 79 + \dots + 274 + 384}{22} = 190.45$$

Para calcular la media recortada a 10%, redondee $(0.1)(24) = 2.4$ a 2. Elimine dos observaciones de cada extremo y después determine el promedio de las 20 restantes:

$$\frac{79 + 80 + \dots + 254 + 274}{20} = 186.55$$

Para calcular la media recortada a 20%, redondee $(0.2)(24) = 4.8$ a 5. Elimine cinco observaciones de cada extremo y después determine el promedio de las 14 restantes:

$$\frac{105 + 126 + \dots + 242 + 245}{14} = 194.07$$

La moda y el rango

La **moda** y el **rango** son resúmenes estadísticos de uso limitado, pero que en ocasiones se aprecian visualmente. La moda muestral es el valor que tiene más frecuencia en una muestra. Si algunos valores tienen una frecuencia igual, cada uno representa una moda. El rango es la diferencia entre los valores más grandes y más pequeños en una muestra. Es una medida de la dispersión, pero rara vez se usa, porque depende solamente de los dos valores extremos y no proporciona ninguna información acerca del resto de la muestra.

Ejemplo

1.13

Encuentre las modas y el rango para la muestra del ejemplo 1.12.

Solución

Hay tres modas: 80, 179 y 232. Cada uno de estos valores aparece dos veces y ningún otro valor aparece más de una vez. El rango es $470 - 30 = 440$.

Cuartiles

La mediana divide la muestra a la mitad. Los **cuartiles** la dividen tanto como sea posible en cuartos. Una muestra tiene tres de aquéllos. Existen diferentes formas de calcular cuartiles, pero todas dan aproximadamente el mismo resultado. El método más simple cuando se calcula

manualmente es el siguiente: Sea n el tamaño de la muestra. Ordene los valores de la muestra del más pequeño al más grande. Para encontrar el primer cuartil, calcule el valor $0.25(n + 1)$. Si éste es un entero, entonces el valor de la muestra en esa posición es el primer cuartil. Si no, tome entonces el promedio de los valores de la muestra de cualquier lado de este valor. El tercer cuartil se calcula de la misma manera, excepto que se usa el valor $0.75(n + 1)$. El segundo cuartil usa el valor $0.5(n + 1)$. El segundo cuartil es idéntico a la mediana. Observe que algunos software usan métodos ligeramente diferentes para calcular cuartiles, por lo que sus resultados no pueden ser exactamente los mismos que los obtenidos por el método que aquí se describió.

Ejemplo

1.14

Encuentre el primer y tercer cuartil de los datos que se refieren al asfalto en el ejemplo 1.12.

Solución

El tamaño de la muestra es $n = 24$. Para encontrar el primer cuartil, calcule $(0.25)(25) = 6.25$. Por tanto, el primer cuartil se encuentra determinando el promedio del 6o. y 7o. puntos de datos, cuando la muestra se arregla en orden creciente. Se obtiene $(105 + 126)/2 = 115.5$. Para encontrar el tercer cuartil, calcule $(0.75)(25) = 18.75$. Promediamos los puntos de los datos 18avo. y 19avo., con lo que se obtiene $(242 + 245)/2 = 243.5$.

Percentiles

El p -ésimo percentil de una muestra, para un número p entre 0 y 100, divide a la muestra tanto como sea posible, el $p\%$ de los valores de la muestra es menor que el p -ésimo percentil y el $(100 - p)\%$ son mayores. Hay muchas maneras para calcular los percentiles; con todas se obtienen resultados similares. Aquí se describe un método similar al método descrito para calcular cuartiles. Ordene los valores de la muestra del más pequeño al más grande y después calcule la cantidad $(p/100)(n + 1)$, donde n es el tamaño de la muestra. Si esta cantidad es un entero, el valor de la muestra en esta posición es el p -ésimo percentil. Por otro lado, promedie los dos valores de la muestra en cualquier lado. Observe que el primer cuartil es el 25avo. percentil, la mediana es el 50avo. percentil y el tercer cuartil es el 75avo. percentil. Algunos software usan métodos ligeramente diferentes para el cálculo de los percentiles, así que sus resultados podrían ser un poco diferentes de los obtenidos por este método.

Los percentiles con frecuencia se usan para interpretar puntajes de exámenes estandarizados. Por ejemplo, si a una estudiante se le informa que su puntaje en un examen de ingreso a la universidad está en el 64avo. percentil, esto significa que 64% de los estudiantes que presentaron el examen obtuvo puntajes inferiores.

Ejemplo

1.15

Encuentre el 65avo. percentil de los datos del caso de asfalto en el ejemplo 1.12.

Solución

El tamaño de la muestra es $n = 24$. Para encontrar el 65avo. percentil, calcule $(0.65)(25) = 16.25$. Por tanto, el 65avo. percentil se encuentra al determinar el promedio del 16avo. y 17avo. puntos de datos, cuando la muestra se arregla en orden creciente. Se obtiene $(236 + 240)/2 = 238$.

En la práctica, con frecuencia un resumen estadístico como se ha descrito aquí se calcula en computadora, usando un software estadístico. El resumen estadístico de datos se conoce como **estadística descriptiva** porque describe los datos. Presentamos un ejemplo del cálculo del resumen estadístico del software MINITAB. Después mostraremos cómo se pueden usar estas estadísticas para descubrir algunas características importantes de los datos.

Para su tesis doctoral, J. Yanowitz, de la Escuela de Minas, de Colorado, investigó los factores que afectan las emisiones de un vehículo diesel, obtuvo datos acerca de la emisión de partículas (EP) para una muestra de 138 vehículos conducidos en bajas altitudes (cerca del nivel del mar) y para una muestra de 62 vehículos conducidos a grandes altitudes (aproximadamente a una milla del nivel del mar). Todos los vehículos se fabricaron entre 1991 y 1996. Las muestras contenían proporciones parecidas de vehículos de bajo y alto kilometraje. Los datos, en unidades de gramos de partículas por galón de combustible consumido, se presentan en las tablas 1.1 y 1.2. A grandes altitudes, la presión barométrica es más baja, así la razón de eficiencia aire/combustible también es más baja. Por esta razón se pensó que la emisión de partículas podría ser mayor a grandes altitudes. Nos gustaría comparar las muestras para determinar si los datos apoyan este supuesto. Es difícil hacer esto con sólo examinar los datos de las tablas. Calcular el resumen estadístico de los datos hace el trabajo mucho más fácil. La figura 1.4 presenta el cálculo con MINITAB de este resumen estadístico para ambas muestras.

TABLA 1.1 Emisión de partículas (EP) (en g/gal) para 138 vehículos conducidos a altitud baja

1.50	0.87	1.12	1.25	3.46	1.11	1.12	0.88	1.29	0.94	0.64	1.31	2.49
1.48	1.06	1.11	2.15	0.86	1.81	1.47	1.24	1.63	2.14	6.64	4.04	2.48
2.98	7.39	2.66	11.00	4.57	4.38	0.87	1.10	1.11	0.61	1.46	0.97	0.90
1.40	1.37	1.81	1.14	1.63	3.67	0.55	2.67	2.63	3.03	1.23	1.04	1.63
3.12	2.37	2.12	2.68	1.17	3.34	3.79	1.28	2.10	6.55	1.18	3.06	0.48
0.25	0.53	3.36	3.47	2.74	1.88	5.94	4.24	3.52	3.59	3.10	3.33	4.58
6.73	7.82	4.59	5.12	5.67	4.07	4.01	2.72	3.24	5.79	3.59	3.48	2.96
5.30	3.93	3.52	2.96	3.12	1.07	5.30	5.16	7.74	5.41	3.40	4.97	11.23
9.30	6.50	4.62	5.45	4.93	6.05	5.82	10.19	3.62	2.67	2.75	8.92	9.93
6.96	5.78	9.14	10.63	8.23	6.83	5.60	5.41	6.70	5.93	4.51	9.04	7.71
7.21	4.67	4.49	4.63	2.80	2.16	2.97	3.90					

TABLA 1.2 Emisión de partículas (EP) (en g/gal) para 62 vehículos conducidos a gran altitud

7.59	6.28	6.07	5.23	5.54	3.46	2.44	3.01	13.63	13.02	23.38	9.24	3.22
2.06	4.04	17.11	12.26	19.91	8.50	7.81	7.18	6.95	18.64	7.10	6.04	5.66
8.86	4.40	3.57	4.35	3.84	2.37	3.81	5.32	5.84	2.89	4.68	1.85	9.14
8.67	9.52	2.68	10.14	9.20	7.31	2.09	6.32	6.53	6.32	2.01	5.91	5.60
5.61	1.50	6.46	5.29	5.64	2.07	1.11	3.32	1.83	7.56			

Descriptive Statistics: LowAltitude, HiAltitude

Variable	N	Mean	SE Mean	TrMean	StDev
LoAltitude	138	3.715	0.218	3.526	2.558
HiAltitude	62	6.596	0.574	6.118	4.519

Variable	Minimum	Q1	Median	Q3	Maximum
LoAltitude	0.250	1.468	3.180	5.300	11.230
HiAltitude	1.110	3.425	5.750	7.983	23.380

FIGURA 1.4 Resultados arrojados por MINITAB que presenta la estadística descriptiva para los datos de EP de las tablas 1.1 y 1.2.

En la figura 1.4, la letra “N” denota el tamaño muestral. Enseguida está la media muestral. La siguiente cantidad (SE Mean) es el **error estándar de la media**. Este último es igual a la desviación estándar dividido entre la raíz cuadrada del tamaño muestral. Ésta es una cantidad que no se usa mucho como una estadística descriptiva, aunque es muy importante en aplicaciones como la construcción de intervalos de confianza y pruebas de hipótesis, que se presentarán en los capítulos 5 y 6. Enseguida del error estándar de la media está la media recortada al 5% (TrMean) y la desviación estándar. Por último, en el segundo bloque de resultados se proporciona el mínimo, mediano y máximo, así como el primer y tercer cuartiles (Q1 y Q3). Se observa que los valores de los cuartiles producidos por el software son ligeramente diferentes de los valores que se calcularían con los métodos que hemos descrito. Esto no es sorprendente, debido a que hay diferentes maneras de calcular estos valores. Las diferencias no son lo suficientemente grandes para que tengan alguna importancia práctica.

El resumen estadístico indica muchas diferencias en las emisiones de EP entre los vehículos de gran y baja altura. Primero, observe que la media es efectivamente más grande para vehículos a gran altura que para vehículos de baja altura (6.60 contra 3.71), lo que apoya la hipótesis de que las emisiones tienden a ser mayores a grandes altitudes. Ahora observe que el valor máximo para vehículos en grandes altitudes (23.38) es mucho más grande que el valor máximo para vehículos a baja altitud (11.23). Esto muestra que hay uno o más vehículos a grandes altitudes, cuyas emisiones son mayores que la más alta de los vehículos de bajas altitudes. ¿La diferencia en las medias de las emisiones podría ser atribuible totalmente a estos vehículos? Para contestar esto, compare las medianas, el primero y el tercer cuartil y la media recortada. Estos estadísticos no están muy afectados por algunos valores grandes, todos son notablemente más grandes para los vehículos a grandes altitudes. Por tanto, podemos concluir que los vehículos a grandes altitudes no sólo tienen emisiones muy grandes, también en general tienen emisiones más grandes que los vehículos a baja altitud. Por último, observe que la desviación estándar es más grande para vehículos a grandes altitudes, lo que indica que los valores para vehículos a grandes altitudes están más dispersos que los de los vehículos a bajas altitudes. Al menos algo de la diferencia en la dispersión se debe a uno o más vehículos a gran altitud con emisiones muy altas.

Resumen estadístico para datos categórico

Con datos categóricos, a cada elemento de la muestra se le asigna una categoría en lugar de un valor numérico. Es necesario trabajar con datos categóricos y resúmenes numéricos. Los dos más comunes son las **frecuencias** y las **proporciones muestrales** (algunas veces llamadas **frecuencias relativas**). La frecuencia para una categoría dada es sólo el número de elementos de la muestra que cae dentro de esa categoría. La proporción muestral es la frecuencia dividida entre el tamaño de la muestra.

Ejemplo 1.16

Un proceso fabrica cojinetes de cigüeñal para un motor de combustión interna. Los cojinetes, cuyo espesor está entre 1.486 y 1.490 mm, se clasifican para ajustar, ello significa que satisfacen la especificación. Los cojinetes más gruesos se rectifican y los más angostos se descartan. En una muestra de mil cojinetes, 910 se ajustaron, 53 se rectificaron y 37 se descartaron. Encuentre las frecuencias y las proporciones muestrales.

Solución

Las frecuencias son 910, 53 y 37. Las proporciones muestrales son $910/1\ 000 = 0.910$, $53/1\ 000 = 0.053$ y $37/1\ 000 = 0.037$.

Estadística muestral y parámetros poblacionales

Cada uno de los estadísticos de la muestra que hemos analizado tiene una contraparte en la población. Esto es fácil de ver cuando la población es finita. Por ejemplo, para una población finita de valores numéricos, la **media** de la población es sólo el promedio de todos los valores de la población; la **mediana** es el valor de en medio, o el promedio de los dos valores de en medio, y así sucesivamente. En efecto, cualquier resumen numérico que se usa para una muestra se puede usar para una **población finita**, con sólo aplicar los métodos de cálculo a los valores de la población en lugar de a los valores de la muestra. Existe una pequeña excepción para la varianza de la población cuando se divide entre n en lugar de $n - 1$. Existe una diferencia en la terminología que se usa para los resúmenes numéricos de las poblaciones con la que se usan para las muestras. Los resúmenes numéricos de una muestra se llaman **estadísticos**, mientras que los resúmenes numéricos de una población se llaman **parámetros**. Por supuesto, en la práctica, nunca se observa toda la población por lo que los parámetros de población no se pueden calcular directamente. En cambio, los estadísticos de la muestra se usan para calcular los valores de los parámetros de la población.

Los métodos para calcular los estadísticos de la muestra requieren que la muestra sea finita. Por tanto, cuando una población contiene un número infinito de valores, los métodos para calcular los estadísticos de una muestra no se pueden aplicar para calcular los parámetros de una población. Para poblaciones infinitas, los parámetros como la media y la varianza se calculan con procedimientos que generalizan los métodos usados para calcular las estadísticas de una muestra y que implican sumas infinitas o integrales. Describiremos estos procedimientos en el capítulo 2.

Resumen

- Un resumen numérico de una muestra se llama **estadístico**.
- Un resumen numérico de una población se llama **parámetro**.
- Con frecuencia los estadísticos se usan para estimar los parámetros.

Ejercicios para la sección 1.2

1. Verdadero o falso: para cualquier lista de números, la mitad de ellos estará debajo de la media.
2. ¿Es la media de la muestra siempre el valor que ocurre con más frecuencia? Si es así, explique por qué. Si no, dé un ejemplo.
3. ¿Es la media de la muestra siempre igual a uno de los valores que está en la muestra? Si es así, explique por qué. Si no, dé un ejemplo.
4. ¿La mediana de la muestra siempre es igual a uno de los valores de la muestra? Si es así, explique por qué. Si no, dé un ejemplo.
5. Encuentre un tamaño de la muestra para el cual la mediana siempre sea igual a uno de los valores en la muestra.
6. En cierta compañía, cada trabajador recibió un aumento de \$50 por semana. ¿Cómo afecta esto la media de los sueldos? ¿Y la desviación estándar de los sueldos?
7. En otra compañía, cada trabajador recibió un aumento de 5%. ¿Cómo afecta esto la media de los sueldos? ¿Y la desviación estándar de los sueldos?
8. El puntaje de Apgar se usa para evaluar reflejos y respuestas de recién nacidos. A cada bebé un profesional de la medicina le asigna un puntaje y los valores posibles son enteros entre cero y diez. Se toma una muestra de mil bebés nacidos en cierto condado y el número con cada puntaje es el siguiente:

Puntaje	0	1	2	3	4	5	6	7	8	9	10
Número de bebés	1	3	2	4	25	35	198	367	216	131	18
- a) Encuentre la media de la muestra de los puntajes de Apgar.
- b) Encuentre la desviación estándar de la muestra de los puntajes de Apgar.
- c) Encuentre la mediana muestral de los puntajes de Apgar.
- d) ¿Cuál es el primer cuartil de los puntajes?
- e) ¿Qué proporción de puntajes es más grande que la media?
- f) ¿Qué proporción de puntaje es mayor en una desviación estándar que la media?
- g) ¿Qué proporción de puntaje está dentro de una desviación estándar de la media?
9. Una clase de estadística con 40 estudiantes realizó una prueba. El puntaje posible más alto era de cuatro puntos. Diez estudiantes obtuvieron cuatro puntos, 12 lograron tres puntos, ocho alcanzaron dos puntos, seis se beneficiaron con un punto y cuatro obtuvieron cero puntos. Calcule la media, la mediana y la desviación estándar de los puntajes.
10. Otra clase de estadística de 60 estudiantes realizó la misma prueba. En este clase, 15 estudiantes obtuvieron cuatro puntos, 18 alcanzaron tres puntos, 12 lograron dos puntos, nueve obtuvieron un punto y seis resultaron con cero puntos. Calcule la media, la mediana y la desviación estándar de los puntajes.
11. En otra clase de estadística, el número total de estudiantes no se conoce. En esta clase, 25% obtuvo cuatro puntos, 30% alcanzó tres puntos, 20% se benefició con dos puntos, 15% logró un punto y 10% resultó con cero puntos.
 - a) ¿Es posible calcular la media de los puntajes para esta clase? Si es así, calcúlela. Si no, explique por qué.
 - b) ¿Es posible calcular la mediana de los puntajes para esta clase? Si es así, calcúlela. Si no, explique por qué.
 - c) ¿Es posible calcular la desviación estándar de la muestra de los puntajes para esta clase? Si es así, calcúlela. Si no, explique por qué.

12. Cada uno de los 16 estudiantes mide la circunferencia de una pelota de tenis por cuatro métodos diferentes, éstos fueron:

Método A: Estimar la circunferencia a simple vista.

Método B: Medir el diámetro con una regla y después calcular la circunferencia.

Método C: Medir la circunferencia con una regla y cuerda.

Método D: Medir la circunferencia haciendo rodar la pelota a lo largo de una regla.

Los resultados (en cm) son los siguientes, en orden creciente para cada método:

Método A: 18.0, 18.0, 18.0, 20.0, 22.0, 22.0, 22.5, 23.0, 24.0, 24.0, 25.0, 25.0, 25.0, 25.0, 26.0, 26.4.

Método B: 18.8, 18.9, 18.9, 19.6, 20.1, 20.4, 20.4, 20.4, 20.4, 20.5, 21.2, 22.0, 22.0, 22.0, 22.0, 23.6.

Método C: 20.2, 20.5, 20.5, 20.7, 20.8, 20.9, 21.0, 21.0, 21.0, 21.0, 21.5, 21.5, 21.5, 21.5, 21.6.

Método D: 20.0, 20.0, 20.0, 20.0, 20.2, 20.5, 20.5, 20.7, 20.7, 20.7, 21.0, 21.1, 21.5, 21.6, 22.1, 22.3.

- Calcule la media de las mediciones para cada método.
- Calcule la mediana de las mediciones para cada método.
- Calcule la media recortada a 20% de las mediciones para cada método.
- Calcule el primero y el tercer cuartil para cada método.
- Calcule la desviación estándar de las mediciones para cada método.
- ¿En qué método es la desviación estándar más grande? ¿Por qué se esperaría que este método tenga la desviación estándar más grande?

- Sin que nada cambie ¿es preferible un método de medición que tenga una desviación estándar más pequeña o uno con una desviación estándar más grande? ¿O no importa? Explique.

13. Con referencia al ejercicio 12.

- Si las mediciones para uno de los métodos se convirtieran a pulgadas (1 pulgada = 2.54 cm), ¿cómo afectaría esto la media? ¿Y la mediana? ¿Y los cuartiles? ¿Y la desviación estándar?
- Si los estudiantes midieran nuevamente la pelota, usando una regla marcada en pulgadas, ¿los efectos sobre la media, la mediana, los cuartiles y la desviación estándar serían los mismos que los del inciso a)? Explique.

14. Una lista de diez números tiene una media de 20, una mediana de 18 y una desviación estándar de 5. El número más grande en la lista es 39.27. Accidentalmente, este número se cambia a 392.7.

- ¿Cuál es el valor de la media después del cambio?
- ¿Cuál es el valor de la mediana después del cambio?
- ¿Cuál es el valor de la desviación estándar después del cambio?

15. ¿Por qué nadie habla del cuarto cuartil? ¿O lo hacen?

16. En cada uno de los siguientes conjuntos de datos, diga si el dato atípico parece ser atribuible a un error, o si se podría suponer que es correcto.

- Una roca se pesa cinco veces. Las lecturas en gramos son 48.5, 47.2, 4.91, 49.5, 46.3.
- Un sociólogo muestrea cinco familias en cierto pueblo y registra sus ingresos anuales. Los ingresos son \$34 000, \$57 000, \$13 000, \$12 000 000, \$62 000.

1.3 Resúmenes gráficos

Diagramas de tallo y hoja

La media, mediana y la desviación estándar son resúmenes numéricos de una muestra o de una población. Los resúmenes gráficos también se usan para ayudar a visualizar una lista de números. El resumen gráfico del que hablaremos primero es el **diagrama de tallo y hoja**. Una gráfica de tallos y hojas constituye una manera simple de resumir un conjunto de datos.

Como ejemplo, los datos de la tabla 1.3 tratan del géiser Old Faithful en el parque nacional Yellowstone. Este géiser alterna periodos de erupción, que duran normalmente de 1.5 a cuatro minutos, con espacios de tiempo de inactividad, que son considerablemente más grandes. La tabla 1.3 presenta la duración, en minutos, de 60 periodos de inactividad. La lista se presenta en orden numérico.

TABLA 1.3 Duración (en minutos) de los periodos de inactividad del géiser Old Faithful

42	45	49	50	51	51	51	51	53	53
55	55	56	56	57	58	60	66	67	67
68	69	70	71	72	73	73	74	75	75
75	75	76	76	76	76	76	79	79	80
80	80	80	81	82	82	82	83	83	84
84	84	85	86	86	86	88	90	91	93

La figura 1.5 presenta un diagrama de tallos y hojas de los datos de géiser. Cada elemento de la muestra se divide en dos partes: un **tallo**, que consta de uno o dos dígitos que están en el extremo izquierdo, y la **hoja**, que consta del siguiente dígito significativo. En los datos del géiser, el tallo es el dígito de las decenas y las hojas, una unidad. Cada renglón del diagrama de tallos y hojas contiene todos los elementos de la muestra con un tallo dado. El diagrama de tallos y hojas es una forma compacta de representar los datos. También indica un poco su forma. Para los datos de géiser podemos ver que relativamente hay pocas duraciones en el intervalo 60-69 minutos, comparado con los intervalos 50-59, 70-79 u 80-89 minutos.

Stem	Leaf
4	259
5	0111133556678
6	067789
7	01233455556666699
8	0000122233444456668
9	013

FIGURA 1.5 Gráfica de tallo y hojas para los datos del géiser de la tabla 1.3.

Cuando hay un gran número de elementos en la muestra con el mismo tallo, con frecuencia se necesita asignar más de un renglón a ese tallo. Por ejemplo, la figura 1.6 presenta una gráfica de tallos y hojas generada por computadora, con el software MINITAB, para los datos EP de la tabla 1.2 de la sección 1.2. La columna de en medio, que tiene los 0, 1 y 2, contiene los tallos, que son los dígitos de las decenas. A la derecha de los tallos están las hojas, que son los dígitos para cada uno de los elementos de la muestra. Como consecuencia de que muchos números son menores de 10, al tallo 0 se le deben asignar varios renglones, cin-

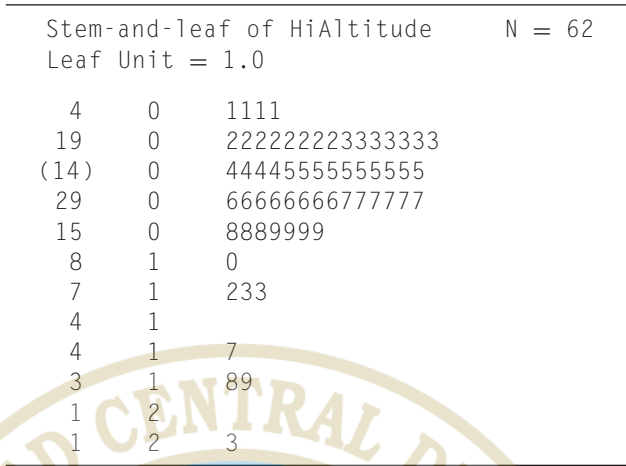


FIGURA 1.6 Diagrama de tallos y hojas de los datos de EP de la tabla 1.2 de la sección 1.2 generado por MINITAB.

co en este caso. Específicamente, el primer renglón tiene los elementos de la muestra cuyos dígitos son 0 o 1, el renglón siguiente tiene los elementos cuyos dígitos son 2 o 3, y así sucesivamente. Por consistencia, a todos los tallos se les asignan varios renglones del mismo modo, aunque sean pocos valores para los tallos 1 y 2, tallos que se podían haber hecho en pocos renglones.

El resultado de la figura 1.6 contiene una columna de frecuencia acumulada a la izquierda del diagrama de tallos y hojas. La parte superior de esta columna proporciona un recuento del número de elementos en o arriba del renglón actual y la parte más baja de la columna proporciona un recuento del número de elementos en o debajo del renglón actual. Después del renglón que contiene a la mediana mostrada entre paréntesis, se encuentra el recuento de elementos en este renglón.

Una buena característica de los diagramas de tallo y hojas es que exhiben todos los valores de la muestra. Se puede reconstruir la muestra totalmente a partir de un diagrama de tallo y hojas, con una excepción importante: el orden con el cual se muestrearon los elementos no se puede determinar.

Diagramas de puntos

Un **diagrama de puntos** es un gráfico que se puede usar para tener una impresión aproximada de la forma de una muestra. Es útil cuando el tamaño de la muestra no es demasiado grande y cuando la muestra contiene algunos valores repetidos. La figura 1.7 presenta un diagrama de puntos para los datos del géiser de la tabla 1.3. Para cada valor de la muestra se dibuja una columna vertical de puntos, con el número de puntos de la columna igual al número que aparece el valor en la muestra. El diagrama de puntos da una buena indicación de dónde se concentran los valores de la muestra y dónde hay separaciones. Por ejemplo, es evidente que en la figura 1.7 la muestra no tiene ningún periodo de inactividad entre los 61 y 65 minutos de duración.

Los diagramas de tallo y hojas y los diagramas de puntos son buenos métodos para revisar una muestra informalmente y se pueden dibujar bastante rápido con lápiz y papel. Sin

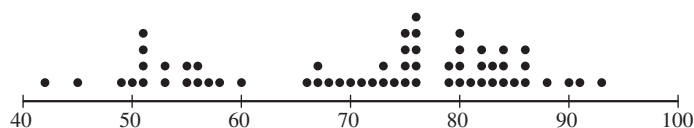


FIGURA 1.7 Diagrama de puntos para los datos del géiser de la tabla 1.3.

embargo, éstos rara vez se usan en presentaciones formales. Los gráficos comúnmente usados en las presentaciones formales son los histogramas y el diagrama de caja, que trataremos a continuación.

Histogramas

Un **histograma** es una gráfica que da una idea de la “forma” de una muestra, indicando las regiones donde los puntos de la muestra están concentrados y las regiones donde son escasos. Construiremos un histograma para las emisiones EP de 62 vehículos conducidos en una gran altitud, presentadas en la tabla 1.2 en la sección 1.2. El rango de la muestra va desde un mínimo 1.11 a un máximo de 23.38, en unidades de gramos de emisiones por galón de combustible. El primer paso es construir la **tabla de frecuencias**, que se muestra en la tabla 1.4.

TABLA 1.4 Tabla de frecuencias para las emisiones EP de 62 vehículos conducidos a mayor altitud

Intervalo de clase (g/gal)	Frecuencia	Frecuencia relativa	Densidad
1–< 3	12	0.194	0.0970
3–< 5	11	0.177	0.0885
5–< 7	18	0.290	0.1450
7–< 9	9	0.145	0.0725
9–< 11	5	0.081	0.0405
11–< 15	3	0.048	0.0120
15–< 25	4	0.065	0.0065

Los intervalos en la columna de la izquierda se llaman **intervalos de clase**. Dividen la muestra en grupos. La notación $1 - < 3$, $3 - < 5$ y así sucesivamente, indica que un punto que está en el límite entrará en la clase de su derecha. Por ejemplo, un valor de la muestra igual a 3 entrará en la clase $3 - < 5$, y no en la $1 - < 3$.

No hay ninguna regla ni difícil ni fácil respecto de cómo escoger los puntos finales de los intervalos de clase. En general, es bueno tener más intervalos en vez de menos, pero es también bueno tener un número grande de puntos de la muestra en los intervalos. Llegar al balance adecuado es una cuestión de criterio y de ensayo y error. En muchos casos es razonable tomar el número de intervalos de clase aproximadamente igual a la raíz cuadrada del tamaño de la muestra. Para los datos EP, los intervalos de clase de dos unidades funcionan bien, excepto para valores más grandes (por ejemplo, mayores que 11), donde los datos se reducen. Por tanto, se han agrupado los valores entre 11 y 15 respecto de una clase y todos los valores superiores a 15 en otra clase.

En la columna “Frecuencia”, en la tabla 1.4, se presentan los números de puntos de datos que están en cada uno de los intervalos de clase. En la columna “Frecuencia relativa” se presentan las frecuencias divididas entre el número total de puntos de datos, que para estos

datos es de 62. La frecuencia relativa de un intervalo de clase es la proporción de puntos de datos que están en el intervalo. Observe que debido a que cada punto de los datos está exactamente en un intervalo de clase, las frecuencias relativas deben sumar 1. Por último, en la columna “Densidad” se presenta la frecuencia relativa dividida entre el ancho de clase. Por ejemplo, en el primer renglón la frecuencia relativa es 0.194 y el ancho de clase es 2 ($3 - 1 = 2$). Por tanto, la densidad es $0.194/2 = 0.0970$. La última clase tiene un ancho de 10 y una frecuencia relativa de 0.065, por lo que su densidad es $0.065/10 = 0.0065$. El propósito de la densidad es ajustar la frecuencia relativa con el ancho de la clase. Sin que nada cambie, las clases anchas tienden a contener más elementos de la muestra que las clases más angostas y, por consiguiente, tienden a tener frecuencias relativas más grandes. Al dividir la frecuencia relativa entre el ancho de la clase se ajusta esta tendencia. La densidad representa la frecuencia relativa por unidad.

La figura 1.8 presenta el histograma para la tabla 1.4. Las unidades en el eje horizontal son las unidades de los datos, en este caso g/galón. Cada intervalo de clase se representa por un rectángulo. La altura de cada rectángulo es la densidad de la muestra en ese intervalo de clase, que está dado en la cuarta columna de la tabla 1.4. El área de cada rectángulo es, por tanto, la frecuencia relativa del intervalo de clase, que se encuentra en la tercera columna de la tabla 1.4. Debido a que las frecuencias relativas suman 1, el área bajo todo el histograma debe ser igual a 1.

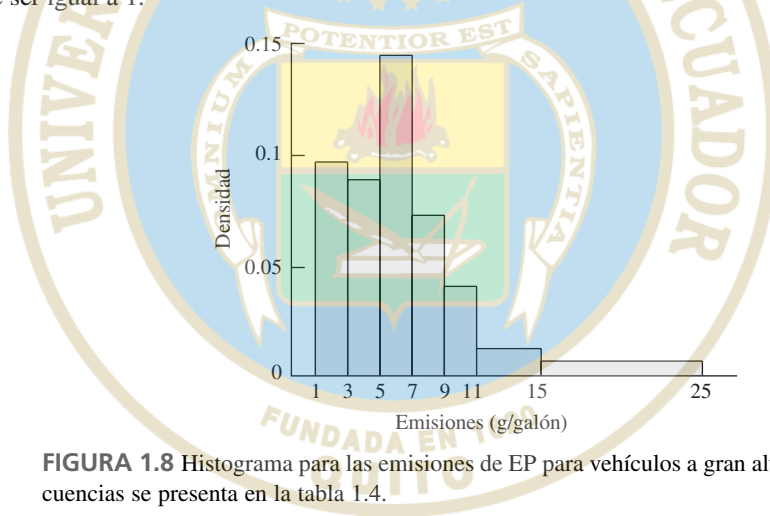


FIGURA 1.8 Histograma para las emisiones de EP para vehículos a gran altitud. La tabla de frecuencias se presenta en la tabla 1.4.

Ejemplo

1.17

Use el histograma de la figura 1.8 para determinar la proporción de los vehículos en la muestra con emisiones entre 7 y 11 g/galón.

Solución

La proporción representa el área bajo el histograma entre 7 y 11. Ésta se encuentra sumando las áreas cubiertas por los rectángulos de los dos intervalos de clase. El resultado es $(2)(0.0725) + (2)(0.0405) = 0.226$. Observe que este resultado también se puede obtener de la tabla de frecuencias. La proporción de puntos de datos con valores entre 7 y 9 es 0.145 y la proporción entre 9 y 11 es 0.081. La proporción entre 7 y 11 es, por tanto, igual a $0.145 + 0.081 = 0.226$.

Ejemplo

1.18

Use el histograma para calcular la proporción de vehículos en la muestra con emisiones entre 6 y 10 g/galón.

Solución

El histograma no da la respuesta exacta, ya que los valores 6 y 10 no son los puntos finales de los intervalos de clase. Calculamos la proporción calculando el área bajo el histograma entre los valores 6 y 10. Éste consta de la mitad del área del rectángulo en el intervalo de 5 a 7, más todo el rectángulo en el intervalo de 7 a 9, más la mitad del rectángulo en el intervalo de 9 a 11. El área es $(1)(0.1450) + (2)(0.0725) + (1)(0.0405) = 0.3305$.

Ejemplo

1.19

¿Cuál es la densidad de la muestra en 6 g/galón?

Solución

La densidad es la altura del histograma en ese punto. Esta altura es 0.1450.

Resumen

Para construir un histograma:

- Elija los puntos límite para los intervalos de clase.
- Calcule las frecuencias y las frecuencias relativas para cada clase.
- Calcule la densidad para cada clase, de acuerdo con la fórmula.

$$\text{Densidad} = \frac{\text{frecuencia relativa}}{\text{intervalo de clase}}$$

- Dibuje un rectángulo para cada clase, cuya altura sea igual a la densidad.

Anchos de clase iguales

La mayoría de los software estadísticos dibujan histogramas y dan la opción de especificar los intervalos de clase o de tener el software elegido para usted. Cuando se pide elegir intervalos de clase, la mayoría de softwares los hacen todos con el mismo ancho. Por ejemplo, la figura 1.9 presenta un histograma de los datos de la figura 1.8, con los intervalos de clase escogidos por MINITAB. Este último elige todos los intervalos con una anchura de dos.

Para muchos conjuntos de datos, las anchuras de clases son iguales. Pero para datos como éstos, con varios datos atípicos que se estiran a la derecha, los anchos de clases iguales son menos deseables. Para comprender por qué, considere que sólo siete puntos de datos de los 62 tienen valores superiores a 11. En la figura 1.9, más de la mitad de los intervalos de clase están dedicados a estos siete puntos y forman una cadena de pequeños rectángulos de diferentes tamaños. Estos rectángulos tienen mucha estructura visual, que distrae de la parte más importante de la estructura en la que está la mayor parte de los datos a la izquierda. Por otra

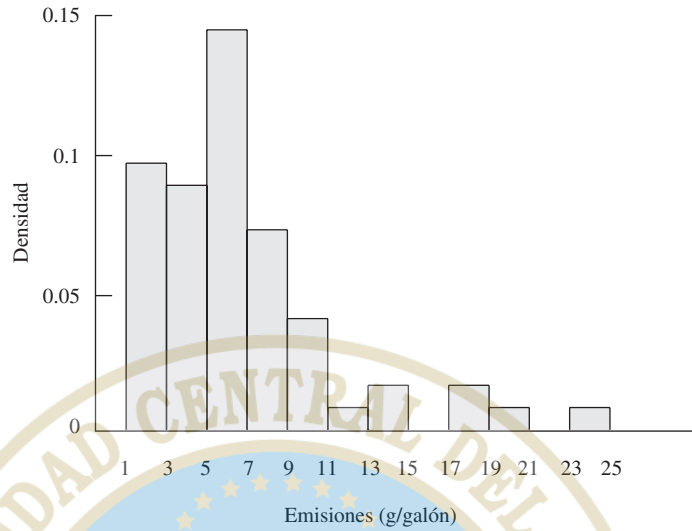


FIGURA 1.9 Histograma para los datos de la figura 1.8, con anchos de clase iguales como los elige MINITAB.

parte, el histograma en la figura 1.8 agrupa estos siete puntos en dos intervalos de clase, ello ofrece una apariencia más suave y permite apreciar mejor, a simple vista, la estructura del conjunto de datos como un todo.

El histograma, la media muestral y la varianza

Tanto la media como la varianza de la muestra tienen interpretaciones físicas con respecto al histograma. Imagine que el histograma de la figura 1.8 es una placa delgada que se sostiene sobre el eje horizontal. Suponga que la masa de cada rectángulo es proporcional a su área. La componente horizontal del centro de masa es el punto sobre el eje de x donde el histograma se balancearía si se apoyara allí. Para encontrar la componente horizontal del centro de masa del histograma, se trataría a cada rectángulo como si su masa estuviese concentrada en su punto medio. Se multiplicaría el punto medio de cada rectángulo por su área y al sumar los productos se obtiene el centro de masa. Los puntos medios de los rectángulos son los puntos medios de los intervalos de clases y las áreas representan las frecuencias relativas (véase la tabla 1.4). Por tanto, el centro de masa del histograma en la figura 1.8 está dado por

$$(2)(0.194) + (4)(0.177) + \cdots + (20)(0.065) = 6.730 \quad (1.6)$$

Este valor está bastante cerca de la media de la muestra, que es 6.596, como lo mostró el resultado de MINITAB (figura 1.4, en la sección 1.2). Para relacionar el centro de masa con la media de la muestra, observe que si cada elemento de la muestra tuviera un valor igual a la mitad de su intervalo de clase, entonces la expresión (1.6) sería igual a la media muestral. Por consiguiente, el centro de masa del histograma es una aproximación a la media muestral. Entre

más angosto sean los rectángulos, más cerca de cada elemento de la muestra estará el centro de su intervalo de clase y más cerca del centro de masa del histograma estará la media de la muestra.

Con el propósito de desarrollar una interpretación física para la varianza de la muestra, imagine una varilla sólida que pasa verticalmente a través del centro de masa del histograma (media de la muestra). Ahora imagine que toma la varilla y que curva al histograma alrededor de ella. A mayor extensión del histograma, sería más difícil darle vuelta. La cantidad física que mide la dificultad para darle vuelta es el momento de inercia. Para cada rectángulo en el histograma, el momento de inercia con respecto al centro de masa está dado por el cuadrado de la distancia del punto medio del rectángulo al centro de masa, multiplicado por el área del rectángulo. El momento de inercia para todo el histograma es la suma de los momentos de los rectángulos, que es

$$(2 - 6.730)^2(0.194) + (4 - 6.730)^2(0.177) + \dots + (20 - 6.730)^2(0.065) = 20.25 \quad (1.7)$$

Este valor está cerca de la varianza de la muestra, que es 20.42. [La varianza de la muestra se puede encontrar en los resultados arrojados por MINITAB (figura 1.4, de la sección 1.2) al elevar al cuadrado la desviación estándar, que es 4.519.] Si el valor de cada elemento de la muestra fuera exactamente igual al punto medio de su intervalo de clase, la ecuación (1.7) daría exactamente la varianza muestral. Es decir, el momento de inercia del histograma con respecto al centro de masa es una aproximación de la varianza muestral. Entre más angostos sean los rectángulos, la aproximación será más cercana.

El hecho de que la media y la varianza de la muestra correspondan a propiedades físicas del histograma es muy útil. En el capítulo 2 se desarrollarán métodos para calcular la media y la varianza poblacional para una población infinita al representar a la población con una curva y calculando el centro de masa y el momento de inercia.

Simetría y sesgo

Un histograma es perfectamente **simétrico** si su mitad derecha es una imagen de espejo de su mitad izquierda. Los histogramas que no son simétricos se llaman **sesgados**. En la práctica, ninguna muestra de datos tiene un histograma perfectamente simétrico; todos presentan algún grado de sesgadura. En un histograma sesgado, un lado, o una cola, es más largo que el otro. Un histograma con una cola larga a la derecha se dice que está **sesgado a la derecha** o **positivamente sesgado**. Un histograma largo con una cola larga a la izquierda se dice que está **sesgado a la izquierda** o **negativamente sesgado**. Aunque existe un método matemático formal para medir el sesgo de un histograma, rara vez se usa; en lugar de eso las personas juzgan el grado de sesgadura informalmente al mirar el histograma. La figura 1.10 presenta algunos histogramas para muestras de datos hipotéticas. Observe que para que un histograma esté sesgado a la derecha (figura 1.10c), la media es mayor que la mediana, porque más de la mitad de los datos estarán a la izquierda del centro de masa. Del mismo modo, la media es menor que la mediana para un histograma sesgado a la izquierda (figura 1.10a). El histograma para los datos de EP (figura 1.8) está sesgado a la derecha. La media de la muestra es 6.596, que es mayor que la mediana de la muestra de 5.75.

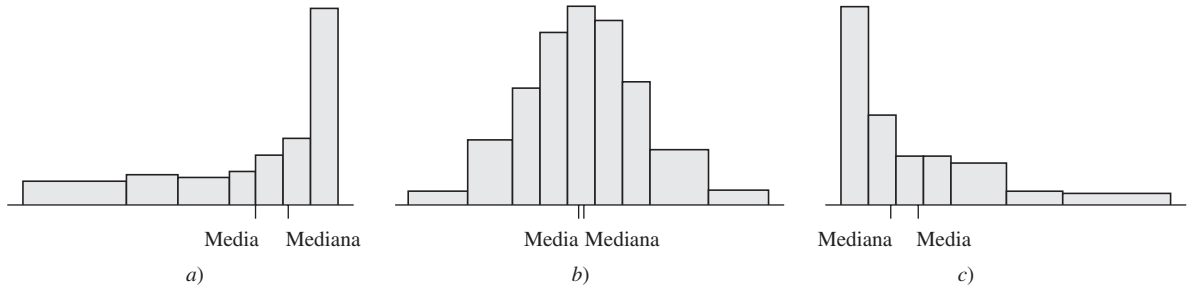


FIGURA 1.10 a) Histograma sesgado a la izquierda, la media es menor que la mediana. b) Histograma casi simétrico, la media y la mediana son aproximadamente iguales. c) Histograma sesgado a la derecha, la media es más grande que la mediana.

Histogramas unimodales y bimodales

Se ha usado el término “moda” para referirnos al valor que ocurre con más frecuencia en una muestra. Este término también se usa en histogramas y otras curvas para referirse a un pico o máximo local. Un histograma es **unimodal** si tiene solamente un pico, o moda, y **bimodal** si tiene evidentemente dos modas distintas. En principio, un histograma puede tener más de dos modas, pero esto no ocurre a menudo en la práctica. Los histogramas de la figura 1.10 son todos unimodales. La figura 1.11 presenta un histograma bimodal para una muestra hipotética.

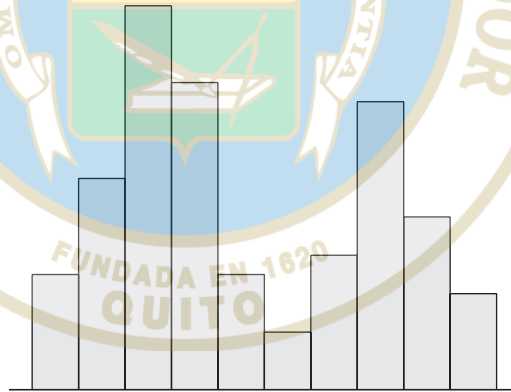


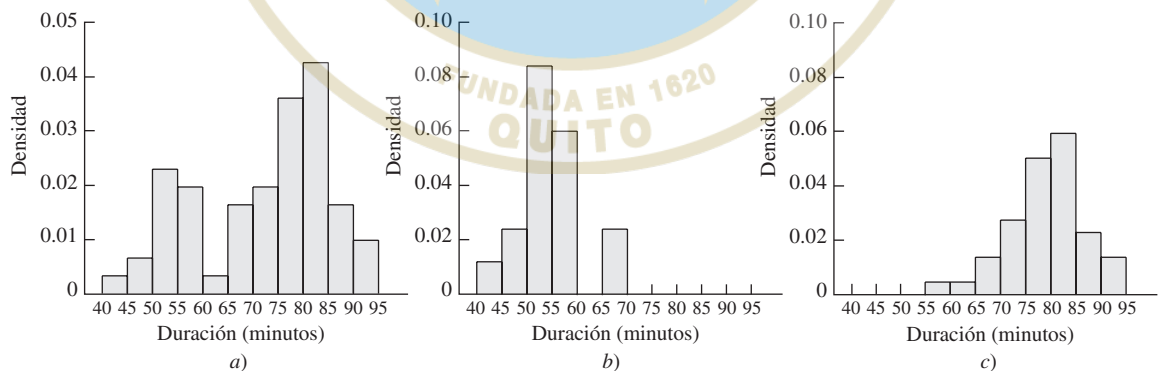
FIGURA 1.11 Histograma bimodal.

En algunos casos, un histograma bimodal indica que la muestra se puede dividir en dos submuestras que son diferentes entre sí, en alguna manera científica. A cada muestra corresponde una de las modas. Como ejemplo, la tabla 1.5 presenta la duración de los 60 periodos de inactividad del géiser Old Faithful (originalmente presentadas en la tabla 1.3). Junto con las duraciones del periodo de inactividad, en minutos, la duración de la erupción inmediatamente precedida por un periodo de inactividad se clasifica como breve (menos de tres minutos) o como larga (más de tres minutos).

TABLA 1.5 Duración de los periodos de inactividad (en minutos) y de las erupciones previas del géiser Old Faithful

Inactividad	Erupción	Inactividad	Erupción	Inactividad	Erupción	Inactividad	Erupción
76	Larga	90	Larga	45	Breve	84	Larga
80	Larga	42	Breve	88	Larga	70	Larga
84	Larga	91	Larga	51	Breve	79	Larga
50	Breve	51	Breve	80	Larga	60	Larga
93	Larga	79	Larga	49	Breve	86	Larga
55	Breve	53	Breve	82	Larga	71	Larga
76	Larga	82	Larga	75	Larga	67	Breve
58	Breve	51	Breve	73	Larga	81	Larga
74	Larga	76	Larga	67	Larga	76	Larga
75	Larga	82	Larga	68	Larga	83	Larga
80	Larga	84	Larga	86	Larga	76	Larga
56	Breve	53	Breve	72	Larga	55	Breve
80	Larga	86	Larga	75	Larga	73	Larga
69	Larga	51	Breve	75	Larga	56	Breve
57	Larga	85	Larga	66	Breve	83	Larga

La figura 1.12a presenta un histograma para las 60 duraciones. La figura 1.12b y c presenta histogramas para las duraciones de inactividad seguidas de erupciones breves y largas, respectivamente. El histograma para todas las duraciones es bimodal. Los histogramas para las duraciones seguidas de erupciones breves o largas son ambos unimodales y sus modas forman las dos modas del histograma para toda la muestra.

**FIGURA 1.12** a) Histograma para las 60 duraciones de la tabla 1.5. Este histograma es bimodal. b) Histograma para las duraciones de la tabla 1.5 que siguen a las erupciones breves. c) Histograma para las duraciones en la tabla 1.5 que siguen a las erupciones largas. Tanto los histogramas para las duraciones seguidas de erupciones breves como para los que le siguen erupciones largas son unimodales, pero las modas están en lugares diferentes. Cuando las dos muestras se combinan, el histograma es bimodal.

Resumen

Los histogramas tienen las propiedades siguientes:

- El área de cada rectángulo representa la proporción de la muestra que está en el intervalo de clase correspondiente.
- La altura de cada rectángulo representa la densidad de la muestra en el intervalo de clase correspondiente.
- El área total bajo el histograma es igual a 1.
- La media muestral es casi igual al centro de masa del histograma. La aproximación se hace más cercana cuando se hacen rectángulos más angostos.
- La varianza muestral es aproximadamente igual al momento de inercia del histograma con respecto a su centro de masa. La aproximación es más cercana cuando se hacen rectángulos más angostos.

Haciendo las alturas iguales a las frecuencias

En este libro se usa el término “histograma” para referirse a una gráfica en la que las alturas de los rectángulos representan las *densidades* (así las áreas representan las frecuencias relativas). Sin embargo, algunas personas dibujan histogramas con las alturas de los rectángulos iguales a las *frecuencias*. En efecto, éste es el método usual en muchos software. El hacer las alturas iguales a las frecuencias (o a las frecuencias relativas) puede producir una imagen distorsionada de los datos. Por ejemplo, la figura 1.13 presenta un histograma para los datos de EP en los que las alturas son iguales a las frecuencias, usando los mismos intervalos de clase que se muestran en el histograma de la figura 1.8. Este histograma exagera visualmente la proporción de vehículos en las dos clases más grandes. La razón de esto es que dichos intervalos de clase son más amplios que el resto y el histograma no se ajusta a ese dato.

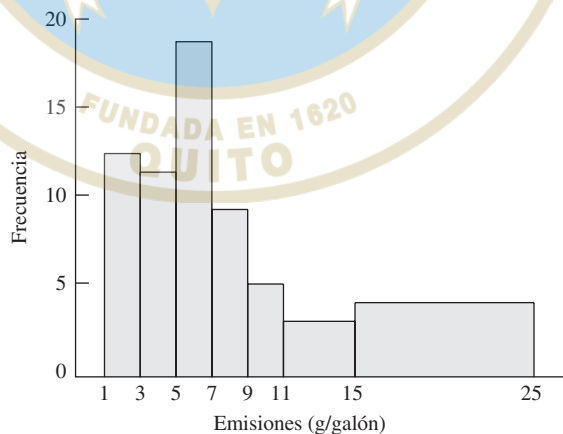


FIGURA 1.13 Histograma para la muestra de la tabla 1.4, con las alturas iguales a las frecuencias. El rectángulo grande sobre el intervalo 15-25 es engañoso; en efecto, solamente 6.5% de la muestra está en ese intervalo. Compárelo con el histograma dibujado correctamente en la figura 1.8 de la página 28.

Si todos los intervalos de clase tienen el mismo ancho, entonces el histograma tendrá la misma forma si las alturas representan las densidades o las frecuencias. En este caso, hacer las alturas iguales a las frecuencias no es engañoso. Sin embargo, cuando se desea hacer intervalos de clase de anchos diferentes, es importante que las alturas de los rectángulos sean iguales a las *densidades* y no a las frecuencias o las frecuencias relativas.

Por último, se indica que cuando las alturas sean iguales a las frecuencias o a las frecuencias relativas, el área total de los rectángulos no es en general igual a 1. Si todos los intervalos de clase tienen el mismo ancho, entonces la media muestral estará aproximadamente en el centro de masa; de otra forma generalmente no estará.

Diagramas de caja

Un **diagrama de caja** constituye una gráfica que incluye la mediana, el primero y el tercer cuartil y cualquier dato atípico que se presente en una muestra. Los diagramas de caja son fáciles de comprender, pero hay una pequeña terminología asociada con ellos. El **rango intercuartil** es la diferencia entre el tercer y el primer cuartil. Observe que debido a que 75% de los datos son menores que el tercer cuartil y que 25% de los datos son menores que el primer cuartil, la mitad de los datos está entre el primero y el tercer cuartil. Por tanto, el rango intercuartil representa la distancia necesaria para atravesar la mitad de los datos de en medio.

Se ha definido a los datos atípicos como puntos que son inusualmente grandes o pequeños. Si denota IQR el rango intercuartil, entonces con el propósito de dibujar diagramas de caja, cualquier punto que está a más de 1.5 IQR por arriba del tercer cuartil, o que está a más de 1.5 IQR por debajo del primer cuartil, se considera un dato atípico. Algunos textos definen a un punto que está a más de 3 IQR del primero o del tercer cuartil como un **dato atípico extremo**. Estas definiciones de datos atípicos son sólo convenciones para el dibujo de los diagramas de caja y no se necesitan usar en otras situaciones.

La figura 1.14 presenta un diagrama de caja para algunos datos hipotéticos. El diagrama consta de una caja cuyo lado inferior es el primer cuartil y el lado superior es el tercer cuartil. La mediana se dibuja como una línea horizontal. Los datos “atípicos” se grafican por separado y se indican con cruces en la figura. Los que se extienden desde la parte superior a la inferior de la caja son líneas verticales llamadas “bigotes”. Éstos terminan en los puntos más extremos que no son atípicos.

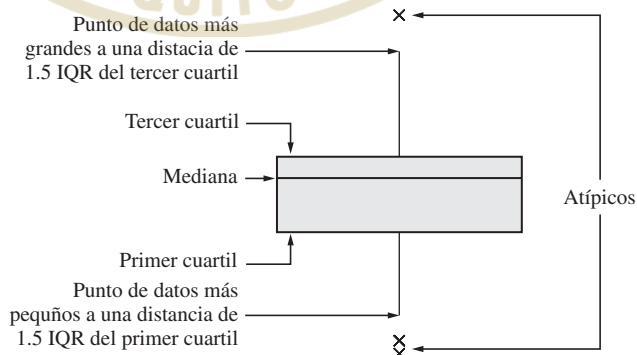


FIGURA 1.14 Anatomía de un diagrama de caja.

Además de cualquier dato atípico, un diagrama de caja está formado por cuatro partes: las dos partes de la caja separadas por la línea de la mediana y los dos “bigotes”. Nuevamente, aparte de los datos atípicos, cada una de estas cuatro partes representa una cuarta parte de los datos. Por tanto, el diagrama de caja indica la longitud de un intervalo para cada cuarta parte de datos y de esta manera se puede usar para determinar las regiones en las que hay mayor y menor densidad de valores de la muestra.

Pasos para la construcción de un diagrama de caja

- Calcule la mediana, el primero y tercer cuartil de la muestra. Indique éstos con líneas horizontales. Dibuje líneas verticales para completar la caja.
- Encuentre el valor de la muestra más grande que no esté a más de 1.5 IQR arriba del tercer cuartil y el valor de la muestra más pequeño que no esté a más de 1.5 IQR debajo del primer cuartil. Extienda líneas verticales (“bigotes”) desde las líneas de los cuartiles a estos puntos.
- Puntos a más de 1.5 IQR arriba del tercer cuartil, o a más de 1.5 IQR por debajo del primer cuartil, se denominan datos atípicos. Dibuje cada dato atípico por separado.

La figura 1.15 presenta un diagrama de caja para los datos del géiser que se presentó en la tabla 1.5. Primero observe que no hay ningún dato atípico en estos datos. Al comparar las cuatro partes del diagrama de caja, se puede decir que los valores de la muestra tienen casi la misma densidad de valores entre la mediana y el tercer cuartil y menos densidad entre la mediana y el primer cuartil. El “bigote” más bajo es un poco más largo que el superior, lo que indica que los datos tienen la cola inferior ligeramente más larga que la cola superior. Como consecuencia de que la distancia entre la mediana y el primer cuartil es más grande que la que existe entre la mediana y el tercer cuartil y debido a que la cuarta parte inferior de los datos produce un “bigote” más largo que la cuarta parte superior, este diagrama de caja indica que los datos están sesgados a la izquierda.

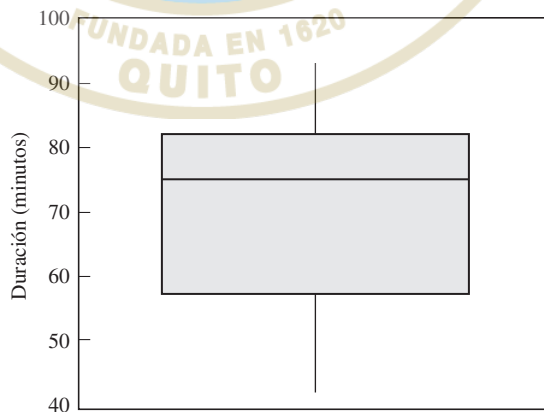


FIGURA 1.15 Diagrama de caja para los datos de los periodos de inactividad del Old Faithful que se presentaron en la tabla 1.5.

Se presentó un histograma para estos datos en la figura 1.12a. El histograma presenta una impresión más general de la dispersión de los datos. En forma importante, el histograma indica que los datos son bimodales, lo que un diagrama de caja no puede hacer.

Diagramas de caja comparativos

Una ventaja muy importante de los diagramas de caja es que se pueden presentar varios juntos, ello permite la fácil comparación visual de las características de varias muestras. Las tablas 1.1 y 1.2 (en la sección 1.2) presentan datos de emisiones de EP para vehículos conducidos en grandes y bajas altitudes. La figura 1.16 presenta una comparación de los dos diagramas de caja de estas dos muestras.

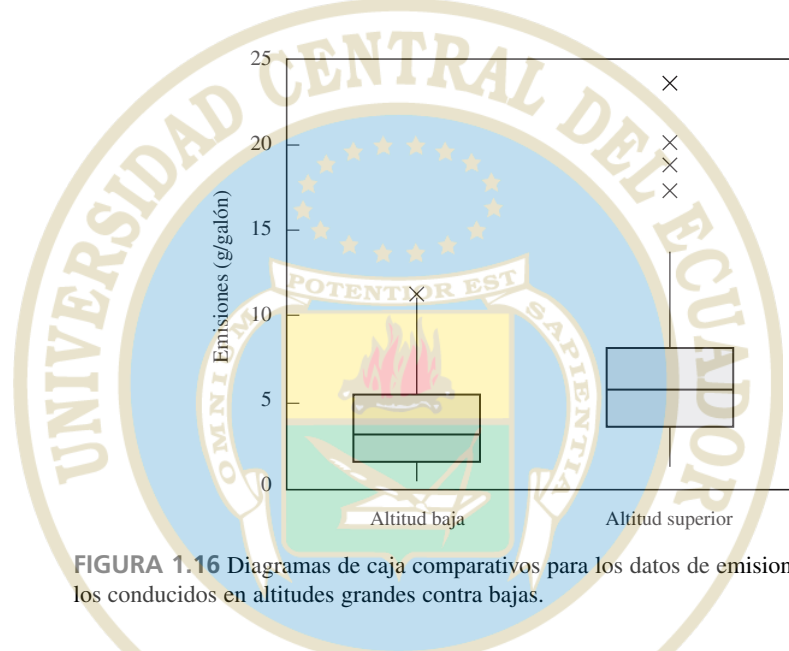


FIGURA 1.16 Diagramas de caja comparativos para los datos de emisiones de EP para vehículos conducidos en altitudes grandes contra bajas.

Los diagramas de caja comparativos en la figura 1.16 muestran que los vehículos conducidos a baja altitud tienden a tener emisiones inferiores. Además, hay algunos datos atípicos entre los datos para los vehículos en grandes altitudes, cuyos valores son superiores que cualquiera de los valores para los vehículos de baja altitud (también hay un valor de baja altitud que apenas reúne las condiciones necesarias para considerarlo un dato atípico). Se concluye que a grandes altitudes, los vehículos tienen emisiones mayores en general y que cuando consideramos los datos atípicos es mucho mayor. La caja para vehículos a grandes altitudes es un poco más grande y el “bigote” inferior es un poco más largo que para vehículos a baja altitud. Se concluye que además de los datos atípicos, la dispersión en los valores es ligeramente más grande para los vehículos a grandes altitudes y mucho más grande cuando se considera a los datos atípicos.

En la figura 1.4 (en la sección 1.2) se comparan los valores de algunos estadísticos descriptivos numéricos para estas dos muestras y se obtienen algunas conclusiones similares a las anteriores. La naturaleza visual de los diagramas de caja de la figura 1.16 hace las comparaciones de las características de las muestras mucho más fácil.

Se ha mencionado que es importante examinar a los datos atípicos para determinar si son resultado de errores y, en tal caso, eliminarlos. Para identificar a los datos atípicos, pueden ser útiles los diagramas de caja. El siguiente ejemplo proporciona un caso.

El artículo “Virgin Versus Recycled Wafers for Furnace Qualification: Is the Expense Justified?” (V. Czitrom y J. Reece, en *Statistical Case Studies for Industrial Process Improvement*, ASA y SIAM, 1997:87-104) describe un proceso para el crecimiento de una capa delgada de dióxido de silicio sobre placas de silicio que se usan en la fabricación de semiconductores. La tabla 1.6 presenta las mediciones del espesor, en angstroms (Å), de la capa de óxido para 24 placas. Se hicieron nueve mediciones en cada placa. Las placas se fabricaron en dos corridas distintas, con 12 placas por cada corrida.

TABLA 1.6 Espesor de las capas de óxido de silicio en placas de silicio

Placa		Espesor (Å)								
Corrida 1	1	90.0	92.2	94.9	92.7	91.6	88.2	82.0	98.2	96.0
	2	91.8	94.5	93.9	77.3	92.0	89.9	87.9	92.8	93.3
	3	90.3	91.1	93.3	93.5	87.2	88.1	90.1	91.9	94.5
	4	92.6	90.3	92.8	91.6	92.7	91.7	89.3	95.5	93.6
	5	91.1	89.8	91.5	91.5	90.6	93.1	88.9	92.5	92.4
	6	76.1	90.2	96.8	84.6	93.3	95.7	90.9	100.3	95.2
	7	92.4	91.7	91.6	91.1	88.0	92.4	88.7	92.9	92.6
	8	91.3	90.1	95.4	89.6	90.7	95.8	91.7	97.9	95.7
	9	96.7	93.7	93.9	87.9	90.4	92.0	90.5	95.2	94.3
	10	92.0	94.6	93.7	94.0	89.3	90.1	91.3	92.7	94.5
	11	94.1	91.5	95.3	92.8	93.4	92.2	89.4	94.5	95.4
	12	91.7	97.4	95.1	96.7	77.5	91.4	90.5	95.2	93.1
Corrida 2	1	93.0	99.9	93.6	89.0	93.6	90.9	89.8	92.4	93.0
	2	91.4	90.6	92.2	91.9	92.4	87.6	88.9	90.9	92.8
	3	91.9	91.8	92.8	96.4	93.8	86.5	92.7	90.9	92.8
	4	90.6	91.3	94.9	88.3	87.9	92.2	90.7	91.3	93.6
	5	93.1	91.8	94.6	88.9	90.0	97.9	92.1	91.6	98.4
	6	90.8	91.5	91.5	91.5	94.0	91.0	92.1	91.8	94.0
	7	88.0	91.8	90.5	90.4	90.3	91.5	89.4	93.2	93.9
	8	88.3	96.0	92.8	93.7	89.6	89.6	90.2	95.3	93.0
	9	94.2	92.2	95.8	92.5	91.0	91.4	92.8	93.6	91.0
	10	101.5	103.1	103.2	103.5	96.1	102.5	102.0	106.7	105.4
	11	92.8	90.8	92.2	91.7	89.0	88.5	87.5	93.8	91.4
	12	92.1	93.4	94.0	94.7	90.8	92.1	91.2	92.3	91.1

Las 12 placas en cada corrida eran de varios tipos y se procesaron en diferentes posiciones en el horno. El propósito en la recopilación de datos fue determinar si el espesor de la capa de óxido se afectaba ya sea por el tipo de placa o por la posición en el horno. Por tanto, éste fue un experimento factorial, con los factores, tipo de placa y posición en el horno y como resultado el espesor de la capa de óxido. El experimento se diseñó de tal manera que no se supuso ninguna diferencia sistemática entre las capas de una corrida a otra. El primer paso en el análisis fue construir un diagrama de caja para los datos de cada corrida con el propósito de ayudar a determinar si esta condición se satisfacía realmente y si ninguna de las observaciones se debía eliminar. Los resultados se presentan en la figura 1.17.

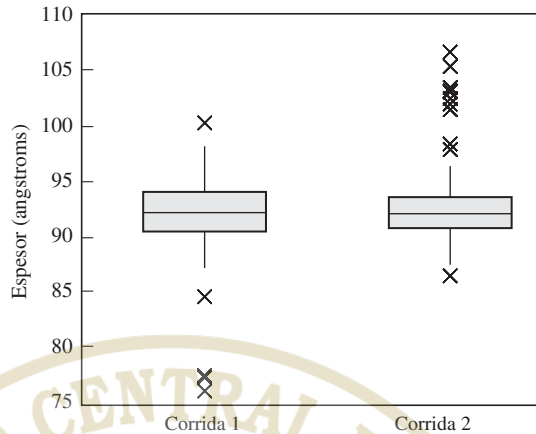


FIGURA 1.17 Diagramas de caja comparativos para los datos del espesor de la capa de óxido.

Los diagramas de caja muestran que había algunos datos atípicos en cada corrida. Observe que, aparte de estos datos atípicos, no hay ninguna diferencia sorprendente entre las muestras y, por tanto, ninguna prueba de alguna diferencia sistemática entre las corridas. La siguiente tarea es inspeccionar los datos atípicos, para determinar cuál, si hay alguno, se debe eliminar. Al examinar los datos de la tabla 1.6, se observa que las ocho mediciones más grandes en dos corridas ocurrieron en una sola placa: la número 10.

Se determinó entonces que esta placa se había contaminado con un residuo de la película, lo que ocasionó grandes mediciones del espesor. Por tanto, sería apropiado eliminar estas mediciones. En el experimento real, los ingenieros tenían a su disposición datos de algunas otras corridas y, por razones técnicas, decidieron eliminar toda la corrida, en vez de analizar una corrida a la que le faltaba una placa. En la corrida 1 se encontró que las tres mediciones más bajas se habían dado por un calibrador descompuesto y, por tanto, se eliminaron adecuadamente. No se pudo determinar ninguna causa para los dos datos atípicos restantes de la corrida 1, así que permanecieron en el análisis.

Datos multivariados

A veces los elementos de una población pueden tener algunos valores asociados entre sí. Por ejemplo, imagine que elige una muestra aleatoria de días y determine el promedio de la temperatura y de la humedad para cada día. Cada día la población proporciona dos valores, la temperatura y la humedad. Por tanto, la muestra aleatoria constaría de pares de números. Si las precipitaciones también se midieran todos los días, la muestra constaría de tripletes. En principio, se podría medir todos los días cualquier número de cantidades, lo que produciría una muestra en la que cada elemento representa una lista de números.

Los datos para cada elemento que constan de más de un valor se llaman **datos multivariados**. Cuando cada elemento es un par de valores, se dice que los datos son **bivariados**. Uno de los resúmenes gráficos más útiles por los datos bivariados numéricos es el **diagrama de dispersión**. Si los datos constan de pares arreglados $(x_1, y_1), \dots, (x_n, y_n)$, entonces un diagrama de dispersión se construye sólo al trazar cada punto en un sistema coordenado bidi-

mensional. Los diagramas de dispersión también se pueden usar para resumir los datos multivariados cuando cada elemento consta de más de dos valores. Simplemente se construirían diagramas de dispersión distintos para cada par de valores.

El siguiente ejemplo muestra la utilidad de los diagramas de dispersión. El artículo “Advances in Oxygen Equivalence Equations for Predicting the Properties of Titanium Welds” (D. Harwig, W. Ittiwattana y H. Castner, *The Welding Journal*, 2001:126s-136s) presenta los datos con respecto a la composición química y las características de la intensidad de diferentes soldaduras de titanio. La figura 1.18 presenta dos diagramas de dispersión. La figura 1.18a constituye un diagrama de la intensidad producida [en miles de libras por pulgada cuadrada (ksi)] contra el contenido de carbono (en %) para algunas de estas soldaduras. La figura 1.18b representa un diagrama de la intensidad producida (en ksi) contra el contenido de nitrógeno (en %) para las mismas soldaduras.

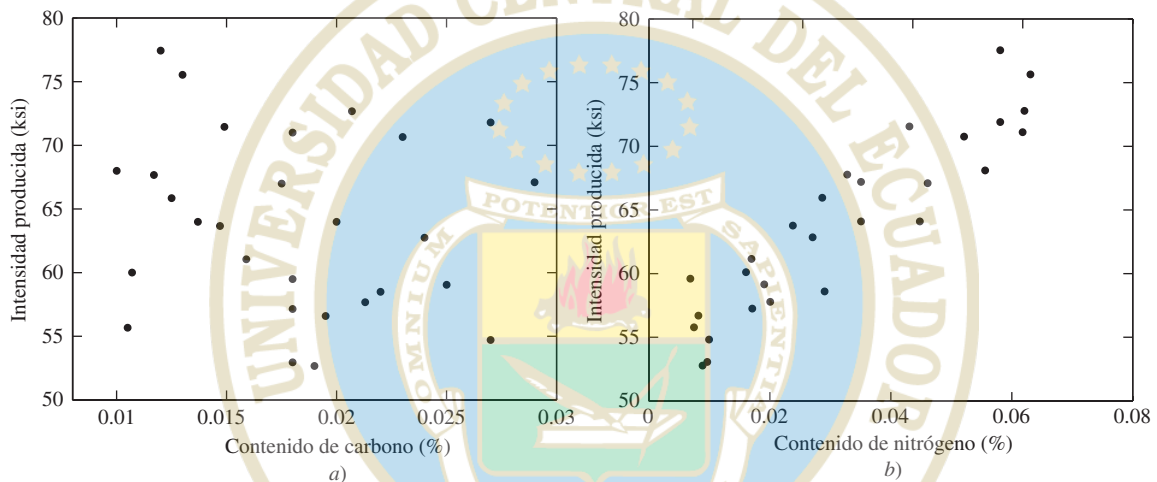


FIGURA 1.18 a) Un diagrama de dispersión que muestra que no existe mucha relación entre el contenido de carbono y la intensidad producida para determinado grupo de soldaduras. b) Un diagrama de dispersión que muestra que para estas mismas soldaduras el contenido de nitrógeno más alto está relacionado con la mayor intensidad producida.

La gráfica de la intensidad producida contra contenido de nitrógeno (figura 1.18b) muestra con cierta claridad la estructura, los puntos parecen estar siguiendo una recta desde la parte izquierda más baja a la parte derecha superior. De este modo, la gráfica muestra una relación entre el contenido de nitrógeno y la intensidad producida: las soldaduras con más alto contenido de nitrógeno tienden a tener una mayor intensidad producida. Este diagrama de dispersión podría conducir a que los investigadores intenten predecir la intensidad a partir del contenido de nitrógeno o que traten de aumentar el contenido de nitrógeno para incrementar la intensidad. (El hecho de que exista una relación en la gráfica de dispersión no garantiza que estos intentos tengan éxito, como se analizará en la sección 7.1.) Por otra parte, no parece haber mucha estructura en la gráfica de dispersión de la intensidad producida contra el contenido de carbono y, por tanto, no hay evidencia de que exista una relación entre estas dos cantidades. Esta gráfica de dispersión desanimaría a los investigadores a tratar de predecir la intensidad a partir del contenido de carbono.

Ejercicios para la sección 1.3

1. Parte de un estudio de control de calidad tuvo como objetivo mejorar una línea de producción, se midieron los pesos (en onzas) de 50 barras de jabón. Los resultados son los siguientes, ordenados del más pequeño al más grande.

11.6 12.6 12.7 12.8 13.1 13.3 13.6 13.7 13.8 14.1
14.3 14.3 14.6 14.8 15.1 15.2 15.6 15.6 15.7 15.8
15.8 15.9 15.9 16.1 16.2 16.2 16.3 16.4 16.5 16.5
16.5 16.6 17.0 17.1 17.3 17.3 17.4 17.4 17.4 17.6
17.7 18.1 18.3 18.3 18.3 18.5 18.5 18.8 19.2 20.3

- a) Construya un diagrama de tallos y hojas para estos datos.
b) Construya un histograma para estos datos.
c) Construya un diagrama de puntos para estos datos.
d) Construya un diagrama de caja para estos datos. ¿El diagrama de caja indica datos atípicos?

2. Siguiendo una lista del número de sitios de residuos peligrosos en cada uno de los 50 estados federales de los Estados Unidos en abril de 1995. Los datos se tomaron de *The World Almanac and Book of Facts 1996* (World Almanac Books, Mahwah, NJ, 1996). La lista se ha ordenado numéricamente.

1 2 3 4 4 5 6 8 8 9
10 10 10 11 11 11 12 12 12 12
13 13 14 15 16 17 17 18 18 19
19 20 22 23 24 25 29 30 33 37
38 39 40 55 58 77 81 96 102 107

- a) Construya un diagrama de tallos y hojas para estos datos.
b) Construya un histograma para estos datos.
c) Construya un diagrama de puntos para estos datos.
d) Construya un diagrama de caja para estos datos. ¿El diagrama de caja muestra algún dato atípico?

3. Con referencia a la tabla 1.2 (p. 20). Construya un diagrama de tallos y hojas con uno de los dígitos como tallo (para valores iguales o superiores a 10, el tallo tendrá dos dígitos) y el dígito de las decenas como hoja. ¿Cuántos tallos hay (asegúrese de incluir tallos deshojados)? ¿Cuáles son algunas de las ventajas y desventajas de este diagrama, comparado con el de la figura 1.6 (p. 26)?

4. Se estudiaron dos métodos de recuperación de proteína. Se hicieron trece corridas usando cada método y se registró la

parte de proteína recuperada para cada corrida. Los resultados son los siguientes:

Método 1	Método 2
0.32	0.25
0.35	0.40
0.37	0.48
0.39	0.55
0.42	0.56
0.47	0.58
0.51	0.60
0.58	0.65
0.60	0.70
0.62	0.76
0.65	0.80
0.68	0.91
0.75	0.99

- a) Construya un histograma para los resultados de cada método.
b) Construya diagramas de caja para comparar los dos métodos.
c) Usando los diagramas de caja, ¿qué diferencias se observan entre los resultados de los dos métodos?

5. Cada uno de los 32 estudiantes, que forman dos secciones de laboratorio con 16 estudiantes cada una, calculó la circunferencia de una pelota de tenis a simple vista. Aquí se muestran los resultados, en centímetros. (Los resultados para el primer grupo de estudiantes también se muestran en el ejercicio 12 de la sección 1.2.)

Grupo 1	Grupo 2
18.0	15.0
18.0	18.0
18.0	18.0
20.0	19.0
22.0	19.0
22.0	19.0
22.5	19.0
23.0	19.5
24.0	20.0
24.0	20.0
25.0	20.0
25.0	20.0
25.0	20.0
25.0	22.0
26.0	24.0
26.4	25.0

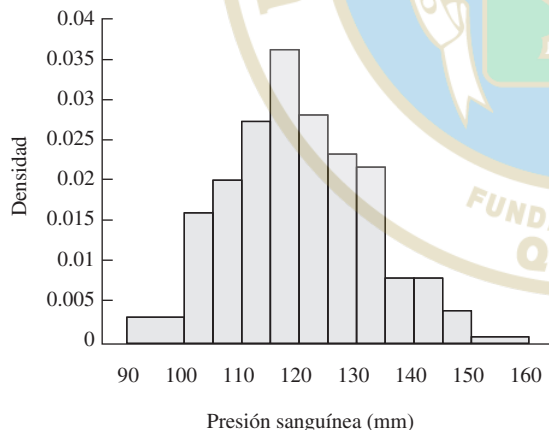
- Construya un histograma para cada grupo.
- Construya diagramas de caja comparativos para los dos grupos.
- Usando diagramas de caja, ¿qué diferencias se pueden ver entre los resultados del primero y del segundo grupo?

6. Dibuje un histograma para el cual:

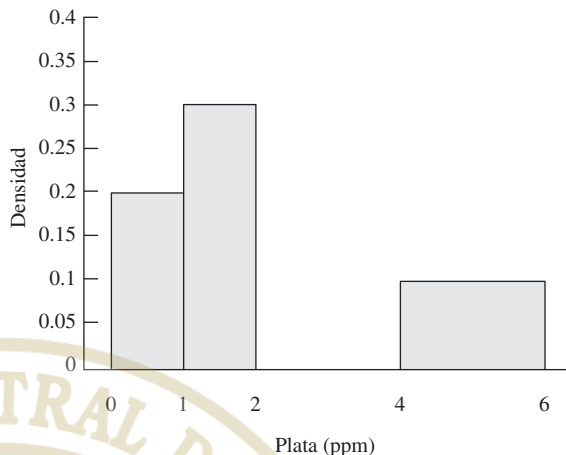
- La media es más grande que la mediana.
- La media es menor que la mediana.
- La media es aproximadamente igual a la mediana.

7. El siguiente histograma presenta la distribución de la presión sanguínea sistólica para una muestra de mujeres. Use éste para responder las siguientes preguntas:

- ¿Es el porcentaje de mujeres con presión sanguínea superior a 130 mm cercano a 25, 50 o 75%?
- ¿En qué intervalo hay más mujeres: 130-135 o 140-150 mm?
- En el intervalo 125-130 mm, la altura del histograma es de 0.024. ¿Qué porcentaje de mujeres tenía presión sanguínea en este intervalo?



8. El siguiente histograma presenta las cantidades de plata [en partes por millón (ppm)] encontradas en una muestra de rocas. Falta un rectángulo del histograma. ¿Cuál es su altura?



9. Una muestra de 100 hombres tiene una estatura promedio de 70 pulgadas y una desviación estándar de 2.5 pulgadas. Una muestra de 100 mujeres tiene una estatura promedio de 64 pulgadas y desviación estándar de 2.5 pulgadas. Si se combinan ambas muestras, la desviación estándar de la estatura de las 200 personas será _____

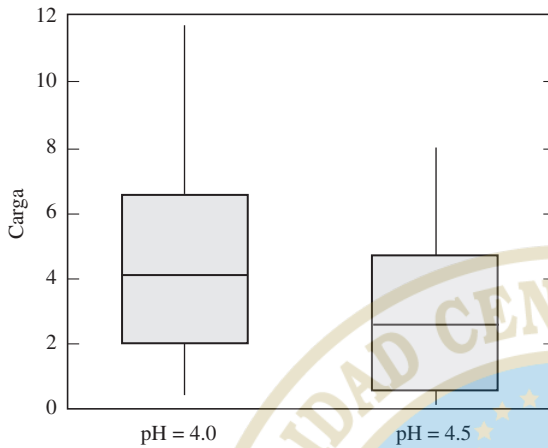
- menor que 2.5 pulgadas
- mayor que 2.5 pulgadas
- igual a 2.5 pulgadas
- no se puede decir a partir de la información dada.

(Sugerencia: No realice ningún cálculo. Sólo trate de dibujar histogramas para cada muestra por separado, y después para la muestra combinada.)

10. Los siguientes son diagramas de caja que comparan la carga [en coulombs entre mol (C/mol) $\times 10^{-25}$] a pH 4.0 y a pH 4.5 para un conjunto de proteínas (del artículo "Optimal Synthesis of Protein Purification Processes", E. Vasquez-Alvarez, M. Leinqueo y J. Pinto, en *Biotechnology Progress* 2001:685-695). Verdadero o falso:

- La mediana de la carga para el pH 4.0 es mayor que el 75avo. percentil de la carga para el pH 4.5.
- Aproximadamente 25% de las cargas para el pH 4.5 son menores que la carga más pequeña en el pH 4.0.
- Cerca de la mitad de los valores de la muestra para el pH 4.0 están entre 2 y 4.
- Hay una proporción mayor de valores fuera de la caja para el pH 4.0 que para el pH 4.5.

- e) Ambas muestras están sesgadas a la derecha.
f) Ambas muestras contienen datos atípicos.

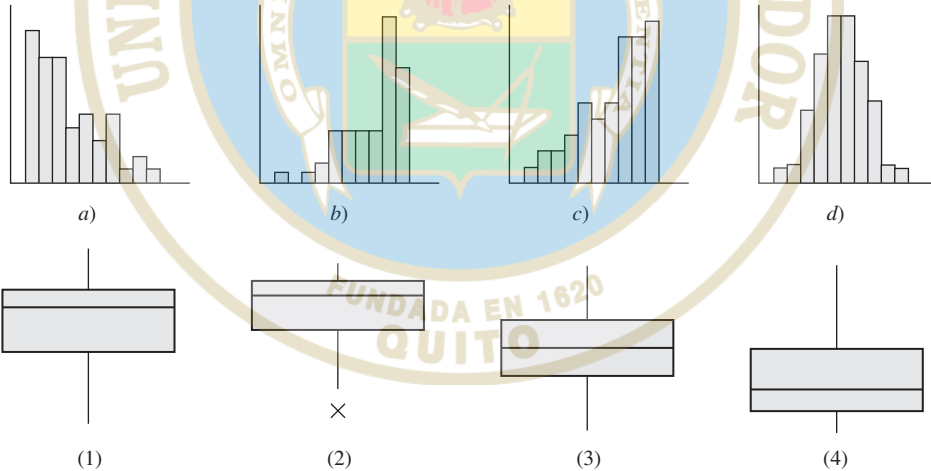


11. Los siguientes son resúmenes estadísticos para dos conjuntos de datos, A y B.

	A	B
Mínimo	0.066	-2.235
1er. cuartil	1.42	5.27
Mediana	2.60	8.03
3er. cuartil	6.02	9.13
Máximo	10.08	10.51

- a) Calcule los rangos intercuartiles tanto para A como para B.
b) ¿Los resúmenes estadísticos para A proporcionan suficiente información para construir un diagrama de caja? Si es así, construya el diagrama de caja. Si no, explique por qué.
c) ¿Los resúmenes estadísticos para B proporcionan la suficiente información para construir un diagrama de caja? Si es así, construya el diagrama de caja. Si no, explique por qué.

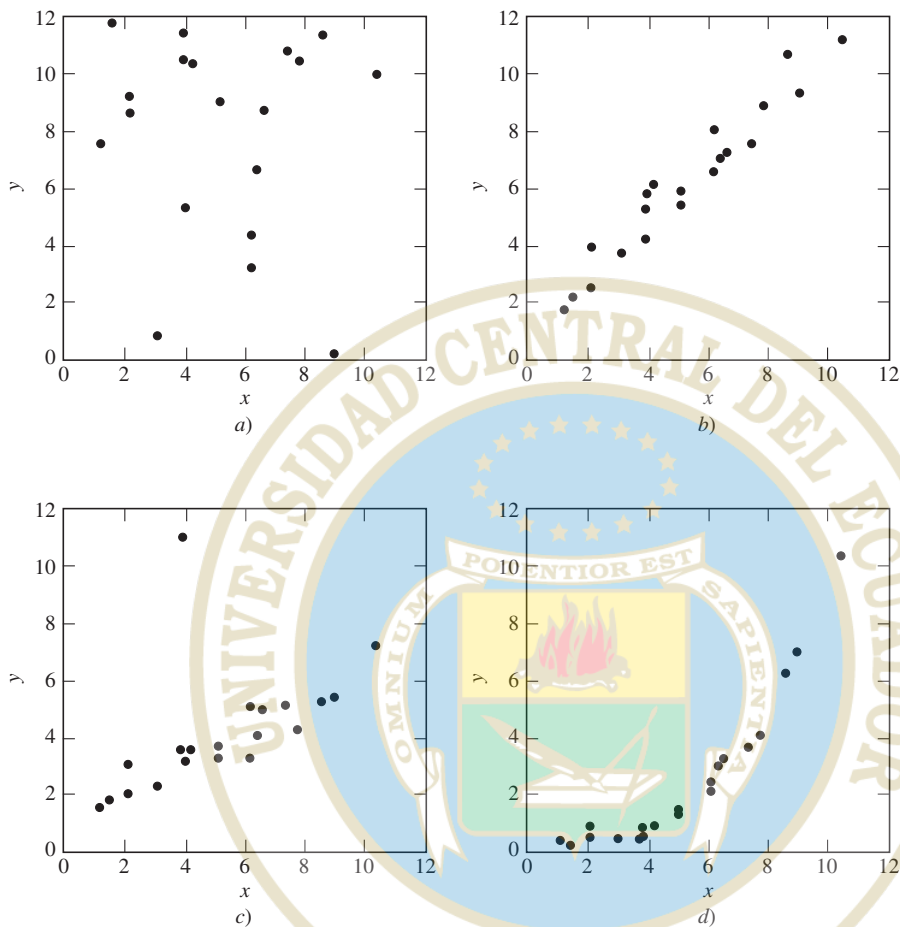
12. Relacione cada histograma con el diagrama de caja que representa el mismo conjunto de datos.



13. Refiérase a los datos de asfalto en el ejemplo 1.12 (p. 17).

- a) Construya un diagrama de caja para los datos de asfalto.
b) ¿Qué valores, si los hay, son atípicos?
c) Construya una gráfica de puntos para los datos de asfalto.
d) Con el fin de construir diagramas de caja, se define un dato atípico como un punto cuya distancia al cuartil más cercano es mayor a 1.5 IQR. Una definición más general y menos precisa es que un dato atípico es cualquier punto que está separado de la mayor parte de los datos. ¿Hay puntos en el conjunto de datos del asfalto que son datos atípicos bajo esta definición más general, pero no bajo la definición del diagrama de caja? ¿Si es así, cuáles son?

14. Relacione cada diagrama de dispersión con el enunciado que lo describa mejor.



- i) La relación entre x y y es casi lineal.
- ii) La relación entre x y y no es lineal.
- iii) No hay mucho que relacione a x con y .
- iv) La relación entre x y y es casi lineal, menos un dato atípico.

15. Para los siguientes datos:

x	1.4	2.4	4.0	4.9	5.7	6.3	7.8	9.0	9.3	11.0
y	2.3	3.7	5.7	9.9	6.9	15.8	15.4	36.9	34.6	53.2

- a) Realice una gráfica de dispersión de y contra x . ¿Es la relación entre x y y casi lineal, o no lo es?
- b) Calcule el logaritmo natural de cada valor de y . Esto se conoce como hacer una *transformación de log de y*. Realice una gráfica de dispersión de $\ln y$ contra x . ¿La relación entre x y $\ln y$ es casi lineal, o no lo es?
- c) En general, es más fácil trabajar con cantidades que tienen una relación lineal aproximada que con cantidades que no tienen una relación lineal. ¿Para estos datos, pensaría que es más fácil trabajar con x y y o con x y $\ln y$? Explique.

Ejercicios adicionales para el capítulo 1

1. Una vendedora convierte los pesos de los paquetes que vende de libras a kilogramos ($1 \text{ kg} \approx 2.2 \text{ lb}$).
 - a) ¿Cómo afecta esto la media del peso de los paquetes?
 - b) ¿Cómo afecta esto la desviación estándar de los pesos?
2. Con referencia al ejercicio 1. La vendedora empieza a usar un empaquetado más pesado, lo que aumenta el peso de cada paquete en 50 gramos.
 - a) ¿Cómo afecta esto a la media del peso de los paquetes?
 - b) ¿Cómo afecta esto a la desviación estándar de los pesos?
3. Los circuitos integrados constan de canales eléctricos que están grabados en placas de silicio. Cierta proporción de circuitos está defectuosa debido al “corte sesgado”, lo que sucede cuando se graba demasiado material lejos de los canales, ello da como resultado partes no grabadas en las placas, que son demasiado angostas. Se está investigando un proceso rediseñado que implica presión más baja en la cámara de grabado. El objetivo es reducir la tasa de corte sesgado a menos de 5%. De los primeros 100 circuitos fabricados por el nuevo proceso, solamente cuatro indican evidencias de corte sesgado. Verdadero o falso:
 - a) Debido a que solamente 4% de los 100 circuitos tenían corte sesgado, se puede concluir que se ha alcanzado el objetivo.
 - b) Aunque los porcentajes de la muestra están debajo de 5%, esto podría representar la variación del muestreo, así que es posible que el objetivo aún no se ha alcanzado.
 - c) No hay caso en evaluar el nuevo proceso, porque no importa cuál es el resultado, podía sólo atribuirse a la variación del muestreo.
 - d) Si muestreamos un número lo suficientemente grande de circuitos y si los porcentajes de circuitos defectuosos están suficientemente por debajo de 5%, entonces es razonable concluir que se ha alcanzado el objetivo.
4. Se lanza una moneda dos veces y cae “cara” ambas veces. Alguien dice: “Hay algo deshonesto en esta moneda. Se supone que una moneda cae cara sólo la mitad de las veces y no siempre.”
 - a) ¿Es razonable concluir que la moneda es deshonest? Explique.
 - b) Si al lanzar la moneda caen 100 “caras” consecutivas, ¿es razonable concluir que la moneda es deshonest? Explique.
5. El número más pequeño en una lista se cambia de 12.9 a 1.29.
 - a) ¿Es posible determinar cuánto cambia la media? Si es así, ¿en cuánto cambia?
 - b) ¿Es posible determinar cuánto cambia la mediana? Si es así, ¿en cuánto cambia? ¿Qué pasa si la lista consta de solamente dos números?
 - c) ¿Es posible determinar cuánto cambia la desviación estándar? Si es así, ¿cuánto cambia?
6. Hay 15 números en una lista y el número más pequeño se cambia de 12.9 a 1.29.
 - a) ¿Es posible determinar cuánto cambia la media? Si es así, ¿cuánto cambia?
 - b) ¿Es posible determinar el valor de la media después del cambio? Si es así, ¿cuál es el valor?
 - c) ¿Es posible determinar cuánto cambia la mediana? Si es así, ¿cuánto cambia?
 - d) ¿Es posible determinar cuánto cambia la desviación estándar? Si es así, ¿cuánto cambia?
7. Hay 15 números en una lista y la media es 25. El número más pequeño en la lista se cambia de 12.9 a 1.29.
 - a) ¿Es posible determinar cuánto cambia la media? Si es así, ¿cuánto cambia?
 - b) ¿Es posible determinar el valor de la media después del cambio? Si es así, ¿cuál es el valor?
 - c) ¿Es posible determinar cuánto cambia la mediana? Si es así, ¿cuánto cambia?
 - d) ¿Es posible determinar cuánto cambia la desviación estándar? Si es así, ¿cuánto cambia?
8. El artículo “The Selection of Yeast Strains for the Production of Premium Quality South African Brandy Base Products” (C. Steger y M. Lambrechts, *Journal of Industrial Microbiology and Biotechnology*, 2000:431-440) presenta información detallada acerca de la composición compuesta volátil de la base de vinos hechos para cada una de las 16

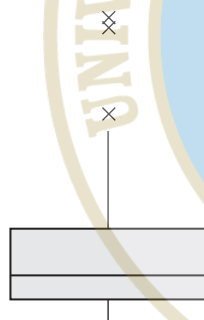
clases de levadura seleccionadas. Las siguientes son las concentraciones de ésteres totales (en mg/L) en cada uno de los vinos.

284.34	173.01	229.55	312.95	215.34	188.72
144.39	172.79	139.38	197.81	303.28	256.02
658.38	105.14	295.24	170.41		

- Calcule la media de la concentración.
- Calcule la mediana de la concentración.
- Calcule el primer cuartil de las concentraciones.
- Calcule el tercer cuartil de las concentraciones.
- Construya un diagrama de caja para las concentraciones. ¿Qué características revela?

9. Con respecto a los datos que se representan en el siguiente diagrama de caja, ¿cuál de los enunciados siguientes es verdadero?

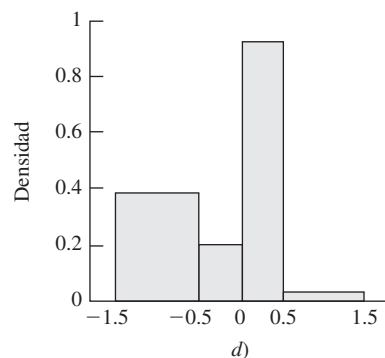
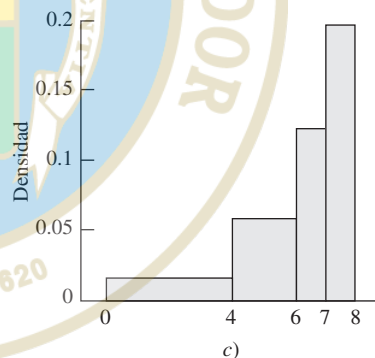
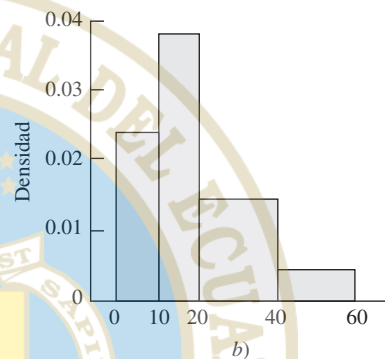
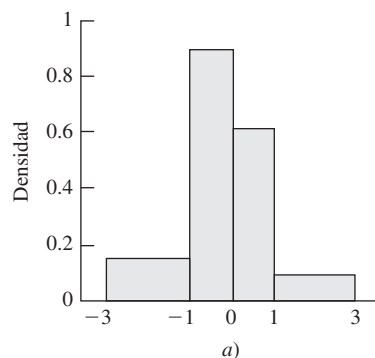
- La media es más grande que la mediana.
- La media es menor que la mediana.
- La media es aproximadamente igual a la mediana.



10. Verdadero o falso: en cualquier diagrama de caja,

- La longitud de los “bigotes” es igual a 1.5 IQR, donde el IQR es el rango intercuartil.
- La longitud de los “bigotes” podría ser más grande que 1.5 IQR, donde el IQR es el rango intercuartil.
- La longitud de los “bigotes” podría ser menor que 1.5 IQR, donde el IQR es el rango intercuartil.
- Los valores en los extremos de los “bigotes” son siempre valores del conjunto de datos que se usan para construir el diagrama de caja.

11. Para cada uno de los histogramas siguientes, determine si el eje vertical ha sido etiquetado correctamente.



12. En el artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, *Water Environment Research*, 2002:177-186), se midieron concentraciones de amonio (en mg/L) en un total de 349 pozos aluviales en el estado de Iowa. La media de la concentración fue de 0.27, la mediana fue de 0.10 y la desviación estándar fue de 0.40. Si se dibujara un histograma con estas 349 mediciones,

- i) estaría sesgado a la derecha.
- ii) estaría sesgado a la izquierda.
- iii) sería aproximadamente simétrico.
- iv) su forma no se podría determinar sin conocer las frecuencias relativas.

13. En el artículo “Vehicle-Arrival Characteristics at Urban Uncontrolled Intersections” (V. Rengaraju y V. Rao, *Journal of Transportation Engineering*, 1995:317-323) se presentan los datos sobre las características de tráfico en diez cruces en Madras, India. Una de las características medidas fueron las velocidades de los vehículos que se desplazaban por los cruces. La tabla siguiente da 15avo., 50avo. y 85avo. percentiles de la velocidad (en km/h) para dos cruces.

Intersección	Percentil		
	15avo.	50avo.	85avo.
A	27.5	37.5	40.0
B	24.5	26.5	36.0

- a) Si se hubiera dibujado un histograma para las velocidades de vehículos que pasan a través de la intersección A, ¿pensaría que está sesgado a la izquierda, sesgado a la derecha, o casi simétrico? Explique.
- b) Si se hubiera dibujado un histograma para las velocidades de vehículos que pasan a través de la intersección B, ¿pensaría que está sesgado a la izquierda, sesgado a la derecha, o casi simétrico? Explique.

14. La frecuencia acumulada y la frecuencia relativa acumulada para un intervalo de clase dado son las sumas de las frecuencias y las frecuencias relativas, respectivamente, sobre todas las clases, incluyendo la clase dada. Por ejemplo, si hay cinco clases, con frecuencias 11, 7, 3, 14 y 5, las frecuencias acumulativas serían 11, 18, 21, 35 y 40, y las frecuencias relativas acumulativas serían 0.275, 0.450, 0.525, 0.875 y 1.000. Construya una tabla para presentar las frecuencias, las frecuencias relativas y las frecuencias relativas acumuladas, para los datos del ejercicio 1 de la sección 1.3, usando los intervalos de clase $11 < 12$, $12 < 13$, ..., $20 < 21$.

15. El artículo “Computing and Using Rural versus Urban Measures in Statistical Applications” (C. Goodall, K. Kafadar y J. Tukey, *The American Statistician*, 1998:101-111) analiza los métodos para medir el grado a los cuales los condados de los Estados Unidos son urbanos más que rurales. La siguiente tabla de frecuencias presenta las frecuencias de población de los condados de los Estados Unidos. Las poblaciones están en la escala $\log_2$; por consiguiente, el primer intervalo contiene condados cuyas poblaciones son al menos de $2^6 = 64$, pero menores que $2^{12.4} = 5\,404$ y así sucesivamente.

$\log_2$ de la población	Núm. de condados
6.0–< 12.4	305
12.4–< 13.1	294
13.1–< 13.6	331
13.6–< 14.0	286
14.0–< 14.4	306
14.4–< 14.8	273
14.8–< 15.3	334
15.3–< 16.0	326
16.0–< 17.0	290
17.0–< 23.0	323

- a) Construya un histograma a partir de la tabla de frecuencias.
- b) Calcule la proporción de condados cuyas poblaciones son mayores a cien mil.
- c) ¿Está el histograma sesgado a la izquierda, a la derecha, o casi simétrico?
- d) Construya un histograma usando las poblaciones reales en vez de su logaritmo. ¿Por qué piensa usted que el artículo transformó las poblaciones a la escala logarítmica?

16. El artículo “Hydrogeochemical Characteristics of Groundwater in a Mid-Western Coastal Aquifer System” (S. Jeon, J. Kim y colaboradores, *Geosciences Journal*, 2001:339-348) presenta las mediciones de las propiedades de algunas aguas subterráneas poco profundas en cierto sistema de acuífero de Corea. Las siguientes son las mediciones de la conductividad eléctrica (en microsiemens por centímetro) para 23 muestras de agua.

2099	528	2030	1350	1018	384	1499
1265	375	424	789	810	522	513
488	200	215	486	257	557	260
461	500					

- a) Encuentre la media.
- b) Encuentre la desviación estándar.

- c) Encuentre la mediana.
d) Construya una gráfica de puntos.
e) Encuentre la media recortada a 10%.
f) Encuentre el primer cuartil.
g) Encuentre el tercer cuartil.
h) Encuentre el rango intercuartil.
i) Construya un diagrama de caja.
j) ¿Cuáles puntos, si los hay, son atípicos?
- k) Si se hubiera construido un histograma, ¿estaría sesgado a la izquierda, sesgado a la derecha, o casi simétrico?
17. La falta de agua ha sido tradicionalmente una preocupación muy importante en las Islas Canarias. Los derechos sobre el agua se dividen en acciones, que son posesión privada. El artículo “The Social Construction of Scarcity. The Case of Water in Tenerife (Canary Islands)” (F. Aguilera-Klink, E. Pérez-Moriana, y J. Sánchez-García, *Ecological Economics*, 2000:233-245) analiza que la extensión de muchas de las acciones están concentradas entre pocos propietarios. La tabla siguiente presenta el número de propietarios que poseen diferentes números de acciones. (Había 15 propietarios que poseían 50 acciones o más; que se han omitido.) Observe que es posible poseer un número no entero de acciones; por ejemplo, el intervalo $2- < 3$ contiene 112 personas que poseían al menos dos pero menos de tres acciones.
18. El informe del editor, en noviembre de 2003, de la edición de *Technometrics* proporciona la siguiente información respecto de la cantidad de tiempo que le toma revisar artículos que fueron presentados para su publicación durante 2002. Para propósitos computacionales, interprete la última categoría (> 9) como $9 - < 15$.

Número de acciones	Número de propietarios
$0- < 1$	18
$1- < 2$	165
$2- < 3$	112
$3- < 4$	87
$4- < 5$	43
$5- < 10$	117
$10- < 15$	51
$15- < 20$	32
$20- < 25$	10
$25- < 30$	8
$30- < 50$	8

Tiempo (meses)	Número de artículos
$0- < 1$	45
$1- < 2$	17
$2- < 3$	18
$3- < 4$	19
$4- < 5$	12
$5- < 6$	14
$6- < 7$	13
$7- < 8$	22
$8- < 9$	11
> 9	18

- a) Construya un histograma para estos datos.
b) Aproxime la mediana del número de acciones poseídas encontrando el punto para el cual las áreas de cualquier lado son iguales.
c) Aproxime el primer cuartil del número de acciones poseídas encontrando el punto para el cual 25% del área está a la izquierda.
- a) Construya un histograma para estos datos.
b) Aproxime la mediana del tiempo de revisión encontrando el punto para el cual las áreas en cualquier lado son iguales.
c) Aproxime el primer cuartil de los tiempos de revisión, encontrando el punto para el cual 25% del área está a la izquierda.
d) Aproxime el tercer cuartil del tiempo de revisión encontrando el punto para el que 75% del área está a la izquierda.
e) Aproxime la media del tiempo de revisión calculando el centro de masa del histograma.
f) Aproxime la varianza del tiempo de revisión, calculando el momento de inercia con respecto a la media, suponiendo que toda la masa de un rectángulo está concentrada en su punto medio.

19. En el artículo “The Ball-on-Three-Ball Test for Tensile Strength: Refined Methodology and Results for Three Hohokam Ceramic Types” (M. Beck, *American Antiquity*, 2002:558-569) se analiza la resistencia de la cerámica antigua. Se evaluaron algunas muestras de cada uno de tres tipos de cerámica. Las cargas (en kg) requeridas para romper las muestras son las que aparecen en la tabla de la siguiente columna.

- Construya diagramas de caja comparativos para las tres muestras.
- ¿Cuántos datos atípicos contiene cada muestra?
- Comente acerca de las características de las tres muestras.

Tipo de cerámica	Cargas (kg)
Sacaton	15, 30, 51, 20, 17, 19, 20, 34, 17, 15, 23, 19, 15, 18, 16, 22, 29, 15, 13, 15
Planicie del río Gila	27, 18, 28, 25, 55, 21, 18, 34, 23, 30, 20, 30, 31, 25, 28, 26, 17, 19, 16, 24, 19, 9, 31, 19, 27, 20, 45, 15
Casa Grande	20, 16, 20, 36, 27, 35, 66, 15, 18, 24, 21, 30, 20, 24, 23, 21, 13, 21



Capítulo 2

Probabilidad

Introducción

El desarrollo de la teoría de la probabilidad fue financiada por apostadores en el siglo XVII, quienes contrataron a algunos matemáticos famosos para que calculasen la probabilidad correcta de ciertos juegos de azar. Con el tiempo, la gente se dio cuenta de que los procesos científicos también son azarosos y desde entonces se han empleado métodos de probabilidad para estudiar el entorno físico.

Actualmente, la probabilidad constituye una gran rama de las matemáticas. Existen muchos libros al respecto y numerosos investigadores han dedicado bastante de su tiempo con el propósito de ampliar su desarrollo. En este capítulo se presenta una introducción de los conceptos de probabilidad más relevantes para el estudio de la estadística.

2.1 Ideas básicas

Para realizar un estudio sistemático de probabilidad, se necesita cierta terminología. Un **experimento** constituye un proceso con un resultado que no se puede predecir certeramente con anterioridad. El hecho de lanzar una moneda al aire, arrojar un dado, medir el diámetro de un perno, pesar los contenidos de una caja de cereal, o medir la resistencia de una cuerda de pescar, son ejemplos de experimentos. Con la finalidad de analizar un experimento en términos probabilísticos, se debe especificar sus posibles resultados.

Definición

Al conjunto de todos los posibles resultados de un experimento se le llama **espacio muestral**.

Al lanzar al aire una moneda se puede utilizar el conjunto {"caras", "cruces"} como el espacio muestral. Para arrojar un dado de seis caras, se puede usar al conjunto {1, 2, 3, 4, 5, 6}. Estos espacios muestrales son finitos. Algunos experimentos tienen espacios muestrales

con un número infinito de resultados. Por ejemplo, imagine que un buril con diámetro de 10 mm hace perforaciones en una lámina de metal. Debido a las variaciones en el ángulo de la perforación y a los pequeños movimientos en la lámina de metal, los diámetros de los agujeros varían entre 10 y 10.2 mm. Por tanto, para el experimento de perforación sería razonable un espacio muestral que esté en el intervalo (10.0, 10.2), o en notación de conjuntos, $\{x \mid 10.0 < x < 10.2\}$. Obviamente, este conjunto contiene un número infinito de resultados.

En muchos experimentos se puede escoger entre diversos espacios muestrales. Por ejemplo, suponga un proceso que produce clavos de acero cuyas longitudes varían entre 5.20 y 5.25 cm. Una opción obvia para el espacio muestral de la longitud de un clavo sería el conjunto $\{x \mid 5.20 < x < 5.25\}$. Sin embargo, si el objetivo fuera simplemente determinar si el clavo es demasiado corto, demasiado largo o está dentro de ciertos límites específicos, una buena elección sería que el espacio muestral fuera {demasiado corto, demasiado largo, dentro de las especificaciones}.

Con frecuencia, al estudiar experimentos, se está interesado en un subconjunto particular de resultados. Por ejemplo, se puede tener interés en la probabilidad de que un dado caiga en un número par. El espacio muestral para el experimento es $\{1, 2, 3, 4, 5, 6\}$ y el correspondiente a que caiga en un número par es el subconjunto $\{2, 4, 6\}$. En el ejemplo del buril usado para perforar, se puede tener interés en la probabilidad de que un hueco tenga un diámetro menor a 10.1 mm. Esto último corresponde al subconjunto $\{x \mid 10.0 < x < 10.1\}$. Existe un nombre especial para el subconjunto de un espacio muestral:

Definición

Un subconjunto de un espacio muestral se denomina **evento**.

Observe que para cualquier espacio muestral, el conjunto vacío $\emptyset$ es un evento, como lo es todo el espacio muestral. Se dice que un evento ocurrió si el resultado del experimento es alguno de los resultados en el evento. Por ejemplo, si un dado cae en el número 2, habrán ocurrido los eventos $\{2, 4, 6\}$ y $\{1, 2, 3\}$, junto con cualquier otro evento que contenga el resultado “2”.

Ejemplo

2.1

Un ingeniero eléctrico tiene en su mano dos cajas de resistores, cada una con cuatro de éstos. Los resistores de la primera caja están etiquetados con 10 Ω (ohms), pero, de hecho, sus resistencias son de 9, 10, 11 y 12 Ω . Los resistores de la segunda caja tienen la etiqueta de 20 Ω , pero sus resistencias son de 18, 19, 20 y 21 Ω . El ingeniero elige un resistor de cada caja y determina la resistencia de cada uno.

Sea A el evento para el cual el primer resistor tiene una resistencia mayor a 10, sea B el evento en el que el segundo resistor tiene una resistencia menor a 19 y sea C el evento en el cual la suma de las resistencias es igual a 28. Determine un espacio muestral para este experimento y especifique los subconjuntos que corresponden a los eventos A , B y C .

Solución

Un buen espacio muestral es el conjunto de pares ordenados en el que el primer elemento representa la resistencia del primer resistor y el segundo elemento constituye la del segundo resistor. Se denotará a este espacio muestral con $\mathcal{S}$.

$$\mathcal{S} = \{(9, 18), (9, 19), (9, 20), (9, 21), (10, 18), (10, 19), (10, 20), (10, 21), \\ (11, 18), (11, 19), (11, 20), (11, 21), (12, 18), (12, 19), (12, 20), (12, 21)\}$$

Los eventos A , B y C están dados por

$$A = \{(11, 18), (11, 19), (11, 20), (11, 21), (12, 18), (12, 19), (12, 20), (12, 21)\} \\ B = \{(9, 18), (10, 18), (11, 18), (12, 18)\} \\ C = \{(9, 19), (10, 18)\}$$

Combinación de eventos

Con frecuencia se construyen eventos mediante la combinación de eventos más sencillos. Debido a que aquéllos son subconjuntos de espacios muestrales, es usual emplear la notación de conjuntos para describir los eventos **construidos** de esta forma. A continuación se repasará la notación necesaria.

- La **unión** de dos eventos A y B , se denota por $A \cup B$, es el conjunto de resultados que pertenecen ya sea a A o B , o a ambos. Esto es, $A \cup B$ significa “ A o B ”. Por tanto, el evento $A \cup B$ se presenta siempre que ocurre A o B (o ambos).
- La **intersección** de dos eventos A y B se denota como $A \cap B$; es decir, constituye el conjunto de resultados que pertenece tanto a A como a B . Por consecuencia, $A \cap B$ significa “ A y B ”. Por consiguiente, el evento $A \cap B$ se presenta siempre que A y B ocurren.
- El **complemento** de un evento A se denota por A^c , es el conjunto de resultados que no pertenecen a A . Es decir, A^c significa “no A ”. Por consiguiente, el evento A^c se presenta siempre que *no* ocurra A .

Los eventos se pueden **mostrar gráficamente** con los diagramas de Venn. La figura 2.1 muestra los eventos $A \cup B$, $A \cap B$ y $B \cap A^c$.

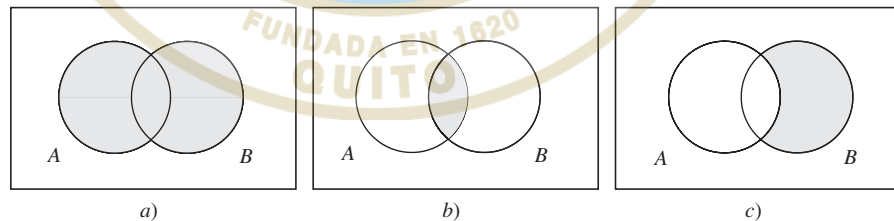


FIGURA 2.1 Diagramas de Venn que muestran varios eventos: a) $A \cup B$, b) $A \cap B$, c) $B \cap A^c$.

Ejemplo 2.2

Con referencia al ejemplo 2.1, determine a $B \cup C$ y $A \cap B^c$.

Solución

El evento $B \cup C$ contiene todos los resultados que pertenecen a B o a C , o a ambos. Por tanto,

$$B \cup C = \{(9, 18), (10, 18), (11, 18), (12, 18), (9, 19)\}$$

El evento B^c contiene los resultados en el espacio muestral que no pertenecen a B . De ahí que el evento $A \cap B^c$ contenga los resultados que pertenecen a A y no pertenecen a B . Por consiguiente,

$$A \cap B^c = \{(11, 19), (11, 20), (11, 21), (12, 19), (12, 20), (12, 21)\}$$

Eventos mutuamente excluyentes

Existen ciertos eventos que nunca se presentan simultáneamente. Por ejemplo, es imposible que una moneda que se arroje al aire caiga a la vez en “cruz” y “cara”, al igual que un clavo de acero sea al mismo tiempo demasiado largo y corto. A eventos de este tipo se les llama **mutuamente excluyentes**.

Definición

- Se dice que los eventos A y B son **mutuamente excluyentes** si no tienen resultados en común.
- De forma más general, se dice que una colección de eventos $A_1, A_2, \dots, A_n$ es mutuamente excluyente si dos de ellos no tienen resultados en común.

El diagrama de Venn en la figura 2.2 muestra eventos mutuamente excluyentes.

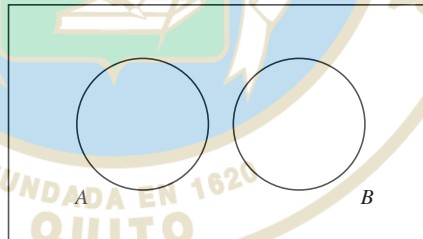


FIGURA 2.2 Los eventos A y B son mutuamente excluyentes.

Ejemplo

2.3

Con referencia al ejemplo 2.1, si se realiza el experimento, ¿es posible que los eventos A y B ocurran al mismo tiempo? ¿Qué pasa con B y C ? ¿ A y C ? ¿Qué par de eventos es mutuamente excluyente?

Solución

Si el resultado es $(11, 18)$ o $(12, 18)$, entonces tanto el evento A como el B ocurren. Si el resultado es $(10, 18)$, entonces ocurren los eventos B y C . Es imposible que ocurran al mismo tiempo A y C , ya que estos eventos son mutuamente excluyentes al no tener ningún resultado en común.

Probabilidades

Todo evento en un espacio muestral tiene una **probabilidad** de ocurrir. Intuitivamente, la probabilidad es una medida cuantitativa de qué tan probable es que ocurra un evento. Formalmente hablando, hay varias interpretaciones de la probabilidad; la primera que se adoptará es que la probabilidad de un evento representa la proporción de veces que se presentaría el evento a largo plazo, si el experimento se repitiera una y otra vez.

Con frecuencia se usa la letra P para representar la probabilidad. Por tanto, cuando se lanza una moneda al aire la notación " $P(\text{"cara"}) = 1/2$ " significa que la probabilidad de que la moneda caiga en "cara" es igual a $1/2$.

Resumen

Dado un experimento y cualquier evento A :

- La expresión $P(A)$ denota la probabilidad de que ocurra el evento A .
- $P(A)$ constituye la proporción de veces que se presenta el evento A en el tiempo, si es que el experimento se realizara una y otra vez.

En muchas situaciones, la única forma de calcular la probabilidad de un evento es repetir el experimento muchas veces y determinar la proporción de veces que ocurre. Por ejemplo, si se deseara calcular la probabilidad de que un tablero de circuitos impresos fabricado por cierto proceso esté defectuoso, usualmente se necesitaría producir cierta cantidad de tableros y probarlos para determinar la proporción de los defectuosos. En algunos casos, las probabilidades se pueden determinar si se conoce la naturaleza física del experimento. Por ejemplo, si se sabe que la forma de un dado es casi igual a la de un cubo perfecto y que su masa está distribuida aproximadamente en forma homogénea, se puede suponer que cada una de sus seis caras tiene la misma probabilidad de salir cuando se lanza el dado.

Una vez que se han encontrado las probabilidades de ciertos eventos mediante el conocimiento científico o la experiencia, se puede calcular matemáticamente las probabilidades de otros eventos. Por ejemplo, si se ha calculado a través de la experimentación que la probabilidad de que un tablero de circuitos impresos esté defectuoso es de 0.10, se puede calcular que la probabilidad de que un tablero no esté defectuoso es de 0.90. Como otro ejemplo, suponga que los clavos de acero producidos por determinado proceso no cumplen con la longitud especificada, ya sea porque son demasiado cortos o demasiado largos. Al medir gran cantidad de clavos, se calculó que la probabilidad de que uno de ellos sea demasiado corto es de 0.02 y que la probabilidad de que otro sea demasiado largo es de 0.03. Entonces puede calcularse que la probabilidad de que un clavo no cumpla con la especificación es de 0.05.

En la práctica, los científicos e ingenieros calculan las probabilidades de ciertos eventos con base en el conocimiento científico y la experiencia, y posteriormente utilizan reglas matemáticas para calcular las estimaciones de las probabilidades de otros eventos. En el resto de esta sección y en la sección 2.2, se explican algunas de estas reglas y se muestra cómo utilizarlas.

Axiomas de la probabilidad

El tema de la probabilidad se basa en tres reglas de sentido común, conocidas como axiomas. Éstas son:

Axiomas de la probabilidad

1. Sea $\mathcal{S}$ un espacio muestral. Entonces $P(\mathcal{S}) = 1$.
2. Para cualquier evento A , $0 \leq P(A) \leq 1$.
3. Si A y B son eventos mutuamente excluyentes, entonces $P(A \cup B) = P(A) + P(B)$.
De forma más general, si $A_1, A_2, \dots$ son eventos mutuamente excluyentes, entonces $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$

Si se piensa un poco, es fácil ver que los tres axiomas en realidad concuerdan con el sentido común. El primero establece que el resultado de un experimento siempre está en el espacio muestral. Esto es obvio, puesto que, por definición, el espacio muestral contiene todos los resultados posibles del experimento. El segundo dice que la frecuencia a largo plazo de cualquier evento siempre se encuentra entre 0 y 100%. Un ejemplo que demuestra al tercer axioma, que ya se analizó anteriormente, es el del proceso que produce clavos de acero, en donde la probabilidad de que un clavo sea demasiado corto es de 0.02 y la de que un clavo es demasiado largo es de 0.03. El tercer axioma establece que la probabilidad de que el clavo sea demasiado corto o muy largo es $0.02 + 0.03 = 0.05$.

Ahora se presentan dos reglas sencillas que son útiles para calcular probabilidades. Estas reglas son intuitivamente obvias y también pueden comprobarse a través de los axiomas. Las demostraciones se encuentran al final de la sección.

Para cualquier evento A ,

$$P(A^c) = 1 - P(A) \quad (2.1)$$

Si $\emptyset$ denota el espacio vacío, entonces

$$P(\emptyset) = 0 \quad (2.2)$$

La ecuación (2.1) establece que la probabilidad de que un evento no ocurra es igual a 1 menos la probabilidad de que ocurra. Por ejemplo, si existe una probabilidad de 40% de que llueva, hay una probabilidad de 60% de que no llueva. La ecuación (2.2) establece que es imposible que un experimento no tenga ningún resultado.

Ejemplo

2.4

El objetivo de una prueba de tiro consiste de un blanco rodeado por dos anillos concéntricos. Se dispara un proyectil hacia el objetivo. La probabilidad de que pegue en el blanco es de 0.10, la de que atine en el anillo interior es de 0.25 y la de que acierte en el anillo exterior es de 0.45. ¿Cuál es la probabilidad de que el proyectil haga blanco en el objetivo? ¿Cuál es la probabilidad de que no pegue en este último?

Solución

Pegar en el blanco, acertar en el anillo interior y atinar en el anillo exterior son eventos mutuamente excluyentes, ya que es imposible que más de uno de éstos ocurra a la vez. Por tanto, utilizando el axioma 3,

$$\begin{aligned} P(\text{pega en el objetivo}) &= P(\text{blanco}) + P(\text{anillo interior}) + P(\text{anillo exterior}) \\ &= 0.10 + 0.25 + 0.45 \\ &= 0.80 \end{aligned}$$

Ahora se puede calcular la probabilidad de que el proyectil no pegue en el objetivo utilizando la ecuación (2.1):

$$\begin{aligned} P(\text{no pega en el objetivo}) &= 1 - P(\text{pega en el objetivo}) \\ &= 1 - 0.80 \\ &= 0.20 \end{aligned}$$

Ejemplo 2.5

La siguiente tabla presenta las probabilidades del número de veces en que el sistema de cierta computadora se “caerá” en el transcurso de una semana. Sea A el evento de que haya más de dos “caídas” durante la semana, y B el evento de que el sistema se “caerá” por lo menos una vez. Determine el espacio muestral. Después precise los subconjuntos del espacio muestral que corresponden a los eventos A y B . Posteriormente determine $P(A)$ y $P(B)$.

Número de casos	Probabilidad
0	0.60
1	0.30
2	0.05
3	0.04
4	0.01

Solución

El espacio muestral del experimento es el conjunto $\{0, 1, 2, 3, 4\}$. Los eventos son $A = \{3, 4\}$ y $B = \{1, 2, 3, 4\}$. Para encontrar $P(A)$, advierta que A constituye el evento en que se presenten tres “caídas” o que haya cuatro “caídas”. Los eventos “que se presenten tres caídas” y “que ocurran cuatro caídas” son mutuamente excluyentes. Por tanto, mediante el axioma tres, se concluye que

$$\begin{aligned} P(A) &= P(\text{ocurran tres “caídas” o se presenten cuatro “caídas”}) \\ &= P(\text{ocurran tres “caídas”}) + P(\text{sucedan cuatro “caídas”}) \\ &= 0.04 + 0.01 \\ &= 0.05 \end{aligned}$$

Se calculará $P(B)$ en dos formas. Primero, observe que B^c es el evento de que no haya ninguna caída. Por tanto, utilizando la ecuación (2.1),

$$\begin{aligned} P(B) &= 1 - P(B^c) \\ &= 1 - P(\text{no ocurran “caídas”}) \\ &= 1 - 0.60 \\ &= 0.40 \end{aligned}$$

En la segunda forma para calcular $P(B)$, observe que B es el evento de que haya una “caída” o de que sucedan dos o que ocurran tres “caídas” o haya cuatro de éstas. Estos eventos son mutuamente excluyentes. Por consiguiente, utilizando el axioma tres, se concluye que

$$\begin{aligned} P(B) &= P(\text{una “caída”}) + P(\text{dos “caídas”}) + P(\text{tres “caídas”}) + P(\text{cuatro “caídas”}) \\ &= 0.30 + 0.05 + 0.04 + 0.01 \\ &= 0.40 \end{aligned}$$

En el ejemplo 2.5 se calcularon las probabilidades de los eventos $A = \{3, 4\}$ y $B = \{1, 2, 3, 4\}$ al sumar las probabilidades de los resultados de cada uno de los eventos: $P(A) = P(3) + P(4)$ y $P(B) = P(1) + P(2) + P(3) + P(4)$. En general, este método funciona. Dado que cualesquiera dos resultados en un espacio muestral son mutuamente excluyentes, la probabilidad de que cualquier evento contenga un número finito de resultados se puede determinar mediante la suma de las probabilidades de los resultados que incluyen al evento.

Si A es un evento y $A = \{E_1, E_2, \dots, E_n\}$, entonces

$$P(A) = P(E_1) + P(E_2) + \dots + P(E_n) \quad (2.3)$$

Espacios muestrales con resultados igualmente probables

En algunos experimentos se puede construir un espacio muestral en el cual todos los resultados sean igualmente probables. Un ejemplo sencillo es el lanzamiento de un dado, en el cual el espacio muestral es $\{1, 2, 3, 4, 5, 6\}$ y cada uno de estos resultados tiene una probabilidad de $1/6$. Otro tipo de experimento que tiene resultados igualmente probables es la selección aleatoria de un elemento a partir de una población de elementos. Se puede suponer que los elementos en la población son los resultados en un espacio muestral y que cada elemento tiene la misma probabilidad de ser seleccionado.

Una población a partir de la cual se muestrea un elemento en forma aleatoria constituye un espacio muestral con resultados igualmente probables.

Si un espacio muestral contiene N resultados igualmente probables, la probabilidad de cada resultado es $1/N$. Esto es así porque la probabilidad de todo el espacio muestral debe ser 1 y esta probabilidad se divide equitativamente entre los N resultados. Si A representa un evento que contiene k resultados, entonces se puede determinar $P(A)$ mediante la suma de las probabilidades de los k resultados, de tal forma que $P(A) = k/N$.

Si S es un espacio muestral que contiene N resultados igualmente probables y si A es un evento que contiene k resultados, entonces

$$P(A) = \frac{k}{N} \quad (2.4)$$

Ejemplo
2.6

Un troquel de extrusión se utiliza para producir varillas de aluminio. Existen ciertas especificaciones para la longitud y diámetro de las varillas. Para cada una de éstas, la longitud puede ser demasiado corta, demasiado larga o estar bien y el diámetro se puede clasificar en muy delgado, muy grueso o estar bien. En una población de mil varillas, el número de ellas en cada clase es:

Longitud	Diámetro		
	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

Se toma una varilla aleatoriamente a partir de esta población. ¿Cuál es la probabilidad de que sea demasiado corta?

Solución

Se considera que cada una de las mil varillas representa un resultado en un espacio muestral. Cada uno de los mil resultados tiene la misma probabilidad. Se resolverá el problema contando el número de resultados que corresponde al evento. El número de varillas que son demasiado cortas es $10 + 3 + 5 = 18$. Dado que el número total de varillas es mil,

$$P(\text{demasiado corta}) = \frac{18}{1\,000}$$

Regla de la suma

Si A y B son eventos mutuamente excluyentes, entonces $P(A \cup B) = P(A) + P(B)$. Esta regla se puede generalizar para abarcar el caso en el que A y B no sean mutuamente excluyentes. En el ejemplo 2.7 se muestra este razonamiento.

Ejemplo
2.7

Con referencia al ejemplo 2.6, si se selecciona aleatoriamente una varilla, ¿cuál es la probabilidad de que sea demasiado corta o demasiado gruesa?

Solución

Primero, se resolverá este problema al contar el número de resultados que corresponde al evento. En la siguiente tabla se circuló la cantidad de varillas que son demasiado gruesas y el número de varillas que son muy cortas aparecen en rectángulos. Observe que hay cinco varillas que son muy cortas y demasiado gruesas.

Longitud	Diámetro		
	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

De los mil resultados, el número de varillas que son demasiado cortas o muy gruesas es $10 + 3 + 5 + 4 + 13 = 35$. Por consiguiente,

$$P(\text{demasiado corta o muy gruesa}) = \frac{35}{1\,000}$$

Ahora se resolverá el problema de tal forma que conduzca a un método más general. En el espacio muestral hay $10 + 3 + 5 = 18$ varillas que son demasiado cortas y $5 + 4 + 13 = 22$ varillas muy gruesas. Pero si se trata de encontrar el número de varillas que sean demasiado cortas o muy gruesas al sumar $18 + 22$, se obtiene un número muy grande (40 en vez de 35). La razón es que hay cinco varillas que son demasiado cortas y muy gruesas y éstas se cuentan dos veces. No obstante, se puede resolver el problema al sumar 18 y 22, pero entonces se le debe restar cinco para corregir el doble conteo.

Se replantea este razonamiento al utilizar probabilidades:

$$\begin{aligned}
 P(\text{demasiado corta}) &= \frac{18}{1\,000}, & P(\text{muy gruesa}) &= \frac{22}{1\,000}, & P(\text{demasiado corta y muy gruesa}) &= \frac{5}{1\,000} \\
 P(\text{demasiado corta o muy gruesa}) &= & P(\text{demasiado corta}) &+ & P(\text{muy gruesa}) &- & P(\text{demasiado corta y muy gruesa}) \\
 &= & \frac{18}{1\,000} &+ & \frac{22}{1\,000} &- & \frac{5}{1\,000} \\
 &= & \frac{35}{1\,000} & & & &
 \end{aligned}$$

El método del ejemplo 2.7 es válido para cualesquiera dos eventos en un espacio muestral. En general, para determinar la probabilidad de que ocurran cualesquiera de los dos eventos, se suman las probabilidades de los eventos y después se resta la probabilidad de que ambos ocurran.

Resumen

Sean A y B cualesquiera eventos, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (2.5)$$

Una prueba de este resultado, basada en los axiomas, se encuentra al final de esta sección. Observe que si A y B son mutuamente excluyentes, entonces $P(A \cap B) = 0$, por lo que, en este caso, la ecuación (2.5) se reduce al axioma 3.

Ejemplo

2.8

En un proceso que fabrica latas de aluminio, la probabilidad de que una lata tenga alguna fisura en su costado es de 0.02, la de que otra la tenga en la tapa es de 0.03 y de que una más presente una fisura en el costado y en la tapa es de 0.01. ¿Cuál es la probabilidad de que al elegir una lata en forma aleatoria tenga una fisura? ¿Cuál es la probabilidad de que no la tenga?

Solución

Se tiene que $P(\text{fisura en el costado}) = 0.02$, $P(\text{fisura en la tapa}) = 0.03$ y $P(\text{fisura en el costado y en la tapa}) = 0.01$. Ahora $P(\text{fisura}) = P(\text{fisura en el costado o fisura en la tapa})$. Usando la ecuación (2.5),

$$\begin{aligned} P(\text{fisura en el costado o fisura en la tapa}) &= P(\text{fisura en el costado}) + P(\text{fisura en la tapa}) \\ &\quad - P(\text{fisura en el costado y fisura en la tapa}) \\ &= 0.02 + 0.03 - 0.01 \\ &= 0.04 \end{aligned}$$

Para determinar la probabilidad de que una lata no tenga ninguna fisura, se calcula

$$P(\text{ninguna fisura}) = 1 - P(\text{fisura}) = 1 - 0.04 = 0.96$$

Demostración de $P(A^c) = 1 - P(A)$

Sea $\mathcal{S}$ un espacio muestral y A un evento. Entonces A y A^c son mutuamente excluyentes; por tanto, por el axioma 3,

$$P(A \cup A^c) = P(A) + P(A^c)$$

Pero $A \cup A^c = \mathcal{S}$, y por el axioma 1, $P(\mathcal{S}) = 1$. Por consiguiente,

$$P(A \cup A^c) = P(\mathcal{S}) = 1$$

Se sigue que $P(A) + P(A^c) = 1$, entonces $P(A^c) = 1 - P(A)$.

Demostración de $P(\emptyset) = 0$

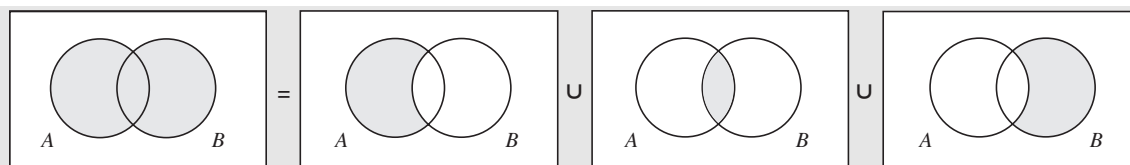
Sea $\mathcal{S}$ un espacio muestral. Entonces $\emptyset = \mathcal{S}^c$. Por consecuencia, $P(\emptyset) = 1 - P(\mathcal{S}) = 1 - 1 = 0$.

Prueba de $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Sean A y B cualesquiera dos eventos. La clave de la demostración es escribir $A \cup B$ como la unión de tres eventos mutuamente excluyentes: $A \cap B^c$, $A \cap B$ y $A^c \cap B$.

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B) \quad (2.6)$$

La siguiente figura ilustra la ecuación (2.6).



Por el axioma 3,

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) = P(A^c \cap B) \quad (2.7)$$

Ahora $A = (A \cap B^c) \cup (A \cap B)$ y $B = (A^c \cap B) \cup (A \cap B)$. Por tanto,

$$P(A) = P(A \cap B^c) + P(A \cap B) \quad (2.8)$$

y

$$P(B) = P(A^c \cap B) + P(A \cap B) \quad (2.9)$$

Al sumar las ecuaciones (2.8) y (2.9) se obtiene

$$P(A) + P(B) = P(A \cap B^c) + P(A^c \cap B) + 2P(A \cap B) \quad (2.10)$$

Al comparar las ecuaciones (2.10) y (2.7) se demuestra que

$$P(A) + P(B) = P(A \cup B) + P(A \cap B) \quad (2.11)$$

De ahí que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Ejercicios para la sección 2.1

- La probabilidad de que un microcircuito esté defectuoso es 0.08. ¿Cuál es la probabilidad de que no presente defectos?
- Un dado-octaedro (de ocho caras) tiene el número 1 pintado en dos de sus caras, el 2 en tres de sus caras, el 3 en dos de sus caras y el 4 en una cara. Se lanza el dado. Suponga que cada cara tiene la misma probabilidad de salir.
 - Determine el espacio muestral de este experimento.
 - Determine $P(\text{número par})$.
 - Si el dado estuviera cargado de tal forma que la cara con el número 4 tuviera el doble de probabilidad de salir que cada una de las otras siete caras, ¿cambiaría esto al espacio muestral? Explique.
 - Si el dado estuviera cargado de manera que la cara con el número 4 tuviera el doble de probabilidad de salir que cada una de las otras siete caras, ¿cambiaría esto el valor de $P(\text{número par})$? Explique.
- Sesenta por ciento de las grandes compras hechas a un vendedor de computadoras son PC, 30% son portátiles y 10% son accesorios, como impresoras. Como parte de una auditoría, se elige una muestra aleatoria del registro de una compra.
 - ¿Cuál es la probabilidad de que se trate de una computadora personal?
 - ¿Cuál es la probabilidad de que se trate de una computadora personal o de una portátil?
- Una unidad producida en cierto proceso tiene una probabilidad de 0.10 de que sea defectuosa. Verdadero o falso:
 - Si se toma una muestra de 100 unidades, exactamente 10 de ellas serán defectuosas.
 - Si se toma una muestra de 100 unidades, el número de unidades defectuosas será aproximadamente 10, pero no exactamente 10.
 - A medida que se toman más y más unidades, la proporción de unidades defectuosas se acercará a 10 por ciento.

5. Un ingeniero que vigila el control de calidad toma una muestra de 100 unidades fabricadas por determinado proceso y encuentra que 15 de ellas son defectuosas. Verdadero o falso.

a) La probabilidad de que una unidad fabricada por este proceso esté defectuosa es 0.15.

b) La probabilidad de que una unidad fabricada por este proceso esté defectuosa se aproxima a 0.15, pero no es exactamente igual a 0.15.
6. Un sistema contiene dos componentes, A y B . El sistema funcionará siempre y cuando A o B funcionen. La probabilidad de que A funcione es 0.95, que B funcione es 0.90 y que ambos funcionen es 0.88. ¿Cuál es la probabilidad de que el sistema funcione?
7. Un sistema contiene dos componentes, A y B . El sistema sólo funcionará si ambos funcionan. La probabilidad de que A funcione es 0.98, que B funcione es 0.95 y que A o B funcionen es 0.99. ¿Cuál es la probabilidad de que el sistema funcione?
8. El cuerpo humano puede contener uno o dos antígenos, A y B . A la sangre que contiene sólo el antígeno A se le denomina tipo A , a la que contiene sólo el B se le conoce como tipo B , a la que contiene a ambos se le llama tipo AB y a la sangre que no contiene ninguno se le denomina tipo O . En cierto banco de sangre, 35% de los donantes de sangre tiene el tipo de sangre A , 10% el tipo B y 5% el tipo AB .

a) ¿Cuál es la probabilidad de que se elija aleatoriamente a un donante de sangre de tipo O ?

b) Un receptor con sangre tipo A puede recibir sin ningún peligro de un donante sangre que no tenga el antígeno B . ¿Cuál es la probabilidad de que un donante elegido aleatoriamente pueda donar al receptor con sangre tipo A ?
9. Verdadero o falso: Si A y B son mutuamente excluyentes,

a) $P(A \cup B) = 0$

b) $P(A \cap B) = 0$

c) $P(A \cup B) = P(A \cap B)$

d) $P(A \cup B) = P(A) + P(B)$

2.2 Métodos de conteo*

Cuando se calculan probabilidades, algunas veces se necesita determinar el número de resultados en un espacio muestral. En esta sección se describirán diversos métodos con ese propósito. La regla básica, que se conoce como **principio fundamental de conteo**, se presenta por medio del ejemplo 2.9.

Ejemplo

2.9

Cierto tipo de automóvil se encuentra disponible en tres colores: rojo, azul o verde, y puede tener un motor grande o pequeño. ¿De cuántas maneras puede un comprador elegir un automóvil?

Solución

Hay tres opciones de color y dos opciones de motor. Una lista completa de las opciones se muestra en la siguiente tabla de 3×2 . El número total de opciones es $(3)(2) = 6$.

	Rojo	Azul	Verde
Grande	Rojo, grande	Azul, grande	Verde, grande
Pequeño	Rojo, pequeño	Azul, pequeño	Verde, pequeño

* Esta sección es opcional.

Al generalizar el ejemplo 2.9, si hay n_1 elecciones de color y n_2 elecciones de motor, una lista completa de elecciones se puede escribir como una tabla $n_1 \times n_2$, por lo que el número total de elecciones es $n_1 n_2$.

Si una operación se puede realizar en n_1 maneras y si para cada una de esas maneras se puede realizar una segunda operación en n_2 maneras, entonces el número total de maneras en que se realizan las dos operaciones es $n_1 n_2$.

Este razonamiento del principio fundamental del conteo de estados se puede ampliar para cualquier número de operaciones.

El principio fundamental del conteo

Suponga que se pueden realizar k operaciones. Si hay n_1 maneras de realizar la primera operación y si para cada una de esas maneras hay n_2 maneras de realizar la segunda operación y si para cada una de esas elecciones de esas maneras de realizar las dos primeras operaciones hay n_3 maneras de realizar la tercera operación y así sucesivamente, entonces el número total de maneras de realizar la secuencia de las k operaciones es $n_1 n_2 \dots n_k$.

Ejemplo

2.10

Cuando se hace un pedido de cierto tipo de computadora, hay tres elecciones de disco duro, cuatro de la cantidad de memoria, dos de la tarjeta de video y tres de monitor. ¿En cuántas maneras se puede solicitar una computadora?

Solución

El número total es $(3)(4)(2)(3) = 72$.

Permutaciones

Una **permutación** constituye un ordenamiento de un conjunto de elementos. Por ejemplo, hay seis permutaciones de las letras A, B, C: ABC, ACB, BAC, BCA, CAB y CBA. Con solamente tres elementos, es fácil determinar el número de permutaciones, sólo con hacer una lista de todas ellas. Pero con un gran número de elementos esto último no sería factible. El principio fundamental del conteo se puede usar para determinar el número de permutaciones de cualquier conjunto de elementos. Por ejemplo, se puede determinar el número de permutaciones de un conjunto de tres elementos de la siguiente manera. Hay tres elecciones para colocar el primer elemento. Después de que se hace la elección, hay dos elecciones restantes para el elemento del segundo lugar. Entonces queda una elección para el elemento del último lugar. Por tanto, el número total de maneras de ordenar tres objetos es $(3)(2)(1) = 6$. Este razonamiento se puede generalizar. El número de permutaciones de un conjunto de n elementos es

$$n(n-1)(n-2) \dots (3)(2)(1)$$

Éste es el producto de los enteros del 1 al n . Este producto se puede escribir con el símbolo $n!$, que se lee “ n factorial”.

Definición

Para cualquier entero positivo n , $n! = n(n-1)(n-2) \cdots (3)(2)(1)$.

También se define a $0! = 1$.

El número de permutaciones de n objetos es $n!$

Ejemplo

2.11

Cinco personas están en la hilera de un cine. ¿En cuántas maneras diferentes se pueden ordenar?

Solución

El número de permutaciones de un conjunto de cinco personas es $5! = (5)(4)(3)(2)(1) = 120$.

A veces se está interesado en contar el número de permutaciones de los subconjuntos de cierto tamaño elegidos de un conjunto más grande. Lo anterior se muestra en el ejemplo 2.12.

Ejemplo

2.12

Cinco salvavidas están disponibles para la guardia de un sábado por la tarde. Hay tres estaciones salvavidas. ¿De cuántas maneras se pueden elegir y organizar los tres salvavidas entre las estaciones?

Solución

Se usa el principio fundamental del conteo. Hay cinco maneras de elegir a un salvavidas para que ocupe la primera estación, luego cuatro de elegir a un salvavidas para que ocupe la segunda estación y por último tres para elegir un salvavidas que ocupe la tercera estación. El número total de permutaciones de los tres salvavidas elegidos entre los cinco es $(5)(4)(3) = 60$.

El razonamiento usado para resolver el ejemplo 2.12 se puede generalizar. El número de permutaciones de k objetos elegidos de un grupo de n objetos es

$$(n)(n-1) \cdots (n-k+1)$$

Esta expresión se puede simplificar utilizando la notación factorial:

$$\begin{aligned} (n)(n-1) \cdots (n-k+1) &= \frac{n(n-1) \cdots (n-k+1)(n-k)(n-k-1) \cdots (3)(2)(1)}{(n-k)(n-k-1) \cdots (3)(2)(1)} \\ &= \frac{n!}{(n-k)!} \end{aligned}$$

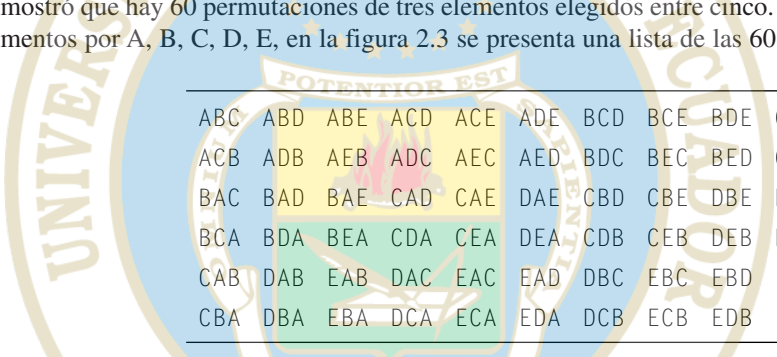
Resumen

El número de permutaciones de k objetos elegidos de un grupo de n elementos es

$$\frac{n!}{(n - k)!}.$$

Combinaciones

En algunos casos, cuando se elige un conjunto de elementos de un conjunto más grande, no se tiene en cuenta el orden de los elementos elegidos; sólo se consideran los elementos que se eligen. Por ejemplo, puede que no importe qué salvavidas ocupe cada estación; puede que sólo sea importante la elección de tres salvavidas. A cada grupo distinto de elementos que se puede seleccionar, sin importar el orden, se le llama **combinación**. A continuación se mostrará cómo determinar el número de combinaciones de k elementos elegidos de un conjunto de n objetos. Se mostrará el razonamiento con el resultado del ejemplo 2.12. En ese ejemplo se mostró que hay 60 permutaciones de tres elementos elegidos entre cinco. Al denotar a los elementos por A, B, C, D, E, en la figura 2.3 se presenta una lista de las 60 permutaciones.



ABC	ABD	ABE	ACD	ACE	ADE	BCD	BCE	BDE	CDE
ACB	ADB	AEB	ADC	AEC	AED	BDC	BEC	BED	CED
BAC	BAD	BAE	CAD	CAE	DAE	CBD	CBE	DBE	DCE
BCA	BDA	BEA	CDA	CEA	DEA	CDB	CEB	DEB	DEC
CAB	DAB	EAB	DAC	EAC	EAD	DBC	EBD	EDB	EDC
CBA	DBA	EBA	DCA	ECA	EDA	DCB	ECB	EDB	EDC

FIGURA 2.3 Las 60 permutaciones de tres elementos elegidos entre cinco.

Las 60 permutaciones de la figura 2.3 están ordenadas en diez columnas de seis permutaciones cada una. Dentro de cada columna, los tres elementos son los mismos y la columna contiene las seis permutaciones diferentes de esos tres elementos. Por tanto, cada columna representa una combinación distinta de tres elementos elegidos entre cinco y hay diez combinaciones de ese tipo. Por tanto, la figura 2.3 muestra que el número de combinaciones de tres elementos elegido entre cinco se puede encontrar al dividir el número de permutaciones de los tres elementos elegidos, o sea $5!/(5 - 3)!$, por el número de permutaciones de los tres elementos, que es $3!$. En resumen, el número de combinaciones de los tres elementos elegi-

dos es $\frac{5!}{3!(5 - 3)!}.$

Con frecuencia el número de combinaciones de k elementos elegidos de n se denota por el símbolo $\binom{n}{k}$. El razonamiento utilizado para deducir el número de combinaciones de los tres elementos elegidos se puede generalizar para deducir una expresión para $\binom{n}{k}$.

Resumen

El número de combinaciones de k elementos elegidos de un grupo de n elementos es

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (2.12)$$

Ejemplo

2.13

A cierto evento asisten 30 personas y se elegirá aleatoriamente a cinco para recibir premios. Estos últimos son iguales, así que el orden en que se elige a las personas no es importante. ¿Cuántos grupos diferentes de cinco personas se puede elegir?

Solución

En virtud de que el orden de las cinco personas elegidas no es importante, se tiene que calcular el número de combinaciones de cinco elegidas entre 30. Esto es

$$\begin{aligned} \binom{30}{5} &= \frac{30!}{5!25!} \\ &= \frac{(30)(29)(28)(27)(26)}{(5)(4)(3)(2)(1)} \\ &= 142\,506 \end{aligned}$$

Elegir una combinación de k elementos de un conjunto de n divide a los n elementos en dos subconjuntos: k que fueron elegidos y $n - k$ que no fueron elegidos. A veces un conjunto se divide en más de dos subconjuntos. Por ejemplo, suponga que en una clase de 12 estudiantes se asignará un proyecto a los estudiantes para trabajar en equipos. Se formarán tres equipos, que constarán de cinco, cuatro y tres estudiantes. Se puede calcular el número de maneras en las que se formarán los equipos del siguiente modo. Se considera el proceso para dividir la clase en tres equipos como una secuencia de dos operaciones. La primera operación es seleccionar una combinación de cinco estudiantes para formar el equipo de cinco. La segunda consiste en elegir una combinación de cuatro estudiantes entre los siete restantes, para formar el equipo de cuatro. El equipo de tres automáticamente constará de los tres estudiantes que quedan.

El número de maneras de realizar la primera operación es

$$\binom{12}{5} = \frac{12!}{5!7!}$$

Después de que se ha realizado la primera operación, el número de las maneras de realizar la segunda operación es

$$\binom{7}{4} = \frac{7!}{4!3!}$$

Por tanto, el número total de maneras de realizar la secuencia de las dos operaciones es

$$\frac{12!}{5!7!} \frac{7!}{4!3!} = \frac{12!}{5!4!3!} = 27\,720$$

Observe que el numerador en la respuesta final es el factorial del tamaño total del grupo, mientras que el denominador constituye el producto de los factoriales de los tamaños de los equipos elegidos de éste. Esto último es válido en general.

Resumen

El número de maneras de dividir un grupo de n elementos en grupos de $k_1, \dots, k_r$ elementos, donde $k_1 + \dots + k_r = n$, es

$$\frac{n!}{k_1! \dots k_r!} \quad (2.12)$$

Ejemplo

2.14

Se lanza un dado 20 veces. En virtud de que en tres de las tiradas salió el número 1, en cinco el 2, en cuatro el 3, en dos el 4 y en tres el 6, ¿cuántos arreglos diferentes de resultados hay?

Solución

Hay 20 resultados. Están divididos en seis grupos; a saber, el grupo de tres resultados en los que salió 1, el de cinco en los que salió 2 y así sucesivamente. El número de maneras de dividir los 20 resultados en seis grupos de tamaños específicos es

$$\frac{20!}{3!5!4!2!3!3!} = 1.955 \times 10^{12}$$

Cuando un espacio muestral consta de resultados igualmente probables, la probabilidad de un evento se puede encontrar al dividir el número de resultados en el evento entre el número total de resultados en el espacio muestral. A veces se pueden usar las reglas de conteo para calcular estos números, como se muestra en el siguiente ejemplo:

Ejemplo

2.15

En una caja de pernos se encuentran ocho gruesos, cinco medianos y tres angostos. Una caja de tuercas contiene seis que ajustan con los pernos gruesos, cuatro que ajustan con los pernos medianos y dos que ajustan con los pernos angostos. Se elige aleatoriamente un perno y una tuerca, ¿cuál es la probabilidad de que la tuerca ajuste con el perno?

Solución

El espacio muestral consta de todos los pares de tuercas y pernos y cada par es igualmente probable de ser elegido. El evento de que la tuerca ajuste con el perno corresponde al conjunto de todos los pares que ajustan de tuercas y pernos. Por tanto,

$$P(\text{tuerca ajusta con perno}) = \frac{\text{número de pares de tuercas y pernos que se ajustan}}{\text{número de pares de tuercas y pernos}}$$

Hay $6 + 4 + 2 = 12$ tuercas y $8 + 5 + 3 = 16$ pernos. Por tanto,

$$\text{Número de pares de tuercas y pernos} = (12)(16) = 192$$

El número de pares que se ajusta se encuentra al sumar el número de pares de tuercas y pernos gruesos, el de pares de tuercas y pernos medianos y el de pares de tuercas y pernos angostos. Estos números son

$$\text{Número de pares de tuercas y pernos gruesos} = (6)(8) = 48$$

$$\text{Número de pares de tuercas y pernos medianos} = (4)(5) = 20$$

$$\text{Número de pares de tuercas y pernos angostos} = (2)(3) = 6$$

Por tanto,

$$P(\text{tuercas que ajustan a los pernos}) = \frac{48 + 20 + 6}{192} = 0.3854$$

Ejercicios para la sección 2.2

- Las moléculas de ADN constan de secuencias químicamente enlazadas a las bases adenina, guanina, citosina y timina, denotadas por A, G, C y T. Una secuencia de tres bases se llama *codón*.
 - ¿Cuántos codones diferentes hay?
 - Las bases A y G son *purinas*, mientras que las C y T son *pirimidínicas*. ¿Cuántos *codones* hay cuya primera y tercera bases son purinas y cuya segunda base es una pirimidínica?
 - ¿Cuántos codones constan de tres bases diferentes?
- Un ingeniero químico está diseñando un experimento para determinar el efecto de temperatura, la razón de activación y el tipo de catalizador en la producción de reacción dada. Quiere estudiar cinco temperaturas diferentes de reacción, dos razones de activación distintas y cuatro catalizadores diferentes. Si cada operación del experimento implica la elección de una temperatura, una razón de activación y un catalizador, ¿cuántas operaciones diferentes son posibles?
- Diez ingenieros han solicitado un puesto administrativo en una gran empresa. Se seleccionará a cuatro de ellos como finalistas para el puesto. ¿De cuántas maneras se puede hacer esta selección?
- Un comité de ocho personas debe elegir un presidente, un vicepresidente y un secretario. ¿De cuántas maneras se puede hacer esta selección?
- Una prueba consta de 15 preguntas. Diez son preguntas verdadero-falso y cinco son de elección múltiple que tienen cuatro opciones cada una. Un estudiante debe seleccionar una respuesta para cada pregunta. ¿De cuántas maneras se puede hacer esta prueba?
- En cierto estado, las placas constan de tres letras seguidas de tres números.
 - ¿Cuántas placas diferentes se pueden hacer?
 - ¿Cuántas placas diferentes se pueden hacer de tal forma que ninguna letra o número aparezca más de una vez?
 - Una placa se elige aleatoriamente. ¿Cuál es la probabilidad de que ninguna letra o número aparezca más de una vez?
- Una contraseña de computadora consta de ocho caracteres.
 - ¿Cuántas contraseñas diferentes son posibles si cada carácter puede ser cualquier letra minúscula o dígito?
 - ¿Cuántas contraseñas diferentes son posibles si cada carácter puede ser cualquier letra minúscula o dígito y al menos un carácter debe ser un dígito?
 - Un sistema de computadora requiere que las contraseñas contengan al menos un dígito. Si se generan ocho caracteres aleatoriamente y cada uno es igualmente probable de ser cualesquiera de las 26 letras o de los diez dígitos, ¿cuál es la probabilidad de que se genere una contraseña válida?
- Una compañía ha contratado a 15 nuevos empleados y debe asignar seis al turno matutino, cinco al vespertino y cuatro al nocturno. ¿De cuántas maneras se puede hacer la asignación?
- Un cajón en un tocador contiene ocho calcetines azules y seis blancos. Un segundo cajón contiene cuatro calcetines azules y dos calcetines blancos. Se elige un calcetín de cada cajón. ¿Cuál es la probabilidad de que combinen?
- Un cajón contiene seis calcetines rojos, cuatro verdes y dos negros. Se elige dos calcetines aleatoriamente. ¿Cuál es la probabilidad de que combinen?

2.3 Probabilidad condicional e independencia

Un espacio muestral contiene todos los resultados posibles de un experimento. A veces se obtiene algo de información adicional acerca de un experimento que indica que los resultados provienen de cierta parte del espacio muestral. En este caso, la probabilidad de un evento está basada en los resultados de esa parte del espacio muestral. Una probabilidad que se basa en una parte de un espacio muestral se llama **probabilidad condicional**. Se analizará esta idea a través de algunos ejemplos.

En el ejemplo 2.6 (en la sección 2.1) se analizó una muestra de mil varillas de aluminio. Para cada varilla, la longitud se clasifica como demasiado corta, demasiado larga o está bien y el diámetro se clasifica como muy delgado, muy grueso o está bien. Esas mil varillas constituyen un espacio muestral en el que cada varilla tiene la misma probabilidad de ser seleccionada. El número de varillas en cada categoría se presenta en la tabla 2.1. De las mil varillas, 928 satisfacen la especificación de diámetro. Por tanto, si se selecciona una varilla, $P(\text{diámetro está bien}) = 928/1\,000 = 0.928$. A esta circunstancia se le llama **probabilidad incondicional**, ya que se basa en todo el espacio muestral. Ahora suponga que se toma una varilla, se mide su longitud y se encuentra que satisface la especificación. ¿Cuál es la probabilidad de que el diámetro también satisfaga la especificación? La clave para calcular esta probabilidad es darse cuenta de que el saber que la longitud satisface la especificación reduce el espacio muestral del que se seleccionó la varilla. La tabla 2.2 presenta esta idea. Una vez que se conoce que se satisface la especificación de la longitud, se sabe que esa varilla será una de las 942 en el espacio muestral que se presenta en la tabla 2.2.

TABLA 2.1 Espacio muestral que contiene mil varillas de aluminio

Longitud	Diámetro		
	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

TABLA 2.2 Espacio muestral reducido que contiene 942 varillas de aluminio que satisfacen la especificación de la longitud

Longitud	Diámetro		
	Muy delgado	Está bien	Muy grueso
Demasiado corta	—	—	—
Está bien	38	900	4
Demasiado larga	—	—	—

De las 942 varillas de este espacio muestral, 900 satisfacen la especificación del diámetro. Por tanto, si se sabe que la varilla satisface la especificación de longitud, la probabilidad

de que la varilla satisfaga la especificación del diámetro es: 900/942. Se dice que la **probabilidad condicional** de que la varilla satisfaga la especificación de un diámetro **dado** que satisfaga la especificación de longitud es igual a 900/942 y se escribe $P(\text{diámetro está bien} \mid \text{longitud está bien}) = 900/942 = 0.955$. Observe que la probabilidad condicional $P(\text{diámetro está bien} \mid \text{longitud está bien})$ es diferente de la probabilidad incondicional $P(\text{diámetro está bien})$, que se calculó para todo el espacio muestral (tabla 2.1) de 0.928.

Ejemplo

2.16

Calcule la probabilidad condicional $P(\text{diámetro está bien} \mid \text{longitud demasiado larga})$. ¿Esta es la misma que la probabilidad incondicional $P(\text{diámetro está bien})$?

Solución

La probabilidad condicional $P(\text{diámetro está bien} \mid \text{longitud demasiado larga})$ se calcula bajo la suposición de que la varilla es demasiado larga. Esto último reduce el espacio muestral a los 40 elementos que se muestran en **negritas** en la tabla siguiente.

Longitud	Diámetro		
	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

De los 40 resultados, 25 satisfacen la especificación de diámetro. Por tanto

$$P(\text{diámetro está bien} \mid \text{longitud demasiado larga}) = \frac{25}{40} = 0.625$$

La probabilidad incondicional $P(\text{diámetro está bien})$ se calcula con base en todos los mil resultados en el espacio muestral y es igual a $928/1\,000 = 0.928$. En este caso, la probabilidad condicional es diferente de la probabilidad incondicional.

Se analizará la solución del ejemplo 2.16 detenidamente. Se encuentra que

$$P(\text{diámetro está bien} \mid \text{longitud demasiado larga}) = \frac{25}{40}$$

En la respuesta 25/40, el denominador, 40, representa el número de resultados que satisfacen la condición de que la varilla es demasiado larga, mientras que el numerador, 25, representa el número de resultados que satisfacen ambas condiciones, que la varilla es demasiado larga y que su diámetro está bien. Si dividimos tanto al numerador como al denominador de esta respuesta entre el número de resultados en todo el espacio muestral, que es de mil, se obtiene

$$P(\text{diámetro está bien} \mid \text{longitud demasiado larga}) = \frac{25/1\,000}{40/1\,000}$$

Ahora $40/1\,000$ representa la *probabilidad* de que se satisface la condición de que la varilla es demasiado larga. Esto es,

$$P(\text{longitud demasiado larga}) = \frac{40}{1\,000}$$

La cantidad $25/1\,000$ representa la *probabilidad* de que se satisfacen tanto la condición de que la varilla es demasiado larga y de que el diámetro está bien. Esto es,

$$P(\text{diámetro está bien y longitud demasiado larga}) = \frac{25}{1\,000}$$

Ahora se puede expresar la probabilidad condicional como

$$P(\text{diámetro está bien} \mid \text{longitud demasiado larga}) = \frac{P(\text{diámetro está bien y longitud demasiado larga})}{P(\text{longitud demasiado larga})}$$

Este razonamiento se puede ampliar para construir una definición de la probabilidad condicional que es válida para cualquier espacio muestral.

Definición

Sean A y B eventos con $P(B) \neq 0$. La probabilidad condicional de A dado B es

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.14)$$

La figura 2.4 presenta diagramas de Venn para ilustrar la idea de la probabilidad condicional.

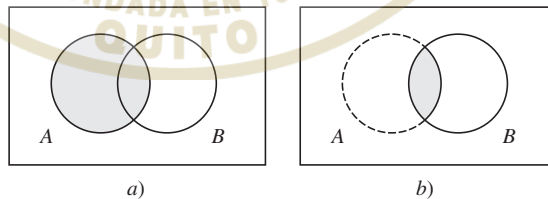


FIGURA 2.4 a) El diagrama representa la probabilidad incondicional $P(A)$. $P(A)$ se muestra al considerar el evento A en proporción con todo el espacio muestral, el cual se representa por el rectángulo. b) El diagrama representa la probabilidad condicional $P(A|B)$. Puesto que se sabe que ocurre el evento B , ahora éste será el espacio muestral. Para que el evento A ocurra, el resultado debe estar en la intersección $A \cap B$. Por tanto, la probabilidad condicional $P(A|B)$ se muestra al considerar la intersección $A \cap B$ en proporción con todo el evento B .

Ejemplo 2.17

En referencia al ejemplo 2.8 de la sección 2.1, ¿cuál es la probabilidad de que una lata tenga una fisura en el costado, dado que tiene una fisura en la tapa?

Solución

Se tiene que $P(\text{fisura en la tapa}) = 0.03$ y que $P(\text{fisura en el costado y fisura en la tapa}) = 0.01$, utilizando la ecuación (2.14),

$$\begin{aligned} P(\text{fisura en el costado} \mid \text{fisura en la tapa}) &= \frac{P(\text{fisura en el costado y fisura en la tapa})}{P(\text{fisura en la tapa})} \\ &= \frac{0.01}{0.03} \\ &= 0.33 \end{aligned}$$

Ejemplo 2.18

Con referencia al ejemplo 2.8 (en la sección 2.1), ¿cuál es la probabilidad de que una lata tenga una fisura en la tapa, dado que tiene una fisura en el costado?

Solución

Se tiene que $P(\text{fisura en el costado}) = 0.02$ y que $P(\text{fisura en el costado y fisura en la tapa}) = 0.01$. Utilizando la ecuación (2.14),

$$\begin{aligned} P(\text{fisura en la tapa} \mid \text{fisura en el costado}) &= \frac{P(\text{fisura en la tapa y fisura en el costado})}{P(\text{fisura en el costado})} \\ &= \frac{0.01}{0.02} \\ &= 0.5 \end{aligned}$$

Los resultados de los ejemplos 2.17 y 2.18 muestran que en general, $P(A \mid B) \neq P(B \mid A)$.

Eventos independientes

Algunas veces el conocimiento de que un evento ha ocurrido no cambia la probabilidad de que ocurra otro. En este caso las probabilidades condicional e incondicional son las mismas y se dice que los eventos son **independientes**. A continuación se presenta un ejemplo.

Ejemplo 2.19

Si una varilla de aluminio se selecciona del espacio muestral que se presenta en la tabla 2.1, determine $P(\text{demasiado larga})$ y $P(\text{demasiado larga} \mid \text{muy angosta})$. ¿Estas probabilidades son diferentes?

Solución

$$P(\text{demasiado larga}) = \frac{40}{1\,000} = 0.040$$

$$\begin{aligned} P(\text{demasiado larga} \mid \text{muy angosta}) &= \frac{P(\text{demasiado larga y muy angosta})}{P(\text{muy angosta})} \\ &= \frac{2/1\,000}{50/1\,000} \\ &= 0.040 \end{aligned}$$

La probabilidad condicional y la probabilidad incondicional son las mismas. La información de que la varilla es muy angosta no cambia la probabilidad de que la varilla es demasiado larga.

El ejemplo 2.19 muestra que el conocimiento de que un evento ocurre a veces no cambia la probabilidad de que ocurra otro evento. En estos casos, se dice que los dos eventos son **independientes**. El evento de que una varilla es demasiada larga y el evento de que una varilla es muy angosta son independientes. A continuación se presenta una definición más precisa del término, tanto en palabras como en símbolos.

Definición

Dos eventos A y B son **independientes** si la probabilidad de cada uno es la misma si ocurren o no los demás eventos.

En símbolos: si $P(A) \neq 0$ y $P(B) \neq 0$, entonces A y B son independientes si

$$P(B \mid A) = P(B) \quad \text{o, de manera equivalente,} \quad P(A \mid B) = P(A) \quad (2.15)$$

Si $P(A) = 0$ o $P(B) = 0$, entonces A y B son independientes.

Si A y B son independientes, entonces los siguientes pares de eventos son también independientes: A y B^c , A^c y B y A^c y B^c . La prueba de este hecho se deja como ejercicio.

El concepto de independencia se puede ampliar para más de dos eventos:

Definición

Los eventos $A_1, A_2, \dots, A_n$ son independientes si la probabilidad de cada uno permanece igual no importando lo que ocurre con los otros.

En símbolos: los eventos $A_1, A_2, \dots, A_n$ son independientes si para cada A_i y cada colección $A_{j1}, A_{j2}, \dots, A_{jm}$ de eventos con $P(A_{j1} \cap \dots \cap A_{jm}) \neq 0$,

$$P(A_i \mid A_{j1} \cap \dots \cap A_{jm}) = P(A_i) \quad (2.16)$$

La regla de la multiplicación

Algunas veces se conoce $P(A|B)$ y se desea encontrar $P(A \cap B)$. Se puede obtener un resultado que sea útil para este propósito al multiplicar ambos lados de la ecuación (2.14) por $P(B)$. Esto conduce a la regla de la multiplicación.

Si A y B son dos eventos con $P(B) \neq 0$, entonces

$$P(A \cap B) = P(B) P(A | B) \quad (2.17)$$

Si A y B son dos eventos con $P(A) \neq 0$, entonces

$$P(A \cap B) = P(A) P(B | A) \quad (2.18)$$

Si $P(A) \neq 0$ y $P(B) \neq 0$, entonces tanto la ecuación (2.17) como la (2.18) son válidas.

Cuando dos eventos son independientes, entonces $P(A|B) = P(A)$ y $P(A|B) = P(B)$, así la regla de la multiplicación se simplifica:

Si A y B son eventos independientes, entonces

$$P(A \cap B) = P(A) P(B) \quad (2.19)$$

Este resultado se puede ampliar para cualquier número de eventos. Si $A_1, A_2, \dots, A_n$ son eventos independientes, entonces para cada colección $A_{j1}, A_{j2}, \dots, A_{jm}$ de eventos

$$P(A_{j1} \cap A_{j2} \cap \dots \cap A_{jm}) = P(A_{j1}) P(A_{j2}) \dots P(A_{jm}) \quad (2.20)$$

En particular,

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2) \dots P(A_n) \quad (2.21)$$

Ejemplo 2.20

Un vehículo tiene dos motores: uno principal y otro auxiliar. El componente del motor falla sólo si fallan ambos motores. La probabilidad de que el motor principal falle es de 0.05 y la de que el motor auxiliar falle es de 0.10. Suponga que los motores principal y auxiliar funcionan de manera independiente. ¿Cuál es la probabilidad de que el componente del motor falle?

Solución

La probabilidad de que el componente del motor falle es la probabilidad de que ambos motores fallen. Por tanto,

$$P(\text{componente del motor falla}) = P(\text{motor principal falla y motor auxiliar falla})$$

Puesto que los motores son independientes, se puede usar la ecuación (2.19):

$$\begin{aligned}
 P(\text{motor principal falla y motor auxiliar falla}) &= P(\text{motor principal falla})P(\text{motor auxiliar falla}) \\
 &= (0.10)(0.05) \\
 &= 0.005
 \end{aligned}$$

Ejemplo

2.21

Un sistema contiene dos componentes, A y B . Ambos componentes deben funcionar para que el sistema trabaje. La probabilidad de que el componente A falle es de 0.08 y de que el B lo haga es de 0.05. Suponga que los dos componentes funcionan de manera independiente, ¿cuál es la probabilidad de que el sistema funcione?

Solución

La probabilidad de que el sistema funcione es la probabilidad de que ambos componentes funcionen. Por tanto,

$$P(\text{funciona el sistema}) = P(\text{funciona } A \text{ y funciona } B)$$

Puesto que los componentes funcionan de manera independiente,

$$\begin{aligned}
 P(\text{funciona } A \text{ y funciona } B) &= P(A \text{ funciona})P(B \text{ funciona}) \\
 &= [1 - P(\text{falla } A)][1 - P(\text{falla } B)] \\
 &= (1 - 0.08)(1 - 0.05) \\
 &= 0.874
 \end{aligned}$$

Ejemplo

2.22

De los microprocesadores fabricados mediante cierto proceso, 20% está defectuoso. Se elige aleatoriamente cinco de aquéllos. Suponga que funcionan independientemente, ¿cuál es la probabilidad de que todos funcionen?

Solución

Para $i = 1, \dots, 5$, sea que A_i denote el evento de que el i -ésimo microprocesador funcione. Entonces

$$\begin{aligned}
 P(\text{los 5 funcionan}) &= P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\
 &= P(A_1)P(A_2)P(A_3)P(A_4)P(A_5) \\
 &= (1 - 0.20)^5 \\
 &= 0.328
 \end{aligned}$$

Ejemplo

2.23

En el ejemplo 2.22, ¿cuál es la probabilidad de que al menos uno de los microprocesadores funcione?

Solución

La manera más fácil de resolver este problema es observar que

$$P(\text{al menos uno funciona}) = 1 - P(\text{todos están defectuosos})$$

Ahora, denotando a D_i como el evento de que el i -ésimo microprocesador es defectuoso,

$$\begin{aligned}
 P(\text{todos están defectuosos}) &= P(D_1 \cap D_2 \cap D_3 \cap D_4 \cap D_5) \\
 &= P(D_1)P(D_2)P(D_3)P(D_4)P(D_5) \\
 &= (0.20)^5 \\
 &= 0.0003
 \end{aligned}$$

Por tanto, $P(\text{al menos uno funciona}) = 1 - 0.0003 = 0.9997$.

Las ecuaciones (2.19) y (2.20) indican cómo calcular las probabilidades cuando se sabe que los eventos son independientes, pero no son generalmente de mucha ayuda para determinar si dos eventos *son realmente independientes*. En la mayoría de los casos, la mejor manera de determinar si los eventos son independientes es comprendiendo el proceso que los produce. He aquí algunos ejemplos:

- Se tira dos veces un dado. Es razonable creer que el resultado de la segunda tirada no se vea afectado por el resultado de la primera. Por tanto, conocer el resultado de la primera tirada no ayuda a predecir el resultado de la segunda. Las dos tiradas son independientes.
- Cierta reacción química se realiza dos veces, utilizando equipos diferentes cada vez. Es razonable creer que el producto de una reacción no afectará la producción de la otra. En este caso los productos son independientes.
- Una reacción química se realiza dos veces sucesivamente, utilizando el mismo equipo. En este caso, no es sensato suponer que las producciones son independientes. Por ejemplo, una producción baja en la primera realización podría indicar que hay más residuos de lo normal. Lo anterior podría tender a hacer que la producción en la siguiente realización fuese más alta. Por tanto, conocer la primera producción puede ayudar a predecir la producción en la segunda realización.
- Los elementos en una muestra aleatoria simple se pueden tratar como independientes, a menos que la población sea finita y la muestra consista de más de 5% de la población (véase el análisis de independencia en la sección 1.1).

Ley de la probabilidad total

La ley de la probabilidad total se muestra en la figura 2.5. Un espacio muestral contiene los eventos A_1, A_2, A_3 y A_4 . Éstos son mutuamente excluyentes, ya que dos no coinciden. Son también **exhaustivos**, ello significa que su unión abarca todo el espacio muestral. Cada resultado en este espacio pertenece a uno y sólo uno de los eventos A_1, A_2, A_3, A_4 .

El evento B es cualquier evento. En la figura 2.5, cada uno de los eventos A_i que intersectan a B , forman los eventos $A_1 \cap B, A_2 \cap B, A_3 \cap B$ y $A_4 \cap B$. Es obvio de la figura 2.5 que los eventos $A_1 \cap B, A_2 \cap B, A_3 \cap B$ y $A_4 \cap B$ son mutuamente excluyentes y que abarcan a B . Cada resultado en B pertenece a uno y sólo uno de los eventos $A_1 \cap B, A_2 \cap B, A_3 \cap B$ y $A_4 \cap B$. Por lo que

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup (A_4 \cap B)$$

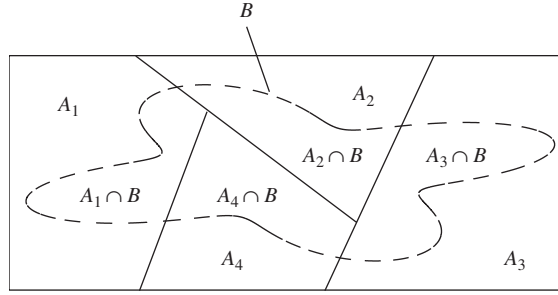


FIGURA 2.5 Los eventos A_1, A_2, A_3, A_4 mutuamente excluyentes y exhaustivos dividen al evento B en subconjuntos mutuamente excluyentes.

el cual es la unión de los eventos mutuamente excluyentes. Por tanto,

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + P(A_4 \cap B)$$

Debido a que $P(A_i \cap B) = P(B|A_i)P(A_i)$,

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) + P(B|A_4)P(A_4) \quad (2.22)$$

La ecuación (2.22) es un caso especial de la ley de la probabilidad total, restringida al caso donde hay cuatro eventos mutuamente excluyentes y exhaustivos. La intuición detrás de la ley de la probabilidad total es muy simple. Los eventos A_1, A_2, A_3, A_4 parten al evento en B piezas. La probabilidad de B se encuentra sumando las probabilidades de todas las piezas.

Se podría dibujar nuevamente la figura 2.5 para tener cualquier número de eventos A_i . Esto conduce al caso general de la ley de la probabilidad total.

Ley de la probabilidad total

Si $A_1, \dots, A_n$ son eventos mutuamente excluyentes y exhaustivos y B es cualquier evento, entonces

$$P(B) = P(A_1 \cap B) + \dots + P(A_n \cap B) \quad (2.23)$$

De manera equivalente, si $P(A_i) \neq 0$ para cada A_i ,

$$P(B) = P(B|A_1)P(A_1) + \dots + P(B|A_n)P(A_n) \quad (2.24)$$

Ejemplo 2.24

Clientes que compran cierta marca de automóvil pueden pedir un motor en cualquiera de tres tamaños. De todos los automóviles vendidos, 45% tiene el motor más pequeño, 35% tamaño mediano y 20% más grande. Los automóviles en una prueba de emisiones dentro de los dos años de su compra fallan 10% con el motor más pequeño, mientras que 12% de los de tamaño mediano y 15% de los de motor más grande. ¿Cuál es la probabilidad de que un automóvil elegido aleatoriamente pueda fallar en una prueba de emisiones dentro de los dos primeros años?

Solución

Sea B el evento de que un automóvil falle en una prueba de emisiones dentro de los dos primeros años. Sea A_1 el evento que es un automóvil con un motor pequeño, A_2 el evento que un automóvil tiene un motor mediano y A_3 el evento que un automóvil tiene un motor grande. Entonces

$$P(A_1) = 0.45 \quad P(A_2) = 0.35 \quad P(A_3) = 0.20$$

La probabilidad de que un automóvil falle una prueba, dado que tiene un motor pequeño, es de 0.10. Es decir, $P(B|A_1) = 0.10$. De manera similar, $P(B|A_2) = 0.12$ y $P(B|A_3) = 0.15$. Por la ley de probabilidad total (ecuación 2.24).

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) \\ &= (0.10)(0.45) + (0.12)(0.35) + (0.15)(0.20) \\ &= 0.117 \end{aligned}$$

A veces problemas como los del ejemplo 2.24 se resuelven con el uso de diagramas de árbol. La figura 2.6 presenta un diagrama de árbol para el ejemplo 2.24. Hay tres ramas principales en el árbol, que corresponde a los tres tamaños de motor. Las probabilidades de los tamaños de motor se colocan en sus ramas respectivas. Al final de cada rama principal están dos ramas secundarias, que representan los eventos de falla y no falla. Las probabilidades condicionales de falla y no falla, dado el tamaño del motor, se colocan en las ramas secundarias. Al multiplicar a lo largo de cada una de las ramas que corresponden al evento $B =$ falla, se obtienen las probabilidades $P(B|A_i)P(A_i)$. Al sumar estas probabilidades se obtiene $P(B)$, como se quería.

Regla de Bayes

Si A y B son dos eventos, se ha visto que en general $P(A|B) \neq P(B|A)$. La regla de Bayes proporciona una fórmula que permite calcular una de las probabilidades condicionales si se conoce la otra. Para ver cómo funciona, suponga que se conoce $P(B|A)$ y que se desea calcular $P(A|B)$. Se inicia con la definición de probabilidad condicional (ecuación 2.14):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Ahora se utiliza la ecuación (2.18) en la que se sustituye $P(B|A)P(A)$ por $P(A \cap B)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.25)$$

La ecuación (2.25) es esencialmente la regla de Bayes. Cuando se escribe la regla de Bayes, la expresión $P(B)$ del denominador se reemplaza con una expresión más complicada obtenida por lo general de la ley de la probabilidad total. De modo específico, puesto que los eventos A y A^c son mutuamente excluyentes y exhaustivos, la ley de la probabilidad total muestra que

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) \quad (2.26)$$

Al sustituir en el lado derecho de la ecuación (2.26) a $P(B)$ de la ecuación (2.25) se obtiene la regla de Bayes. También se puede obtener una versión más general de la regla de Bayes al

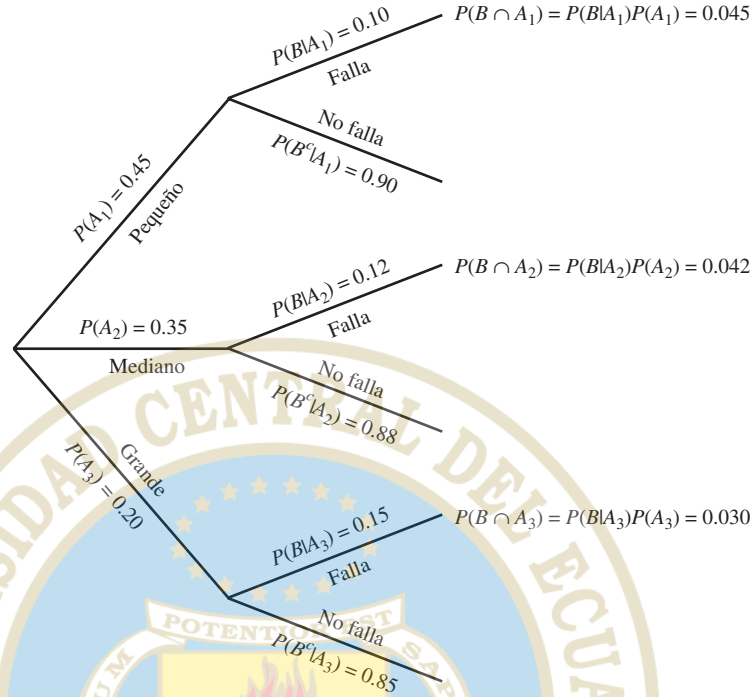


FIGURA 2.6 Diagrama de árbol para la solución del ejemplo 2.24.

considerar una colección $A_1, \dots, A_n$ de eventos mutuamente excluyentes y exhaustivos y utilizando la ley de la probabilidad total sustituyendo a $P(B)$ con la expresión del lado derecho de la ecuación (2.24).

La regla de Bayes

Caso especial: Sean A y B eventos con $P(A) \neq 0$, $P(A^c) \neq 0$ y $P(B) \neq 0$. Entonces

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \quad (2.27)$$

Caso general: Sean $A_1, \dots, A_n$ eventos mutuamente excluyentes y exhaustivos con $P(A_i) \neq 0$ para cada A_i . Sea B cualquier evento con $P(B) \neq 0$. Entonces

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_i)P(A_i)} \quad (2.28)$$

En el ejemplo 2.25 se muestra cómo se puede usar la regla de Bayes para descubrir un importante y sorprendente resultado en el campo de las pruebas médicas.

Ejemplo 2.25

La proporción de personas en una comunidad que tienen cierta enfermedad es 0.005. Está disponible una prueba para diagnosticar la enfermedad. Si una persona la padece, la probabilidad de que la prueba dé una señal positiva es 0.99. Si una persona no está enferma, la probabilidad de que la prueba dé una señal positiva es 0.01. Si una persona sale positiva en la prueba, ¿cuál es la probabilidad de que la persona realmente esté enferma?

Solución

Sea D el evento de que la persona realmente está enferma y sea $+$ el evento que la prueba sale positiva. Se desea encontrar $P(D|+)$. Se dan las probabilidades siguientes:

$$P(D) = 0.005 \quad P(+|D) = 0.99 \quad P(+|D^c) = 0.01$$

Al utilizar la regla de Bayes (ecuación 2.27),

$$\begin{aligned} P(D|+) &= \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)} \\ &= \frac{(0.99)(0.005)}{(0.99)(0.005) + (0.01)(0.995)} \\ &= 0.332 \end{aligned}$$

En el ejemplo 2.25 sólo aproximadamente un tercio de las personas que resulta positivo en el examen, realmente padece la enfermedad. Observe que la prueba es bastante exacta; clasifica correctamente a 99% de las personas tanto enfermas como no enfermas. La razón por la que una proporción grande de pruebas resulta positiva cuando en realidad se está libre de la enfermedad es que la enfermedad es rara: sólo 0.5% de la población la padece. Debido a que cierta cantidad de enfermedades son raras, muchas pruebas médicas resultarán positivas cuando en realidad son falsos positivos, aun cuando la prueba es bastante exacta. Por esta razón, cuando una prueba sale positiva se hace generalmente una segunda prueba antes de que se confirme el diagnóstico.

Ejemplo 2.26

Con referencia al ejemplo 2.24, se elige aleatoriamente un registro de una prueba de emisiones con falla. ¿Cuál es la probabilidad de que éste sea un automóvil con un motor pequeño?

Solución

Sea B el evento que un automóvil falla en una prueba de emisiones. Sea A_1 el evento que un automóvil tiene un motor pequeño, A_2 el evento de que lo tiene mediano y A_3 que su motor es grande. Se desea encontrar $P(A_1|B)$. Las siguientes probabilidades están dadas en el ejemplo 2.24:

$$\begin{aligned} P(A_1) &= 0.45 & P(A_2) &= 0.35 & P(A_3) &= 0.20 \\ P(B|A_1) &= 0.10 & P(B|A_2) &= 0.12 & P(B|A_3) &= 0.15 \end{aligned}$$

Por la regla de Bayes,

$$\begin{aligned}
 P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)} \\
 &= \frac{(0.10)(0.45)}{(0.10)(0.45) + (0.12)(0.35) + (0.15)(0.20)} \\
 &= 0.385
 \end{aligned}$$

Aplicación al análisis de confiabilidad

El análisis de confiabilidad constituye la rama de la ingeniería que se dedica al cálculo de las tasas de fallas de los sistemas. Mientras que algunos problemas en el análisis de confiabilidad requieren de métodos matemáticos avanzados, hay muchos problemas que se pueden resolver con los métodos que se han aprendido hasta ahora. Se inicia con un ejemplo que muestra el cálculo de confiabilidad de un sistema que consiste de dos componentes conectados *en serie*.

Ejemplo 2.27

Un sistema contiene dos componentes, *A* y *B*, conectados en serie como se muestra en el diagrama siguiente.



El sistema funcionará sólo si ambos componentes funcionan. La probabilidad de que *A* funcione está dada por $P(A) = 0.98$ y la probabilidad de que *B* funciona está dada por $P(B) = 0.95$. Suponga que *A* y *B* funcionan de manera independiente. Determine la probabilidad de que el sistema funcione.

Solución

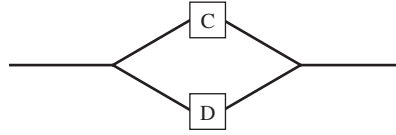
Debido a que el sistema funcionará sólo si ambos componentes funcionan, se tiene que

$$\begin{aligned}
 P(\text{sistema funciona}) &= P(A \cap B) \\
 &= P(A)P(B) \text{ por la suposición de independencia} \\
 &= (0.98)(0.95) \\
 &= 0.931
 \end{aligned}$$

El ejemplo 2.28 muestra el cálculo de confiabilidad de un sistema que consta de dos componentes conectados *en paralelo*.

Ejemplo 2.28

Un sistema contiene dos componentes, *C* y *D*, conectados en paralelo como se muestra en el diagrama siguiente.



El sistema funcionará si alguno, C o D funcionan. La probabilidad de que C funcione es 0.90 y la de que D lo haga es 0.85. Suponga que C y D funcionan de manera independiente. Determine la probabilidad de que el sistema funcione.

Solución

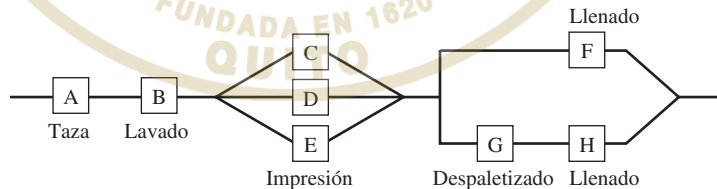
Debido a que el sistema funcionará mientras que cualesquiera de los dos componentes funcione, se tiene que

$$\begin{aligned}
 P(\text{sistema funciona}) &= P(C \cup D) \\
 &= P(C) + P(D) - P(C \cap D) \\
 &= P(C) + P(D) - P(C)P(D) \\
 &\quad \text{por la suposición de independencia} \\
 &= 0.90 + 0.85 - (0.90)(0.85) \\
 &= 0.985
 \end{aligned}$$

Con frecuencia la confiabilidad de sistemas más complejos se puede determinar al descomponer al sistema en una serie de subsistemas, cada uno contiene componentes conectados en serie o en paralelo. El ejemplo 2.29 muestra el método.

Ejemplo 2.29

La tesis “Dynamic, Single-stage, Multiperiod, Capacitated Production Sequencing Problem with Multiple Parallel Resources” (D. Ott, tesis de maestría, Escuela de Minas, de Colorado, 1998) describe un método de producción utilizado en la fabricación de latas de aluminio. El siguiente diagrama esquemático, ligeramente simplificado, muestra el proceso.



La entrada inicial al proceso consta de hojas de aluminio enrolladas, aproximadamente 0.25 mm de espesor. En un proceso que se conoce como “hacer formas de taza”, se desenrollan estas hojas y se les da forma, en cuerpos de lata, que son cilindros cerrados en la parte inferior y abiertos en la parte superior. Después, estos cuerpos de lata se lavan y se envían a la impresión, donde se imprime la etiqueta sobre la lata. En la práctica hay varias impresoras en una línea; el diagrama presenta una línea con tres de aquéllas. El impresor deposita las latas en paletas, que son estructuras de madera que sostienen 7 140 latas cada una. El siguiente paso es llenar las latas. Algunas líneas de llenado pueden aceptar latas directamente de las paletas,

pero otras sólo las aceptan de los recipientes de las latas, que son grandes contenedores que tienen aproximadamente cien mil latas cada uno. Para usar estas líneas de llenado, las latas se deben transportar de las paletas a los recipientes de las latas, en un proceso que se llama despaletizado. En la práctica hay varias líneas de llenado; el diagrama presenta un caso donde hay dos líneas de llenado, una que aceptará latas de las paletas y la otra aceptará latas que no provienen de las paletas. En el proceso de llenado de latas, éstas se llenan y se les coloca la tapa. Después se empaquetan las latas y se envían a los distribuidores.

Se desea calcular la probabilidad de que el proceso funcione un día sin falla. Suponga que el proceso de hacer la lata tiene una probabilidad de 0.995 de funcionar con éxito en un día. Puesto que este componente se denota por “A” en el diagrama, se expresa esta probabilidad como $P(A) = 0.995$. Suponga que las otras componentes del proceso tienen las siguientes probabilidades de funcionar con éxito durante un periodo de un día: $P(B) = 0.99$, $P(C) = P(D) = P(E) = 0.95$, $P(F) = 0.90$, $P(G) = 0.90$, $P(H) = 0.98$. Suponga que los componentes funcionan de manera independiente. Determine la probabilidad de que el proceso funcione con éxito en un día.

Solución

Es posible resolver este problema observando que todo el proceso se puede descomponer en subsistemas, cada uno de éstos consta de sistemas compuestos sólo en serie o en paralelo. Específicamente, el subsistema 1 consiste de los componentes hacer la lata y lavar, que están conectadas en serie; el subsistema 2 consiste de las impresoras que están conectadas en paralelo; el subsistema 3 consiste de líneas de llenado, que están conectadas en paralelo, con una de las dos líneas que consta de dos componentes conectados en serie.

Se calculan las probabilidades del funcionamiento con éxito para cada subsistema, denotando las probabilidades con p_1 , p_2 y p_3 .

$$\begin{aligned} P(\text{subsistema 1 funciona}) &= p_1 = P(A \cap B) \\ &= P(A)P(B) \\ &= (0.995)(0.990) \\ &= 0.985050 \end{aligned}$$

$$\begin{aligned} P(\text{subsistema 2 funciona}) &= p_2 = 1 - P(\text{subsistema 2 falla}) \\ &= 1 - P(C^c \cap D^c \cap E^c) \\ &= 1 - P(C^c)P(D^c)P(E^c) \\ &= 1 - (0.05)^3 \\ &= 0.999875 \end{aligned}$$

El subsistema 3 funciona si F funciona, o si tanto G como H funcionan. Por tanto,

$$\begin{aligned} P(\text{subsistema 3 funciona}) &= p_3 = P(F \cup (G \cap H)) \\ &= P(F) + P(G \cap H) - P(F \cap G \cap H) \\ &= P(F) + P(G)P(H) - P(F)P(G)P(H) \\ &= (0.90) + (0.90)(0.98) - (0.90)(0.90)(0.98) \\ &= 0.988200 \end{aligned}$$

Todo el proceso consta de tres subsistemas conectados en serie. Por tanto, para que el proceso funcione, los tres subsistemas deben funcionar. Se concluye que

$$\begin{aligned}
 P(\text{funciona el sistema}) &= P(\text{los sistemas 1, 2 y 3 todos funcionan}) \\
 &= p_1 p_2 p_3 \\
 &= (0.985050)(0.999875)(0.988200) \\
 &= 0.973
 \end{aligned}$$

Se advierte que la suposición de que los componentes funcionan independientemente es crucial en las soluciones de los ejemplos 2.27, 2.28 y 2.29. Cuando no se satisface esta suposición, puede ser muy difícil hacer que los cálculos tengan una precisión confiable. Si se usa la suposición de independencia sin justificación, los cálculos pueden tener una confiabilidad engañosa.

Ejercicios para la sección 2.3

- Una caja contiene diez fusibles. Ocho de ellos están tasados en 10 amperes (A) y los otros dos están tasados en 15 A. Se seleccionan dos fusibles aleatoriamente.
 - ¿Cuál es la probabilidad de que el primer fusible esté tasado en 15 A?
 - ¿Cuál es la probabilidad de que el segundo fusible esté tasado en 15 A, dado que el primer fusible esté tasado en 10 A?
 - ¿Cuál es la probabilidad de que el segundo fusible esté tasado en 15 A, dado que el primer fusible lo esté en 15 A?
- Con referencia al ejercicio 1, se seleccionan aleatoriamente fusibles de la caja, uno tras otro, hasta que se selecciona uno de 15 A.
 - ¿Cuál es la probabilidad de que los primeros dos fusibles sean ambos de 10 A?
 - ¿Cuál es la probabilidad de que un total de dos fusibles sean elegidos de la caja?
 - ¿Cuál es la probabilidad de que más de tres fusibles sean elegidos de la caja?
- Un día de graduación de una gran universidad, se selecciona aleatoriamente a un graduado. Sea A el evento que el estudiante está por terminar la carrera de ingeniería y sea B el evento que el estudiante tomó un curso de cálculo en la universidad. ¿Qué probabilidad es mayor, $P(A|B)$ o $P(B|A)$? Explique.
- En el artículo “Integrating Risk Assessment and Life Cycle Assessment: A Case Study of Insulation” (Y. Nishioka, J. Levy y colaboradores, *Risk Analysis*, 2002:1003-1017) se calcula que 5.6% de cierta población tiene asma y que un asmático tiene 0.027 probabilidad de sufrir un ataque en un día dado. Se elige aleatoriamente a una persona de esta población. ¿Cuál es la probabilidad de que esta persona tenga un ataque asmático en ese día?
- Los pozos de petróleo perforados en la región A tienen una probabilidad de 0.2 de producir. Los pozos perforados en la región B tienen una probabilidad de 0.09. Se perfora un pozo en cada región. Suponga que los pozos producen de manera independiente.
 - ¿Cuál es la probabilidad de que ambos pozos produzcan?
 - ¿Cuál es la probabilidad de que ninguno produzca?
 - ¿Cuál es la probabilidad de que al menos uno produzca?
- De todas las fallas de un tipo determinado de unidad de disco duro de computadora, se determina que 20% de éstos tiene dañado sólo el sector que contiene la tabla de asignación de archivos, en 70% sólo los sectores no esenciales están dañados y en 10% tanto el sector de asignación como uno o más sectores no esenciales están dañados. Se selecciona aleatoriamente una unidad de disco dañada y se examina.
 - ¿Cuál es la probabilidad de que el sector de asignación esté dañado?
 - ¿Cuál es la probabilidad de que un sector no esencial esté dañado?
 - Si se encuentra que la unidad de disco tiene un sector de asignación dañado, ¿cuál es la probabilidad de que algunos sectores no esenciales también estén dañados?

- d) Si se encuentra que la unidad de disco tiene un sector no esencial dañado, ¿cuál es la probabilidad de que el sector de asignación también esté dañado?
- e) Si se encuentra que la unidad de disco tiene un sector de asignación dañado, ¿cuál es la probabilidad de que ningún sector no esencial esté dañado?
- f) Si se encuentra que la unidad de disco tiene un sector no esencial dañado, ¿cuál es la probabilidad de que el sector de asignación no está dañado?
7. En el proceso de producción de válvulas de motor, éstas se someten a un primer rectificado. Las válvulas cuyos espesores están dentro de la especificación se encuentran listas para la instalación. Las válvulas cuyos espesores están arriba de la especificación se rectifican, mientras que aquellas cuyos espesores están por debajo se desechan. Suponga que después del primer rectificado, 70% de las válvulas satisface la especificación, 20% es nuevamente rectificado y 10% se desecha. Además, suponga que de las válvulas que son nuevamente rectificadas, 90% satisface la especificación y 10% se desecha.
- a) Determine la probabilidad de que una válvula se rectifique sólo una vez.
- b) Dado que una válvula se hace sólo una vez, ¿cuál es la probabilidad de que se deseché?
- c) Determine la probabilidad de que se deseché una válvula.
- d) Dado que una válvula se desecha, ¿cuál es la probabilidad de que se rectifique dos veces?
- e) Determine la probabilidad de que la válvula satisfaga la especificación (después de la primera o de la segunda rectificación).
- f) Dado que una válvula satisface la especificación (después de la primera o segunda rectificación), ¿cuál es la probabilidad de que se haya rectificado dos veces?
- g) Dado que una válvula satisface la especificación, ¿cuál es la probabilidad de que se haya rectificado una vez?
8. Sarah y Thomas tiran un dado cada uno. Quien obtenga el número más alto gana; si ambos tiran el mismo número de veces, ninguno gana.
- a) ¿Cuál es la probabilidad de que Thomas gane?
- b) Si Sarah tira un 3, ¿cuál es la probabilidad de que gane?
- c) Si Sarah tira un 3, ¿qué probabilidad hay de que Thomas gane?
- d) Si Sarah gana, ¿cuál es la probabilidad de que Thomas haya tirado un 3?
- e) Si Sarah gana, ¿cuál es la probabilidad de que Sarah haya tirado un 3?
9. Un sistema de aspersión automático especial tiene dos tipos diferentes de dispositivos de activación para cada regadera. Un tipo tiene una confiabilidad de 0.9; es decir, la probabilidad de que se active cuando debe el aspersor es 0.9. El otro tipo, que opera independientemente del primer tipo, tiene una confiabilidad de 0.8. Si se dispara cualquier dispositivo, el aspersor se activará. Suponga que empieza un fuego cerca de una regadera.
- a) ¿Cuál es la probabilidad de que la regadera se active?
- b) ¿Cuál es la probabilidad de que la regadera no se active?
- c) ¿Cuál es la probabilidad de que ambos dispositivos de activación trabajen adecuadamente?
- d) ¿Cuál es la probabilidad de que sólo el dispositivo con 0.9 de confiabilidad trabaje adecuadamente?
10. Una cadena de restaurantes de comida rápida tiene 600 negocios en los Estados Unidos. En la tabla siguiente se clasifican las ciudades por el tamaño y la ubicación y presenta el número de restaurantes en ciudades para cada categoría. Se elige aleatoriamente un restaurante de los 600 para hacer una prueba de mercado de un nuevo estilo de pollo.

	Región			
Población de la ciudad	NE	SE	SO	NO
Debajo de 50 000	30	35	15	5
50 000-500 000	60	90	70	30
Arriba de 500 000	150	25	30	60

- a) Si el restaurante está ubicado en una ciudad con una población de arriba de 500 000, ¿cuál es la probabilidad de que esté en el noreste?
- b) Si el restaurante está ubicado en el sureste, ¿cuál es la probabilidad de que esté en una ciudad que tenga una población debajo de 50 000?
- c) Si el restaurante está ubicado en el suroeste, ¿cuál es la probabilidad de que esté en una ciudad que tenga una población de 500 000 o menos?
- d) Si el restaurante está ubicado en una ciudad con una población de 500 000 o menos, ¿cuál es la probabilidad de que esté en el suroeste?
- e) Si el restaurante está ubicado en el sur (ya sea en el SE o en el SO), ¿cuál es la probabilidad de que esté en una ciudad con una población de 50 000 o más?
11. Las centrales nucleares tienen componentes superfluos en sistemas importantes para reducir la probabilidad de fallas

catastróficas. Suponga que una planta tiene dos sistemas de medición del nivel de refrigerante en el núcleo del reactor y que cada sistema de medición tiene una probabilidad de 0.01 de fallar. Suponga que una causa potencial para que el sistema de medición falle es que los cables eléctricos que conducen del centro a la sala de control donde se ubican los sistemas de medición podrían quemarse. Alguien desea calcular la probabilidad de que ambos sistemas fallen y hace el cálculo siguiente:

$$\begin{aligned} P(\text{ambos sistemas de medición fallan}) &= P(\text{el primer sistema de medición falla}) \times \\ &\quad P(\text{el segundo sistema de medición falla}) \\ &= (0.01)(0.01) \\ &= 0.0001 \end{aligned}$$

- ¿Qué suposición se debe hacer en este cálculo?
 - Explique por qué esta suposición probablemente no está justificada en este caso.
 - ¿Es la probabilidad 0.0001 posiblemente demasiado alta o demasiado baja? Explique.
- Con referencia al ejercicio 11, ¿es posible que la probabilidad de que fallen ambos sistemas de medición sea más grande de 0.01? Explique.
 - Un lote de diez componentes contiene tres que están defectuosos. Se extraen aleatoriamente dos componentes y se evalúan. Sea A el evento de que el primer componente extraído esté defectuoso y sea B el evento de que el segundo también lo esté.
 - Determine $P(A)$.
 - Determine $P(B|A)$.
 - Determine $P(A \cap B)$.
 - Determine $P(A^c \cap B)$.
 - Determine $P(B)$.
 - ¿Son A y B independientes? Explique.
 - Un lote de mil componentes contiene 300 que están defectuosos. Se sacan aleatoriamente dos componentes y se evalúan. Sea A el evento de que el primer componente extraído esté defectuoso y sea B el evento de que el segundo también lo esté.
 - Determine $P(A)$.
 - Determine $P(B|A)$.
 - Determine $P(A \cap B)$.
 - Determine $P(A^c \cap B)$.
 - Determine $P(B)$.
 - ¿Son A y B independientes? Explique.
 - En un lote de n componentes, 30% está defectuoso. Se extraen aleatoriamente dos componentes y se evalúan. Sea A el evento de que el primer componente extraído esté defectuoso y sea B el evento de que el segundo también lo esté. ¿Para qué tamaño de lote n serán A y B cercanamente independientes: $n = 10$ o $n = 10\,000$? Explique.
 - Dos inspectores de calidad supervisan fallas en artículos. Si se encuentra una, será detectada por el primer inspector con una probabilidad de 0.9 y por el segundo con probabilidad 0.7. Suponga que los inspectores trabajan en forma independiente.
 - Si un artículo tiene una falla, ¿cuál es la probabilidad de que la detecten ambos inspectores?
 - Si un artículo tiene una falla, ¿cuál es la probabilidad de que la detecte al menos uno de los dos inspectores?
 - Suponga que el segundo inspector revisa solamente los artículos que han sido aprobados por el primer inspector. Si un artículo tiene una falla, ¿cuál es la probabilidad de que el segundo inspector la detecte?
 - Con referencia al ejercicio 16, suponga que ambos inspectores supervisan cada artículo y que si uno no tiene falla, entonces ningún inspector la detectará.
 - Suponga que la probabilidad de que un artículo tenga falla es 0.10. Si un artículo es aprobado por el primer inspector, ¿cuál es la probabilidad de que tenga en realidad falla?
 - Suponga que la probabilidad de que un artículo tenga falla es 0.10. Si un artículo es aprobado por ambos inspectores, ¿cuál es la probabilidad de que tenga en realidad una falla?
 - Un programa de control de calidad en una línea de montaje de botellas de plástico implica inspeccionar botellas terminadas para detectar fallas, como huecos microscópicos. La proporción de botellas que tiene tal falla en realidad es de sólo 0.0002. Si una botella tiene una falla, la probabilidad es 0.995 de que no pasará la inspección. Si una botella no tiene falla, la probabilidad es 0.99 de que pasará la inspección.
 - Si una botella no pasa la inspección, ¿cuál es la probabilidad de que tiene falla?

- b) ¿Cuál de las siguientes es la interpretación más correcta de la respuesta del inciso a)?
- La mayoría de las botellas que no pasan la inspección no tienen falla.
 - La mayoría de las botellas que pasan la inspección tienen falla.
- c) Si una botella pasa la inspección, ¿cuál es la probabilidad de que no tenga falla?
- d) ¿Cuál de las siguientes es la interpretación más correcta de la respuesta del inciso c)?
- La mayoría de las botellas que no pasan la inspección tienen falla.
 - La mayoría de las botellas que aprueban la inspección no tienen falla.
- e) Explique por qué una probabilidad pequeña en el inciso a) no es un problema, tan grande como una gran probabilidad del inciso c).

19. Con referencia al ejemplo 2.25 (p. 80).

- Si a un hombre le sale la prueba negativa, ¿cuál es la probabilidad de que en realidad esté enfermo?
- Para muchas pruebas médicas, es un procedimiento usual repetir la prueba cuando sale positiva. Si las pruebas que se repiten son independientes, ¿cuál es la probabilidad de que a un hombre le resulten positivas dos pruebas sucesivas si él está enfermo?
- Suponiendo que las pruebas repetidas son independientes, ¿cuál es la probabilidad de que a un hombre le salgan positivas dos pruebas sucesivas si no tiene la enfermedad?
- Si a un hombre le salen positivas dos pruebas sucesivas, ¿cuál es la probabilidad de que tenga la enfermedad?

20. Un sistema contiene dos componentes, A y B , conectados en serie, como se muestra en el diagrama.

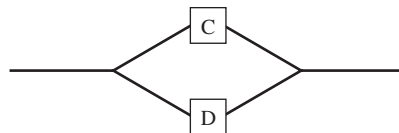


Suponga que A y B funcionan de manera independiente. Para que el sistema funcione, ambos componentes deben funcionar.

- Si la probabilidad de que A falle es 0.05 y la probabilidad de que B falle es 0.03, determine la probabilidad de que el sistema funcione.
- Si tanto A como B tienen probabilidad p de fallar, ¿cuál debe ser el valor de p para que la probabilidad de que el sistema funcione sea 0.90?

- Si tres componentes están conectados en serie y cada uno tiene probabilidad p de fallar, ¿cuál debe ser el valor de p para que la probabilidad de que el sistema funcione sea 0.90?

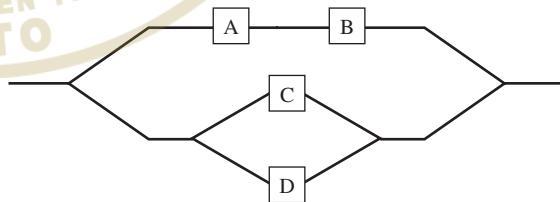
21. Un sistema contiene dos componentes, C y D , conectados en paralelo como se muestra en el diagrama.



Suponga que C y D funcionan de manera independiente. Para que el sistema funcione, deben funcionar C o D .

- Si la probabilidad de que C falle es 0.08 y la probabilidad de que D falle es 0.12, encuentre la probabilidad de que el sistema funcione.
- Si tanto C como D tienen probabilidad p de fallar, ¿cuál debe ser el valor de p para que la probabilidad de que el sistema funcione sea 0.99?
- Si tres componentes están conectados en paralelo, funcionan de manera independiente y cada uno tiene una probabilidad p de fallar, ¿cuál debe ser el valor de p para que la probabilidad de que el sistema funcione sea 0.99?
- Si los componentes funcionan independientemente y cada componente tiene una probabilidad de fallar de 0.5, ¿cuál es el número mínimo de componentes que se debe conectar en paralelo para que la probabilidad de que el sistema funcione sea de al menos 0.99?

22. Un sistema consta de cuatro componentes conectados, como se muestra en el diagrama siguiente.



Suponga que A , B , C y D funcionan de manera independiente. Si las probabilidades de que A , B , C y D fallen son 0.10, 0.05, 0.10 y 0.20, respectivamente, ¿cuál es la probabilidad de que el sistema funcione?

23. Si A y B son eventos independientes, pruebe que los siguientes pares de eventos son independientes: A^c y B , A y B^c y A^c y B^c .

2.4 Variables aleatorias

En muchos casos es deseable asignar un valor numérico a cada resultado de un experimento. Esta asignación se llama **variable aleatoria**. Para aclarar la idea se presenta un ejemplo. Suponga que un ingeniero eléctrico tiene seis resistores a la mano. Tres de ellos tienen etiqueta de $10\ \Omega$ y los otros tres tienen etiqueta de $20\ \Omega$. El ingeniero quiere conectar un resistor de $10\ \Omega$ y un resistor de $20\ \Omega$ en serie, para crear una resistencia de $30\ \Omega$. Ahora se supone que, en efecto, los tres resistores etiquetados con $10\ \Omega$ tienen las resistencias reales de 9, 10 y $11\ \Omega$ y que los tres resistores etiquetados con $20\ \Omega$ tienen las resistencias reales de 19, 20 y $21\ \Omega$. El proceso para seleccionar un resistor de cada tipo es un experimento cuyo espacio muestral consta de nueve resultados igualmente probables. El espacio muestral se presenta en la tabla siguiente.

Resultado	Probabilidad
(9, 19)	1/9
(9, 20)	1/9
(9, 21)	1/9
(10, 19)	1/9
(10, 20)	1/9
(10, 21)	1/9
(11, 19)	1/9
(11, 20)	1/9
(11, 21)	1/9

Ahora lo que es importante para el ingeniero de este experimento es la suma de las dos resistencias, en vez de sus valores individuales. Por tanto, se asigna a cada resultado un número igual a la suma de las dos resistencias seleccionadas. Esta asignación se representa por la letra X y se presenta en la tabla siguiente.

Resultado	X	Probabilidad
(9, 19)	28	1/9
(9, 20)	29	1/9
(9, 21)	30	1/9
(10, 19)	29	1/9
(10, 20)	30	1/9
(10, 21)	31	1/9
(11, 19)	30	1/9
(11, 20)	31	1/9
(11, 21)	32	1/9

La función X , que asigna un valor numérico a cada resultado en el espacio muestral, es una variable aleatoria.

Una **variable aleatoria** asigna un valor numérico a cada resultado en un espacio muestral.

Se acostumbra denotar a las variables aleatorias con letras mayúsculas. Las letras X , Y y Z se usan con más frecuencia.

Se puede calcular las probabilidades para las variables aleatorias de una manera obvia. En el ejemplo anterior, el evento $X = 29$ corresponde al evento $\{(9, 20), (10, 19)\}$ del espacio muestral. Por tanto, $P(X = 29) = P[\{(9, 20), (10, 19)\}] = 2/9$.

Ejemplo 2.30

Haga una lista de los valores posibles de la variable aleatoria X y determine la probabilidad para cada uno de ellos.

Solución

Los valores posibles son 28, 29, 30, 31 y 32. Para encontrar la probabilidad de uno de estos valores, se suman las probabilidades de los resultados en el espacio muestral que corresponden al valor. Los resultados están en la tabla siguiente.

X	$P(X = x)$
28	1/9
29	2/9
30	3/9
31	2/9
32	1/9

La tabla de probabilidades del ejemplo 2.30 contiene toda la información necesaria para calcular cualquier probabilidad que considere a la variable aleatoria X . Observe que los resultados del espacio muestral no se presentan en la tabla. Cuando se conocen las probabilidades en relación con una variable aleatoria, generalmente no se piensa en el espacio muestral: sólo se concentra uno en las probabilidades.

Hay dos tipos importantes de variables aleatorias: **discretas** y **continuas**. Una variable aleatoria discreta es aquella cuyos valores posibles forman un conjunto discreto; en otras palabras, los valores se pueden ordenar y existen separaciones entre los valores adyacentes. La variable aleatoria X , que se acaba de describir, es discreta. En contraparte, los valores posibles de una variable aleatoria continua siempre están contenidos en un intervalo; es decir, son todos los puntos entre dos números. Se proporcionan definiciones precisas de estas clases de variables aleatorias posteriormente en esta sección.

A continuación se presentan más ejemplos de variables aleatorias.

Ejemplo 2.31

Con frecuencia los chips de computadora tienen imperfecciones en su superficie. Para cierto tipo de chip de computadora, 9% no tiene imperfecciones, 22% contiene una imperfección, 26% presenta dos imperfecciones, 20% contiene tres imperfecciones, 12% tiene cuatro im-

perfecciones y 11% presenta cinco imperfecciones. Sea Y el número de imperfecciones en un chip elegido aleatoriamente. ¿Cuáles son los valores posibles de Y ? ¿ Y es discreto o continuo? Determine $P(Y = y)$ para cada valor posible y .

Solución

Los valores posibles para Y son los enteros 0, 1, 2, 3, 4 y 5. La variable aleatoria Y es discreta, ya que solamente toma valores enteros. A 9% de los resultados en el espacio muestral se le asigna el valor 0. Por tanto, $P(Y = 0) = 0.09$. Similarmente $P(Y = 1) = 0.22$, $P(Y = 2) = 0.26$, $P(Y = 3) = 0.20$, $P(Y = 4) = 0.12$ y $P(Y = 5) = 0.11$.

Ejemplo

2.32

Cierto tipo de disco magnético debe funcionar en un ambiente donde está expuesto a gases corrosivos. Se sabe que 10% de estos discos tiene tiempos de vidas menores que o iguales a 100 horas, 50% lo tiene mayor a 100 horas, pero menor o igual a 500, y 40% incluye tiempos superiores a 500 horas. Sea Z el número de horas en tiempo de vida de un disco elegido aleatoriamente. ¿ Z es continua o discreta? Determine $P(Z \leq 500)$. ¿Se pueden calcular todas las probabilidades para Z ? Explique.

Solución

El tiempo de vida de un componente no está limitado a una lista de valores discretamente espaciados; Z es continua. De todos los componentes, 60% tiene vidas menores o iguales a 500 horas. Por tanto, $P(Z \leq 500) = 0.60$. No se tiene la información suficiente para calcular todas las probabilidades para Z . Se pueden calcular algunas de ellas; por ejemplo, $P(Z \leq 100) = 0.10$, $P(100 < Z \leq 500) = 0.50$ y $P(Z > 500) = 0.40$. Pero no se sabe, por ejemplo, la proporción de componentes que tiene tiempos de vida entre 100 y 200 horas, o entre 200 y 300 horas, así que no se puede encontrar la probabilidad de que la variable aleatoria Z tenga cualquiera de estos intervalos. Para calcular todas las probabilidades de Z , se necesitaría calcular la probabilidad para cada intervalo posible; por ejemplo, $P(200 < Z \leq 300)$, $P(200 < Z \leq 201)$, $P(200 < Z \leq 200.1)$ y así sucesivamente. Se verá cómo se puede hacer lo anterior en esta sección, cuando se analicen las variables aleatorias continuas.

Variables aleatorias y poblaciones

Con frecuencia es útil pensar en un valor de una variable aleatoria como que se ha seleccionado de una población. Por ejemplo, considere la variable aleatoria Y descrita en el ejemplo 2.31. Observar un valor para esta variable aleatoria es como seleccionar un valor de una población que consta de los enteros 0, 1, 2, 3, 4 y 5, en las proporciones siguientes: 0's, 9%; 1's, 22%; 2's, 26%; 3's, 20%; 4's, 12%; y 5's, 11%. Para una variable aleatoria continua es adecuado imaginar una población infinita que contiene a todos los valores posibles de la variable aleatoria. Por ejemplo, para la variable aleatoria Z del ejemplo 2.32 se podría imaginar una población que contenga a todos los números positivos, con 10% de los valores de población menores o iguales a 100, 50% superiores a 100 pero menores o iguales a 500, y 40% mayores a 500. La proporción de valores de la población en cualquier intervalo sería igual a la probabilidad de que la variable Z está en ese intervalo.

Los métodos para trabajar con variables aleatorias son diferentes para variables aleatorias discretas y continuas. Se inicia con las variables aleatorias discretas.

Variables aleatorias discretas

Se inicia con la definición de una variable aleatoria discreta.

Definición

Una variable aleatoria es **discreta** si sus valores posibles constituyen un conjunto discreto. Lo anterior significa que si los valores posibles se ordenan, hay una separación entre cada valor y el próximo. El conjunto de valores posibles podría ser infinito; por ejemplo, el conjunto de todos los enteros o el conjunto de todos los enteros positivos.

Es común que los valores posibles de una variable aleatoria discreta sean un conjunto de enteros. Para cualquier variable aleatoria discreta, si se especifica la lista de sus valores posibles junto con la probabilidad que tiene la variable aleatoria en cada uno de estos valores, entonces se ha descrito **completamente a la población** a partir de la cual se seleccionó a la variable aleatoria. Esto se ilustra con un ejemplo.

El número de fallas en un alambre de cobre de 1 pulg de longitud, fabricado en proceso específico, varía de alambre en alambre. En conjunto, 48% de los alambres producidos no tiene falla, 39% presenta una, 12% fue detectado con dos y 1% tiene tres. Sea X el número de fallas en una pieza de alambre seleccionada aleatoriamente. Entonces

$$P(X = 0) = 0.48 \quad P(X = 1) = 0.39 \quad P(X = 2) = 0.12 \quad P(X = 3) = 0.01$$

La lista de valores posible 0, 1, 2, 3, junto con las probabilidades para cada uno, proporciona una descripción completa de la población de la que se tomó a X . Esta descripción se conoce como **función de masa de probabilidad**.

Definición

La **función de masa de probabilidad** de una variable aleatoria discreta X es la función $p(x) = P(X = x)$. A veces a la función de masa de probabilidad se le llama **distribución de probabilidad**.

Por tanto, para la variable aleatoria X que representa el número de fallas en una longitud de alambre, $p(0) = 0.48$, $p(1) = 0.39$, $p(2) = 0.12$, $p(3) = 0.01$ y $p(x) = 0$ para cualquier valor de x diferente de 0, 1, 2 o 3. Observe que si se suman los valores de la función de masa de probabilidad sobre todos los valores posibles de X , la suma es igual a 1. Esto es cierto para cualquier función de masa de probabilidad. La razón es que al sumar los valores de una función de masa de probabilidad sobre todos los valores posibles de la variable aleatoria correspondiente, se obtiene la probabilidad de que la variable aleatoria es igual a uno de sus valores posibles y esta probabilidad es siempre igual a 1.

La función de masa de probabilidad se puede representar por un diagrama en el cual se dibuja una recta vertical para cada uno de los valores posibles de la variable aleatoria. Las alturas de las rectas son iguales a las probabilidades de los valores correspondientes. La inter-

pretación física de este diagrama es que cada recta representa una masa igual a su altura. En la figura 2.7 se muestra un diagrama de la función de masa de probabilidad de la variable aleatoria X .

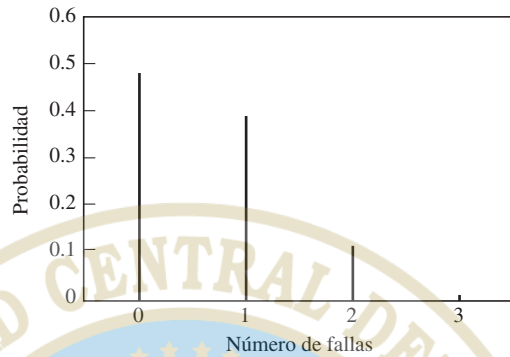


FIGURA 2.7 Función de masa de probabilidad de X , el número de fallas en una elección aleatoria de una pieza de alambre.

Función de distribución acumulativa de una variable aleatoria discreta

La función de masa de probabilidad especifica la probabilidad de que una variable aleatoria sea igual a un valor determinado. La **función de distribución acumulativa** especifica la probabilidad de que una variable aleatoria sea menor o igual a un valor dado. La función de distribución acumulativa de la variable aleatoria X es la función $F(x) = P(X \leq x)$.

Ejemplo

2.33

Sea $F(x)$ la función de distribución acumulativa de la variable aleatoria X que representa el número de fallas en un alambre elegido aleatoriamente. Determine $F(2)$ y $F(1.5)$.

Solución

Puesto que $F(2) = P(X \leq 2)$ se necesita encontrar $P(X \leq 2)$. Con dicho propósito se suman las probabilidades de los valores de X que son menores o iguales a 2, a saber, 0, 1 y 2. Por tanto,

$$\begin{aligned}
 F(2) &= P(X \leq 2) \\
 &= P(X = 0) + P(X = 1) + P(X = 2) \\
 &= 0.48 + 0.39 + 0.12 \\
 &= 0.99
 \end{aligned}$$

Ahora $F(1.5) = P(X \leq 1.5)$. Por tanto, para calcular $F(1.5)$ se debe sumar las probabilidades para los valores de X que son menores o iguales a 1.5, que son 0 y 1. Por lo que,

$$\begin{aligned}
 F(1.5) &= P(X \leq 1.5) \\
 &= P(X = 0) + P(X = 1) \\
 &= 0.48 + 0.39 \\
 &= 0.87
 \end{aligned}$$

En general, para cualquier variable aleatoria discreta X , la función de distribución acumulativa $F(x)$ se puede calcular sumando las probabilidades de todos los valores posibles de X que son menores o iguales a x . Observe que $F(x)$ está definido para cualquier número x , no exactamente para los valores posibles de X .

Resumen

Sea X una variable aleatoria discreta. Entonces

- La función de masa de probabilidad de X es la función $p(x) = P(X = x)$.
- La función de distribución acumulativa de X es la función $F(x) = P(X \leq x)$.
- $F(x) = \sum_{t \leq x} p(t) = \sum_{t \leq x} P(X = t)$.
- $\sum_x p(x) = \sum_x P(X = x) = 1$, donde la sumatoria se realiza sobre todos los valores posibles de X .

Ejemplo 2.34

Dibuje la función de distribución acumulativa $F(x)$ de la variable aleatoria X que representa el número de fallas en un alambre elegido aleatoriamente.

Solución

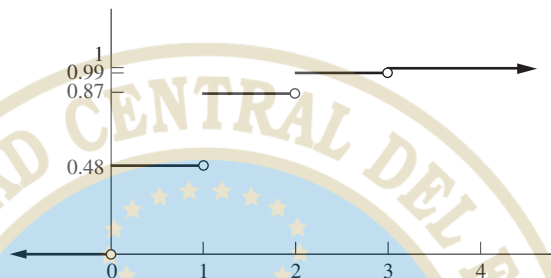
Primero se calcula $F(x)$ para cada uno de los valores posibles de X , que son 0, 1, 2 y 3.

$$\begin{aligned}
 F(0) &= P(X \leq 0) = 0.48 \\
 F(1) &= P(X \leq 1) = 0.48 + 0.39 = 0.87 \\
 F(2) &= P(X \leq 2) = 0.48 + 0.39 + 0.12 = 0.99 \\
 F(3) &= P(X \leq 3) = 0.48 + 0.39 + 0.12 + 0.01 = 1
 \end{aligned}$$

Para cualquier valor x , se calcula a $F(x)$ sumando las probabilidades de todos los valores posibles de X que son menores o iguales a x . Por ejemplo, si $1 \leq x < 2$, los valores posibles de X que son menores o iguales a x son 0 y 1, por lo que $F(x) = P(X = 0) + P(X = 1) = F(1) = 0.87$. Por tanto,

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.48 & 0 \leq x < 1 \\ 0.87 & 1 \leq x < 2 \\ 0.99 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

En la figura siguiente se presenta una gráfica de $F(x)$.



Para una variable aleatoria discreta, la gráfica de $F(x)$ consta de una serie de rectas horizontales (llamadas “pasos”) con saltos en cada uno de los valores posibles de X . Observe que el tamaño del salto en cualquier punto x es igual al valor de la función de masa de probabilidad $p(x) = P(X = x)$.

Media y varianza para variables aleatorias discretas

En la sección 1.3 se vio que la media de una muestra es aproximadamente igual a la componente horizontal del centro de masa de un histograma de la muestra, que es el punto en el eje x en el cual el histograma mantendría el equilibrio si se sostuviera de allí. Por analogía, se define a la **media poblacional** como la componente horizontal del centro de masa de la gráfica de su función de masa de probabilidad de una variable aleatoria discreta. La media poblacional de una variable aleatoria X que también se puede llamar **esperanza**, o **valor esperado**, de X y que se denota por μ_x , por $E(X)$, o simplemente por μ . A veces se puede eliminar la palabra “poblacional” y sólo se hará referencia a la media poblacional como la media.

Ejemplo

2.35

Determine la media de la variable aleatoria X que representa el número de fallas en una pieza de alambre elegida aleatoriamente.

Solución

La media es el centro de masa de la gráfica de la función de masa de probabilidad (figura 2.7). El centro de masa se calcula multiplicando la altura de cada recta por su valor en el eje horizontal y después se suman los productos. Los valores son 0, 1, 2 y 3. Las alturas son $P(X = 0) = 0.48$, $P(X = 1) = 0.39$, $P(X = 2) = 0.12$ y $P(X = 3) = 0.01$. La media es, por tanto,

$$\begin{aligned}
 \mu_X &= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) \\
 &= (0)(0.48) + (1)(0.39) + (2)(0.12) + (3)(0.01) \\
 &= 0.6600
 \end{aligned}$$

En general, la media de una variable aleatoria discreta se encuentra al multiplicar cada valor posible de la variable aleatoria por su probabilidad y después se suman.

Definición

Sea X una variable aleatoria discreta con función de masa de probabilidad $p(x) = P(X = x)$.

La media de X está dada por

$$\mu_X = \sum_x x P(X = x) \quad (2.29)$$

donde la sumatoria se hace sobre todos los valores posibles de X .

A veces la media de X se llama esperanza, o valor esperado, de X y también se denota por $E(X)$ o por μ .

En la sección 1.3 se mostró que la varianza de una muestra es aproximadamente igual al momento de inercia del histograma de la muestra alrededor de la media de la muestra. Se define a la **varianza poblacional** de una variable aleatoria discreta como el momento de inercia de la gráfica de su función de masa de probabilidad con respecto a la media poblacional μ . La varianza poblacional de una variable aleatoria X con frecuencia se denomina simplemente **varianza** de X . Se puede denotar por σ_X^2 , por $V(X)$, o simplemente por σ^2 . Para calcularla se multiplica la altura de cada recta del diagrama de la función de masa de probabilidad por el cuadrado de su distancia horizontal a la poblacional y después se suman los productos. Es más fácil de comprender esto cuando se presenta en una fórmula:

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x)$$

Realizando un poco de álgebra se puede obtener una fórmula alternativa

$$\sigma_X^2 = \sum_x x^2 P(X = x) - \mu_X^2$$

Al final de esta sección se presenta una deducción de esta fórmula.

También se define la **desviación estándar poblacional** como la raíz cuadrada de la varianza poblacional. Se denota a la desviación estándar de la población de una variable aleatoria X por σ_X o simplemente por σ . Como con la media, a veces se eliminará la palabra “poblacional” y sólo se denominará a la poblacional y la desviación estándar poblacional como la varianza y la desviación estándar, respectivamente.

Resumen

Sea X una variable aleatoria discreta con función de masa de probabilidad $p(x) = P(X = x)$. Entonces

- La varianza de X está dada por

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x) \quad (2.30)$$

- Una fórmula alternativa para la varianza está dada por

$$\sigma_X^2 = \sum_x x^2 P(X = x) - \mu_X^2 \quad (2.31)$$

- La varianza de X también se puede denotar por $V(X)$ o por σ^2 .
- La desviación estándar es la raíz cuadrada de la varianza: $\sigma_X = \sqrt{\sigma_X^2}$.

Ejemplo

2.36

Determine la varianza y la desviación estándar para la variable aleatoria X que representa el número de fallas de una pieza de alambre elegida aleatoriamente.

Solución

En el ejemplo 2.35 se calculó la media de X con un valor de $\mu_X = 0.6600$. Se encuentra la varianza utilizando la ecuación (2.30):

$$\begin{aligned} \sigma_X^2 &= (0 - 0.6600)^2 P(X = 0) + (1 - 0.6600)^2 P(X = 1) + (2 - 0.6600)^2 P(X = 2) \\ &\quad + (3 - 0.6600)^2 P(X = 3) \\ &= (0.4356)(0.48) + (0.1156)(0.39) + (1.7956)(0.12) + (5.4756)(0.01) \\ &= 0.5244 \end{aligned}$$

La desviación estándar es $\sigma_X = \sqrt{0.5244} = 0.724$.

Ejemplo

2.37

Use la fórmula alternativa, ecuación (2.31), para calcular la varianza de X , del número de fallas de un alambre elegido aleatoriamente.

Solución

En el ejemplo 2.35 se calculó que la media tenía un valor de $\mu_X = 0.6600$. La varianza es, por tanto,

$$\begin{aligned} \sigma_X^2 &= 0^2 P(X = 0) + 1^2 P(X = 1) + 2^2 P(X = 2) + 3^2 P(X = 3) - (0.6600)^2 \\ &= (0)(0.48) + (1)(0.39) + (4)(0.12) + (9)(0.01) - (0.6600)^2 \\ &= 0.5244 \end{aligned}$$

Ejemplo

2.38

Un resistor en un circuito dado tiene una resistencia en el rango de $99\ \Omega$ - $101\ \Omega$. Un ingeniero obtiene dos resistores. La probabilidad de que ambos satisfagan la especificación es 0.36, la probabilidad de que exactamente uno de ellos satisfaga la especificación es 0.48 y la probabilidad de que ninguno de ellos lo haga es 0.16. Sea X el número de resistores que satisface la especificación. Determine la función de masa de probabilidad, media, varianza y desviación estándar de X .

Solución

La función de masa de la probabilidad es $P(X = 0) = 0.16$, $P(X = 1) = 0.48$, $P(X = 2) = 0.36$ y $P(X = x) = 0$ para $x \neq 0, 1$ o 2 . La media es

$$\begin{aligned}\mu_X &= (0)(0.16) + (1)(0.48) + (2)(0.36) \\ &= 1.200\end{aligned}$$

La varianza es

$$\begin{aligned}\sigma_X^2 &= (0 - 1.200)^2(0.16) + (1 - 1.200)^2(0.48) + (2 - 1.200)^2(0.36) \\ &= 0.4800\end{aligned}$$

La desviación estándar es $\sigma_X = \sqrt{0.4800} = 0.693$.

Histograma de probabilidad

Cuando los valores posibles de una variable aleatoria discreta están espaciados uniformemente, la función de masa de probabilidad se puede representar por medio de un histograma, con rectángulos centrados en los valores posibles de la variable aleatoria. El área de un rectángulo centrado en un valor x es igual a $P(X = x)$. Este histograma se llama un **histograma de probabilidad**, ya que las áreas representan probabilidades. En la figura 2.7 se presenta el diagrama de la función de masa de probabilidad de una variable aleatoria X que representa el número de fallas en un alambre. La figura 2.8 presenta un histograma de probabilidad para esta variable aleatoria.

La probabilidad de que el valor de una variable aleatoria esté en un intervalo específico está dada por el área bajo el histograma de probabilidad. En el ejemplo 2.39 se ilustra la idea.

Ejemplo

2.39

Determine la probabilidad de que un alambre elegido aleatoriamente tenga más de una falla. Indique esta probabilidad como un área bajo el histograma de probabilidad.

Solución

Se desea encontrar $P(X > 1)$. Puesto que ningún alambre tiene más de tres fallas, la proporción de que los alambres tengan más de una se puede encontrar al sumar la proporción de que tengan dos más la proporción de que tengan tres. Simbólicamente, $P(X > 1) = P(X = 2) + P(X = 3)$. La función de masa de probabilidad especifica que $P(X = 2) = 0.12$ y $P(X = 3) = 0.01$. Por tanto, $P(X > 1) = 0.12 + 0.01 = 0.13$.

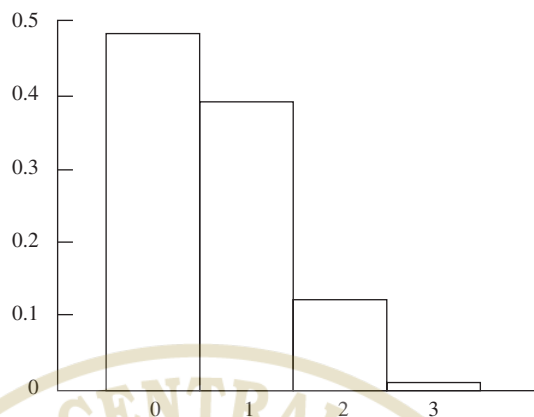


FIGURA 2.8 Histograma de probabilidad para X , el número de fallas en una pieza de alambre elegida aleatoriamente. Compárela con la figura 2.7.

Esta probabilidad está dada por el área bajo el histograma de probabilidad que corresponde a esos rectángulos centrados en valores superiores a 1 (véase la figura 2.9). Hay dos de estos rectángulos; sus áreas son $P(X = 2) = 0.12$ y $P(X = 3) = 0.01$. Ésta es otra manera de mostrar que $P(X > 1) = 0.12 + 0.01 = 0.13$.

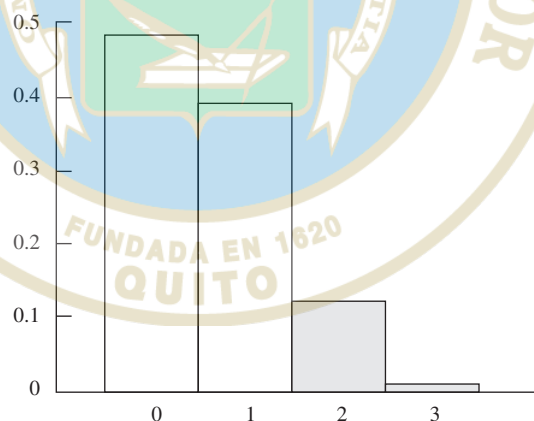


FIGURA 2.9 Histograma de probabilidad para X , el número de fallas en una pieza de alambre elegida aleatoriamente. El área que corresponde a valores de X mayores que 1 está sombreada. Esta área es igual a $P(X > 1)$.

En el capítulo 4 se verá que las probabilidades para las variables aleatorias discretas se pueden aproximar a veces calculando el área bajo una curva. Representar las probabilidades discretas con un histograma de probabilidad hará más fácil comprender cómo se hace esto último.

Variables aleatorias continuas

En la tabla 1.4 (en la sección 1.3) se presentaron intervalos de clase para las emisiones, en gramos de partículas por galón de combustible consumido, de una muestra de 62 vehículos. Observe que las emisiones constituyen una variable continua, ya que sus valores posibles no están restringidos a algún conjunto discretamente espaciado. Los intervalos de clase son elegidos para que cada intervalo contenga un número razonablemente grande de vehículos. Si la muestra fuera más grande, se podrían hacer los intervalos más angostos. En particular, si se tuviera información sobre toda la población, que contiene millones de vehículos, se podrían hacer los intervalos extremadamente angostos. El histograma parecería entonces muy suave y se podría aproximar con una curva, como la que se muestra en la figura 2.10.

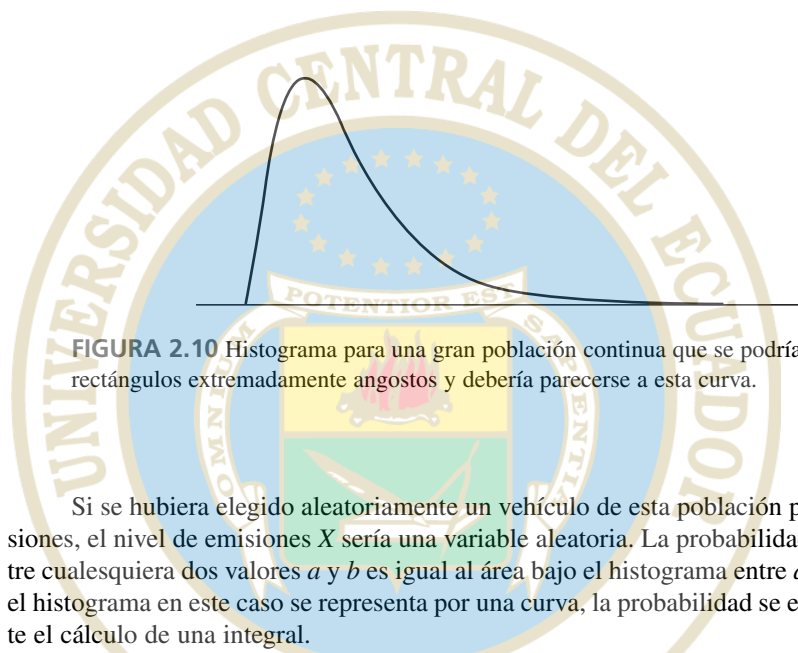


FIGURA 2.10 Histograma para una gran población continua que se podría dibujar con rectángulos extremadamente angostos y debería parecerse a esta curva.

Si se hubiera elegido aleatoriamente un vehículo de esta población para medir sus emisiones, el nivel de emisiones X sería una variable aleatoria. La probabilidad de que X esté entre cualesquiera dos valores a y b es igual al área bajo el histograma entre a y b . Debido a que el histograma en este caso se representa por una curva, la probabilidad se encontraría mediante el cálculo de una integral.

La variable aleatoria X descrita aquí es un ejemplo de una **variable aleatoria continua**. Ésta se define como una variable aleatoria cuyas probabilidades se representan por áreas bajo una curva. Esta curva se llama **función de densidad de probabilidad**. Como consecuencia de que la función de densidad de probabilidad es una curva, los cálculos de las probabilidades implican integrales, en vez de las sumatorias que se usan en el caso discreto.

Definición

Una variable aleatoria es **continua** si sus probabilidades están dadas por áreas bajo una curva. La curva se llama **función de densidad de probabilidad** para la variable aleatoria.

A veces la función de densidad de probabilidad se llama **distribución de probabilidad**.

Cálculo de probabilidades con la función de densidad de probabilidad

Sea X una variable aleatoria continua. Sea la función $f(x)$ la función de densidad de probabilidad de X . Sean a y b cualesquiera dos números, con $a < b$.

La proporción de la población cuyos valores de X están entre a y b están dados por $\int_a^b f(x) dx$, y el área bajo la función de densidad de probabilidad entre a y b . Ésta es la probabilidad de que la variable aleatoria X tome un valor entre a y b . Observe que el área bajo la curva no depende de si los puntos finales a y b estén incluidos en el intervalo. Por tanto, las probabilidades que implican a X no dependen de si se incluyen a los puntos finales.

Resumen

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Sean a y b cualesquiera dos números, con $a < b$. Entonces

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = \int_a^b f(x) dx$$

Además,

$$P(X \leq a) = P(X < a) = \int_{-\infty}^a f(x) dx \quad (2.32)$$

$$P(X \geq a) = P(X > a) = \int_a^{\infty} f(x) dx \quad (2.33)$$

Si $f(x)$ es la función de densidad de probabilidad de una variable aleatoria X , entonces el área bajo toda la curva desde $-\infty$ a ∞ es la probabilidad de que el valor de X esté entre $-\infty$ y ∞ . Esta probabilidad debe ser igual a 1, ya que el valor de X siempre está entre $-\infty$ y ∞ . Por tanto, el área bajo toda la curva $f(x)$ es igual a 1.

Resumen

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Entonces

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Ejemplo

2.40

Se perfora un hueco en un componente de una hoja de metal y después se inserta un eje a través del hueco. La holgura del eje es igual a la diferencia entre el radio del hueco y el radio del eje. Sea X la variable aleatoria que denota a la holgura, en milímetros. La función de densidad de probabilidad de X es

$$f(x) = \begin{cases} 1.25(1 - x^4) & 0 < x < 1 \\ 0 & \text{de otro modo} \end{cases}$$

Los componentes con holguras superiores a 0.8 mm se deben desechar. ¿Cuál es la proporción de componentes que serán desechados?

Solución

En la figura 2.11 se presenta la función de densidad de probabilidad de X . Observe que la densidad $f(x)$ es 0 para $x \leq 0$ y para $x \geq 1$. Esto indica que las holguras siempre están entre 0 y 1 mm. La proporción de componentes que se debe desechar es $P(X > 0.8)$, que es igual al área bajo la función de densidad de probabilidad a la derecha de 0.8.

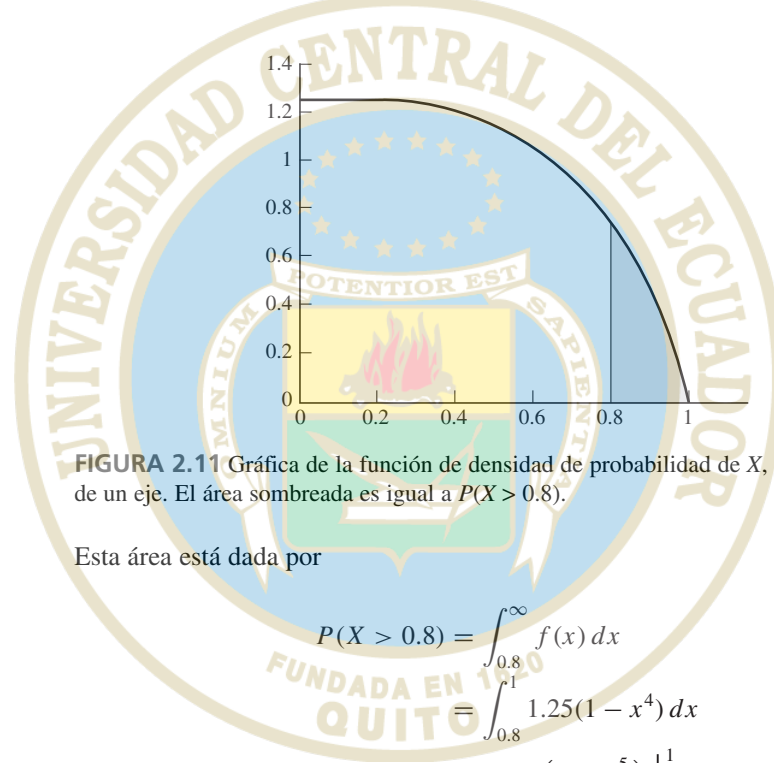


FIGURA 2.11 Gráfica de la función de densidad de probabilidad de X , la holgura de un eje. El área sombreada es igual a $P(X > 0.8)$.

Esta área está dada por

$$\begin{aligned} P(X > 0.8) &= \int_{0.8}^{\infty} f(x) dx \\ &= \int_{0.8}^1 1.25(1 - x^4) dx \\ &= 1.25 \left(x - \frac{x^5}{5} \right) \Big|_{0.8}^1 \\ &= 0.0819 \end{aligned}$$

Función de distribución acumulativa de una variable aleatoria continua

La función de distribución acumulativa de una variable aleatoria continua X es $F(x) = P(X \leq x)$, al igual que para una variable aleatoria discreta. Para esta última, $F(x)$ se puede encontrar al sumar los valores de la función de masa de probabilidad. Para una variable aleatoria continua, el valor de $F(x)$ se obtiene al integrar la función de densidad de probabilidad. Puesto que

$F(x) = P(X \leq x)$ y usando la ecuación (2.32) se tiene que $F(x) = \int_{-\infty}^x f(t) dt$, donde $f(t)$ es la función de densidad de probabilidad.

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. La función de distribución acumulativa de X es la función

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad (2.34)$$

Ejemplo 2.41

Con referencia al ejemplo 2.40, determine la función de distribución acumulativa $F(x)$ y dibújela.

Solución

La función de densidad de probabilidad X está dada por $f(t) = 0$ si $t \leq 0$, $f(t) = 1.25(1 - t^4)$ si $0 < t < 1$ y $f(t) = 0$ si $t \geq 1$. La función de distribución acumulativa está dada por $F(x) = \int_{-\infty}^x f(t) dt$. Puesto que $f(t)$ está definida por partes en tres intervalos diferentes, el cálculo de la función de distribución acumulativa implica tres casos distintos.

Si $x \leq 0$:

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt \\ &= \int_{-\infty}^x 0 dt \\ &= 0 \end{aligned}$$

Si $0 < x < 1$:

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt \\ &= \int_{-\infty}^0 f(t) dt + \int_0^x f(t) dt \\ &= \int_{-\infty}^0 0 dt + \int_0^x 1.25(1 - t^4) dt \\ &= 0 + 1.25 \left(t - \frac{t^5}{5} \right) \Big|_0^x \\ &= 1.25 \left(x - \frac{x^5}{5} \right) \end{aligned}$$

Si $x > 1$:

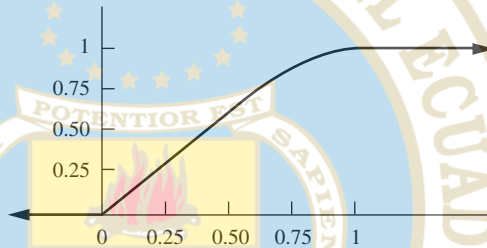
$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt \\ &= \int_{-\infty}^0 f(t) dt + \int_0^1 f(t) dt + \int_1^x f(t) dt \end{aligned}$$

$$\begin{aligned}
 &= \int_{-\infty}^0 0 \, dt + \int_0^1 1.25(1 - t^4) \, dt + \int_1^x 0 \, dt \\
 &= 0 + 1.25 \left(t - \frac{t^5}{5} \right) \Big|_0^1 + 0 \\
 &= 0 + 1 + 0 \\
 &= 1
 \end{aligned}$$

Por tanto,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1.25 \left(x - \frac{x^5}{5} \right) & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

Aquí se presenta una gráfica de $F(x)$.



Observe que la función de distribución acumulativa $F(x)$ del ejemplo 2.41 representa una función continua si su gráfica no presenta ningún salto. Ésta es una característica de las variables aleatorias continuas. La función de distribución acumulativa de una variable aleatoria continua será continua siempre, mientras que la función de distribución acumulativa de una variable aleatoria no continua nunca será continua.

Ejemplo 2.42

Con referencia al ejemplo 2.40, use la función de distribución acumulativa para encontrar la probabilidad de que la holgura del eje es menor a 0.5 mm.

Solución

Sea X la holgura del eje. Se tiene que encontrar $P(X \leq 0.5)$. Éste es equivalente a encontrar $F(0.5)$, donde $F(x)$ es la función de distribución acumulativa. Usando los resultados del ejemplo 2.41, $F(0.5) = 1.25(0.5 - 0.5^5/5) = 0.617$.

Media y varianza para variables aleatorias continuas

La poblacional y la varianza de una variable aleatoria continua están definidas de la misma forma que para una variable aleatoria discreta, excepto que se usa la función de densidad de probabilidad en lugar de la función de masa. Específicamente, si X constituye una variable

aleatoria continua, su poblacional se define como el centro de masa de su función de densidad de probabilidad y su poblacional representa el momento de inercia con respecto a un eje vertical que pasa a través del centro de masa. Las fórmulas son análogas a las ecuaciones de la (2.29) a la (2.31), con las sumatorias reemplazadas por integrales.

Como en el caso de las variables aleatorias discretas, a veces se eliminará la palabra “poblacional” y se hará referencia a la media poblacional, a la varianza poblacional y la desviación poblacional sólo como la media, la varianza y la desviación estándar, respectivamente.

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Entonces la media de X está dada por

$$\mu_X = \int_{-\infty}^{\infty} x f(x) dx \quad (2.35)$$

A la media de X algunas veces se le llama esperanza, o valor esperado, de X y se puede denotar también por $E(X)$ o por μ .

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Entonces

- La varianza de X está dada por

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx \quad (2.36)$$

- Una fórmula alternativa para la varianza está dada por

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu_X^2 \quad (2.37)$$

- La varianza de X también se puede denotar por $V(X)$ o por σ^2 .
- La desviación estándar es la raíz cuadrada de la varianza: $\sigma_X = \sqrt{\sigma_X^2}$.

Ejemplo

2.43

Con referencia al ejemplo 2.40, determine la media y la varianza de la holgura.

Solución

En la ecuación (2.35), la media de la holgura está dada por

$$\begin{aligned} \mu_X &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^1 x [1.25(1 - x^4)] dx \end{aligned}$$

$$\begin{aligned}
 &= 1.25 \left(\frac{x^2}{2} - \frac{x^6}{6} \right) \Big|_0^1 \\
 &= 0.4167
 \end{aligned}$$

Una vez que se ha calculado $\mu_X = 0.4167$, ahora se puede calcular σ_X^2 . Es más fácil usar la fórmula alternativa, ecuación (2.37):

$$\begin{aligned}
 \sigma_X^2 &= \int_{-\infty}^{\infty} x^2 f(x) dx - \mu_X^2 \\
 &= \int_0^1 x^2 [1.25(1 - x^4)] dx - (0.4167)^2 \\
 &= 1.25 \left(\frac{x^3}{3} - \frac{x^7}{7} \right) \Big|_0^1 - (0.4167)^2 \\
 &= 0.0645
 \end{aligned}$$

Mediana poblacional y percentiles

En la sección 1.2 se definió la mediana de una muestra como el número de en medio, el promedio de los dos números intermedios, cuando los valores de la muestra están ordenados del más pequeño al más grande. Intuitivamente, la mediana de la muestra constituye el punto que divide a la muestra por la mitad. La mediana poblacional se define en forma similar. En relación con la función de densidad de probabilidad, la mediana representa el punto para el que la mitad del área bajo la curva está a la izquierda y la otra mitad a la derecha. En este contexto, si X es una variable aleatoria continua con función de densidad de probabilidad $f(x)$, la mediana de X es el punto x_m que resuelve la ecuación $P(X \leq x_m) = \int_{-\infty}^{x_m} f(x) dx = 0.5$.

La mediana es un caso especial de un percentil. Sea $0 < p < 100$. El p -ésimo percentil de una población es el valor x_p de forma que el $p\%$ de los valores de población sean menores o iguales a x_p . Por tanto, si X es una variable aleatoria continua con función de densidad de probabilidad $f(x)$, el p -ésimo percentil de X es el punto x_p que resuelve la ecuación $P(X \leq x_p) = \int_{-\infty}^{x_p} f(x) dx = p/100$. Observe que la mediana es el 50avo. percentil. La figura 2.12 muestra a la mediana y al 90avo. percentil para una población hipotética.

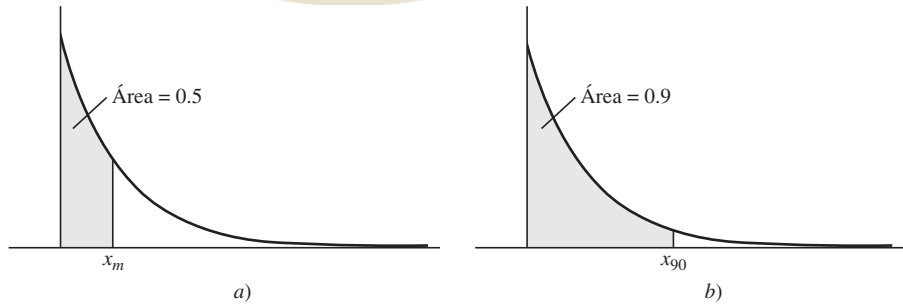


FIGURA 2.12 a) La mitad de los valores de la población son menores que la mediana x_m .
b) 90% de los valores de población que son menores que el 90avo. percentil x_{90} .

Definición

Sea X una variable aleatoria continua con función de masa de probabilidad $f(x)$ y función de distribución acumulativa $F(x)$.

- La mediana de X es el punto x_m que resuelve la ecuación

$$F(x_m) = P(X \leq x_m) = \int_{-\infty}^{x_m} f(x) dx = 0.5.$$

- Si p es cualquier número entre 0 y 100, el p -ésimo percentil es el punto x_p que resuelve la ecuación $F(x_p) = P(X \leq x_p) = \int_{-\infty}^{x_p} f(x) dx = p/100$.
- La mediana es el 50avo. percentil.

Se observa que es posible construir variables aleatorias continuas para las cuales hay un intervalo de puntos que satisfacen la definición de la mediana o de un percentil. Tales variables aleatorias rara vez se encuentran en la práctica.

Ejemplo 2.44

Cierta masa radiactiva emite partículas alfa periódicamente. El tiempo entre emisiones, en segundos, es aleatorio, con función de densidad de probabilidad

$$f(x) = \begin{cases} 0.1e^{-0.1x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Determine la mediana del tiempo entre emisiones. Determine el 60avo. percentil de los tiempos.

Solución

La mediana x_m es la solución de $\int_{-\infty}^{x_m} f(x) dx = 0.5$. Por tanto, se debe resolver

$$\int_0^{x_m} 0.1e^{-0.1x} dx = 0.5$$

$$-e^{-0.1x} \Big|_0^{x_m} = 0.5$$

$$1 - e^{-0.1x_m} = 0.5$$

$$e^{-0.1x_m} = 0.5$$

$$-0.1x_m = \ln 0.5$$

$$0.1x_m = 0.6931$$

$$x_m = 6.931 \text{ s}$$

La mitad de los tiempos entre emisiones es menor a 6.931 s y la mitad es más grande.

El 60avo. percentil x_{60} es la solución de $\int_{-\infty}^{x_{60}} f(x) dx = 0.6$. Al proceder como antes, se sustituye x_{60} por x_m y 0.6 por 0.5. Con lo que se obtiene

$$1 - e^{-0.1x_{60}} = 0.6$$

$$e^{-0.1x_{60}} = 0.4$$

$$-0.1x_{60} = \ln 0.4$$

$$0.1x_{60} = 0.9163$$

$$x_{60} = 9.163 \text{ s}$$

Sesenta por ciento de los tiempos entre emisiones es menor a 9.163 s y 40% es más grande.

Deducción de la fórmula alternativa para la varianza

Para deducir la ecuación (2.31) se inicia con la ecuación (2.30).

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x)$$

Al desarrollar el binomio al cuadrado $(x - \mu_X)^2$ se obtiene

$$\sigma_X^2 = \sum_x (x^2 - 2x\mu_X + \mu_X^2) P(X = x)$$

Al distribuir el término $P(X = x)$ sobre los términos entre paréntesis se tiene que

$$\sigma_X^2 = \sum_x [x^2 P(X = x) - 2x\mu_X P(X = x) + \mu_X^2 P(X = x)]$$

Sumando los términos por separado,

$$\sigma_X^2 = \sum_x x^2 P(X = x) - \sum_x 2x\mu_X P(X = x) + \sum_x \mu_X^2 P(X = x) \quad (2.38)$$

Ahora $\sum_x 2x\mu_X P(X = x) = 2\mu_X \sum_x x P(X = x) = 2\mu_X \mu_X = 2\mu_X^2$, y $\sum_x \mu_X^2 P(X = x) = \mu_X^2 \sum_x P(X = x) = \mu_X^2 (1) = \mu_X^2$.

Sustituyendo en la ecuación (2.38) se obtiene

$$\sigma_X^2 = \sum_x x^2 P(X = x) - 2\mu_X^2 + \mu_X^2$$

Se concluye que

$$\sigma_X^2 = \sum_x x^2 P(X = x) - \mu_X^2$$

Para obtener la fórmula alternativa (2.37) a partir de la ecuación (2.36) para la varianza de una variable aleatoria continua, se pueden usar los mismos pasos; reemplazando $\sum_x$ con $\int_{-\infty}^{\infty}$ y $P(X = x)$ con $f(x)dx$.

Ejercicios para la sección 2.4

1. Determine si cada una de las siguientes variables aleatorias es discreta o continua.
 - a) El número de caras en 100 lanzamientos a “cara” o “cruz” de una moneda.
 - b) La longitud de una varilla elegida aleatoriamente de la producción de un día.
 - c) El puntaje del examen final de un estudiante elegido aleatoriamente de la clase de estadística de ingeniería del último semestre.
 - d) La edad de un estudiante elegido aleatoriamente de la Escuela de Minas, de Colorado.

- c) La edad que tendrá un estudiante de la Escuela de Minas, de Colorado, elegida aleatoriamente en su próximo cumpleaños.
2. La tabla siguiente presenta la función de masa de probabilidad del número de defectos X en un tablero de circuitos impresos elegido aleatoriamente.

x	0	1	2	3
$p(x)$	0.5	0.3	0.1	0.1

- a) Determine $P(X < 2)$.
- b) Determine $P(X \geq 1)$.
- c) Determine μ_X .
- d) Determine σ_X^2 .
3. Una compañía de materiales químicos envía cierto disolvente en tambores de diez galones. Sea X el número de tambores pedidos por un cliente elegido aleatoriamente. Suponga que X tiene la siguiente función de masa de probabilidad:

x	1	2	3	4	5
$p(x)$	0.4	0.2	0.2	0.1	0.1

- a) Determine la media del número de tambores ordenados.
- b) Determine la varianza del número de tambores ordenados.
- c) Determine la desviación estándar del número de tambores ordenados.
- d) Sea Y el número de galones ordenados. Determine la función de masa de probabilidad de Y .
- e) Determine la media del número de galones ordenados.
- f) Determine la varianza del número de galones ordenados.
- g) Determine la desviación estándar del número de galones ordenados.
4. Un automóvil viejo con un motor de cuatro cilindros es llevado a un taller para ajustarlo. Sea X el número de cilindros con compresión baja.

- a) ¿Cuál de las tres funciones dadas en la tabla siguiente es una función de masa de probabilidad posible de X ? Explique.

	x				
	0	1	2	3	4
$p_1(x)$	0.2	0.2	0.3	0.1	0.1
$p_2(x)$	0.1	0.3	0.3	0.2	0.2
$p_3(x)$	0.1	0.2	0.4	0.2	0.1

- b) Para la función de masa de probabilidad posible, calcule μ_X y σ_X^2 .

5. Cierta tipo de componente está empaquetado en lotes de cuatro. Sea X el número de componentes que funcionan de modo adecuado en un lote elegido de manera aleatoria. Suponga que la probabilidad de que exactamente x componentes funcionen es proporcional a x ; en otras palabras, suponga que la función de masa de probabilidad de X es dada por

$$p(x) = \begin{cases} cx & x = 1, 2, 3, \text{ o } 4 \\ 0 & \text{de otro modo} \end{cases}$$

donde c es una constante.

- a) Determine el valor de la constante c para que $p(x)$ sea una función de masa de probabilidad.
- b) Determine $P(X = 2)$.
- c) Determine la media del número de componentes que funcionan adecuadamente.
- d) Determine la varianza del número de componentes que funcionan adecuadamente.
- e) Determine la desviación estándar del número de componentes que funcionan adecuadamente.
6. El producto de un proceso químico se supervisa continuamente para asegurar que la concentración permanezca dentro de los límites aceptables. Cada vez que la concentración se sale de los límites, el proceso se detiene y se calibra nuevamente. Sea X el número de veces en una semana dada que el proceso se calibra nuevamente. La tabla siguiente presenta los valores de la función de distribución acumulativa $F(x)$ de X .

x	$F(x)$
0	0.17
1	0.53
2	0.84
3	0.97
4	1.00

- ¿Cuál es la probabilidad de que el proceso se calibre nuevamente menos de dos veces durante una semana?
- ¿Cuál es la probabilidad de que el proceso se calibre nuevamente más de tres veces durante una semana?
- ¿Cuál es la probabilidad de que el proceso se calibre de nuevo exactamente una vez durante una semana?
- ¿Cuál es la probabilidad de que el proceso no se calibre nuevamente durante una semana?
- ¿Cuál es el número más probable de que ocurran nuevas calibraciones durante una semana?

7. En 100 días diferentes, un ingeniero especializado en el tránsito de automóviles cuenta el número de éstos que pasan por cierto cruceiro entre las 5:00 y 5:05 p.m. Los resultados se presentan en la tabla siguiente.

Número de automóviles	Número de días	Proporción de días
0	36	0.36
1	28	0.28
2	15	0.15
3	10	0.10
4	7	0.07
5	4	0.04

- Sea X el número de automóviles que pasan por el cruceiro entre las 5:00 y las 5:05 p.m. en un día elegido aleatoriamente. Alguien sugiere que para cualquier entero positivo x , la de masa de probabilidad de X es $p_1(x) = (0.2)(0.8)^x$. Usando esta función, calcule $P(X = x)$ para valores de x de 0 a 5 inclusive.
- Otra persona sugiere que para cualquier entero positivo x , la función de masa de probabilidad es $p_2(x) = (0.4)(0.6)^x$. Usando esta función, calcule $P(X = x)$ para valores de x de 0 a 5 inclusive.
- Compare los resultados de los incisos a) y b) con los datos de la tabla. ¿Cuál función de masa de probabilidad parece ser el mejor modelo? Explique.
- Alguien dice que ninguna de las funciones es un buen modelo ya que ninguna coincide exactamente con los datos. ¿Esto es correcto? Explique.

8. Se seleccionan aleatoriamente chips de microprocesadores uno tras otro de una gran población y se prueban para determinar si son aceptables para determinada aplicación. El 90% de los chips en la población es aceptable.

- ¿Cuál es la probabilidad de que el primer chip elegido sea aceptable?
 - ¿Cuál es la probabilidad de que el primer chip sea inaceptable y que el segundo sea aceptable?
 - Sea X el número de chips que se prueba hasta incluir el primer chip aceptable. Determine $P(X = 3)$.
 - Determine la función de masa de probabilidad de X .
9. Con referencia al ejercicio 8, sea Y el número de chips probados hasta incluir el segundo chip aceptable.

- ¿Cuál es el valor más pequeño posible de Y ?
- ¿Cuál es la probabilidad de que Y tome ese valor?
- Sea X el número de chips que se prueba hasta incluir el primer chip aceptable. Determine $P(Y = 3 | X = 1)$.
- Determine $P(Y = 3 | X = 2)$.
- Determine $P(Y = 3)$.

10. Se seleccionan aleatoriamente tres componentes, uno a la vez, de un gran lote. Cuando se selecciona cada componente, se prueba. Si pasa la prueba, ocurre un éxito (S); si no pasa la prueba, ocurre una falla (F). Suponga que 80% de los componentes del lote conseguirá pasar la prueba. Sea X el número de éxitos entre los tres componentes seleccionados.

- ¿Cuáles son los valores posibles para X ?
- Determine $P(X = 3)$.
- Al evento de que falle el primer componente y que los próximos dos pasen con éxito la prueba se denota por FSS. Determine $P(\text{FSS})$.
- Determine $P(\text{SFS})$ y $P(\text{SSF})$.
- Use los resultados de los incisos c) y d) para encontrar $P(X = 2)$.
- Determine $P(X = 1)$.
- Determine $P(X = 0)$.
- Determine μ_X .
- Determine σ_X^2 .
- Sea Y el número de éxitos si se seleccionan cuatro componentes. Determine $P(Y = 3)$.

11. La hidrogenación del benceno para el ciclohexano es promovida con un catalizador de níquel dividido en poros finos. El catalizador de partículas se puede considerar como esferas de diferentes tamaños. Todas las partículas tienen masas entre 10 y 70 μg . Sea X la masa de una partícula ele-

gida aleatoriamente. La función de densidad de probabilidad de X está dada por

$$f(x) = \begin{cases} \frac{x-10}{1800} & 10 < x < 70 \\ 0 & \text{de otro modo} \end{cases}$$

- ¿Cuál es la proporción de partículas que tiene una masa menor a $50 \mu\text{g}$?
- Determine la media de las masas de las partículas.
- Determine la desviación estándar de las masas de las partículas.
- Determine la función de distribución acumulativa de las masas de las partículas.
- Determine la mediana de las masas de las partículas.

12. Las especificaciones piden que el espesor de las hojas de aluminio para hacer latas estén entre 8 y 11 milésimos de pulgada. Sea X el espesor de una hoja de aluminio. Suponga que la función de densidad de probabilidad de X está dada por

$$f(x) = \begin{cases} \frac{x}{54} & 6 < x < 12 \\ 0 & \text{de otro modo} \end{cases}$$

- ¿Qué proporción de hojas cubrirán con la especificación?
- Determine la media del espesor de una hoja.
- Determine la varianza del espesor de una hoja.
- Determine la desviación estándar del espesor de una hoja.
- Determine la función de distribución acumulativa del espesor.
- Determine la mediana el espesor.
- Determine el décimo percentil del espesor.
- Una hoja especial tiene 10 milésimos de pulgada de espesor. ¿Qué proporción de hojas tiene más espesor?

13. Una masa radiactiva emite partículas de tiempo periódicamente. El tiempo entre dos emisiones es aleatorio. Sea T el tiempo en segundos entre dos emisiones. Suponga que la función de densidad de probabilidad de T es dada por

$$f(t) = \begin{cases} 0.2e^{-0.2t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

- Determine la media del tiempo entre emisiones.
- Determine la desviación estándar del tiempo entre emisiones.

- Determine la función de distribución acumulativa del tiempo entre emisiones.
- Determine la probabilidad de que el tiempo entre emisiones sea menor a diez segundos.
- Determine la mediana del tiempo entre emisiones.
- Determine el 90avo. percentil de los tiempos entre emisiones.

14. Un proceso que fabrica anillos de pistón produce anillos cuyos diámetros (en centímetros) varían de acuerdo con la función de densidad de probabilidad

$$f(x) = \begin{cases} 3[1 - 16(x - 10)^2] & 9.75 < x < 10.25 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine la media del diámetro de los anillos que se fabrican mediante este proceso.
- Determine la desviación estándar de los diámetros de anillos fabricados por este proceso. (Sugerencia: La ecuación 2.36 puede ser más fácil de usar que la ecuación 2.37.)
- Determine la función de distribución acumulativa de los diámetros de los anillos de pistón.
- ¿Qué proporción de anillos de pistón tiene diámetros menores a 9.75 centímetros?
- ¿Qué proporción de anillos de pistón tiene diámetros entre 9.75 y 10.25 centímetros?

15. Con referencia al ejercicio 14, un proceso competidor produce anillos cuyos diámetros (en centímetros) varían de acuerdo con la función de densidad de probabilidad

$$f(x) = \begin{cases} 15[1 - 25(x - 10.05)^2]/4 & 9.85 < x < 10.25 \\ 0 & \text{de otro modo} \end{cases}$$

Las especificaciones pedidas para el diámetro son 10.0 ± 0.1 cm. ¿Qué proceso es mejor, éste o el del ejercicio 14? Explique.

16. Las partículas son un componente muy importante de la contaminación atmosférica en muchas áreas. Es interesante estudiar los tamaños de las partículas contaminantes. Sea X el diámetro, en micrómetros, de una partícula elegida aleatoriamente. Suponga que en cierta área, la función de densidad de probabilidad de X es inversamente proporcional al volumen de la partícula; es decir, suponga que

$$f(x) = \begin{cases} \frac{c}{x^3} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

donde c es una constante.

- a) Determine el valor de c para que $f(x)$ sea una función de densidad de probabilidad.
 - b) Determine la media del diámetro de la partícula.
 - c) Determine la función de distribución acumulativa del diámetro de la partícula.
 - d) Determine la mediana del diámetro de la partícula.
 - e) El término PM_{10} se refiere a partículas con diámetros menores o iguales a $10 \mu m$. ¿Qué proporción de partículas contaminantes son PM_{10} ?
 - f) El término $PM_{2.5}$ se refiere a partículas con diámetros menores o iguales a $2.5 \mu m$. ¿Qué proporción de partículas contaminantes son $PM_{2.5}$?
 - g) ¿Qué proporción de partículas PM_{10} son $PM_{2.5}$?
17. Un científico ecologista está preocupado por la tasa a la que se absorbe cierta solución tóxica en la piel. Sea X el volumen en microlitros de la solución absorbida por 1 pulg² de piel en 1 min. Suponga que la función de densidad de probabilidad de X se aproxima bien por la función $f(x) = (2\sqrt{2\pi})^{-1}e^{-(x-10)^2/8}$, definida de $-\infty < x < \infty$.
- a) Determine la media del volumen absorbido en 1 min.
 - b) (Difícil.) Determine la desviación estándar del volumen absorbido en 1 min.

2.5 Funciones lineales de variables aleatorias

En la práctica con frecuencia se construyen nuevas variables aleatorias realizando operaciones aritméticas con otras variables aleatorias. Por ejemplo, se podría sumar una constante a una variable aleatoria, multiplicar una variable aleatoria por una constante, o sumar dos o más variables aleatorias. En esta sección se describe cómo calcular medias y varianzas de variables aleatorias construidas de esta manera y se presentan algunos ejemplos prácticos. La presentación en esta sección es intuitiva. Una presentación más rigurosa se ofrece en la sección 2.6. Quiénes deseen este tipo de presentación pueden cubrir la sección 2.6 además de, o en lugar de, esta sección.

Sumando una constante

Cuando se suma una constante a una variable aleatoria, la media se aumenta por el valor de la constante, pero la varianza y la desviación estándar son iguales. Por ejemplo, suponga que las varillas de acero producidas por cierta máquina tienen una media de la longitud de 5.0 pulgadas y una varianza ($\sigma^2 = 0.003$ pulg²). Cada varilla se fija a una base que tiene exactamente 1.0 pulg de longitud. La media de la longitud del montaje será de $5.0 + 1.0 = 6.0$ pulg. Debido a que cada longitud se ha aumentado en la misma cantidad, la dispersión en las longitudes no cambia, por lo que la varianza es la misma. Para poner esto en términos estadísticos, sea X la longitud de una varilla elegida aleatoriamente y sea $Y = X + 1$ la longitud del montaje. Entonces $\mu_Y = \mu_{X+1} = \mu_X + 1$, y $\sigma_Y^2 = \sigma_{X+1}^2 = \sigma_X^2$. En general, cuando se suma una constante a una variable aleatoria, la media se desplaza en esa constante y la varianza no cambia.

Resumen

Si X es una variable aleatoria y b es una constante, entonces

$$\mu_{X+b} = \mu_X + b \quad (2.39)$$

$$\sigma_{X+b}^2 = \sigma_X^2 \quad (2.40)$$

Multiplicando por una constante

Con frecuencia se tiene que multiplicar una variable aleatoria por una constante. Esto último se podría hacer, por ejemplo, al convertir a un sistema más conveniente de unidades. Se continúa con el ejemplo de la producción de varillas de acero para mostrar cómo se afecta la media, la varianza y la desviación estándar de una variable aleatoria por la multiplicación de una constante.

Si medimos las longitudes de las varillas descritas anteriormente en centímetros en vez de pulgadas, la media de longitud será $(2.54 \text{ cm/pulg})(5.0 \text{ pulg}) = 12.7 \text{ cm}$. En términos estadísticos, sea la variable aleatoria X la longitud en pulgadas de una varilla elegida aleatoriamente y sea $Y = 2.54X$ la longitud en centímetros. Entonces $\mu_Y = 2.54 \mu_X$. En general, cuando una variable aleatoria se multiplica por una constante, su media se multiplica por la misma constante.

Resumen

Si X es una variable aleatoria y a es una constante, entonces

$$\mu_{aX} = a\mu_X \quad (2.41)$$

Cuando la longitud X de una varilla se mide en pulgadas, la varianza σ_X^2 debe tener unidades de pulg^2 . Si $Y = 2.54X$ es la longitud en centímetros, entonces σ_Y^2 debe tener unidades de cm^2 . Por tanto, se obtiene σ_Y^2 al multiplicar σ_X^2 por 2.54^2 , que es el factor de conversión de pulg^2 a cm^2 . En general, cuando una variable aleatoria se multiplica por una constante, su varianza se multiplica por el cuadrado de la constante.

Resumen

Si X es una variable aleatoria y a es una constante, entonces

$$\sigma_{aX}^2 = a^2 \sigma_X^2 \quad (2.42)$$

$$\sigma_{aX} = |a| \sigma_X \quad (2.43)$$

Si una variable aleatoria se multiplica por una constante y después se suma a otra constante, el efecto sobre la media y la varianza se puede determinar al combinar las ecuaciones (2.39) y (2.41) y las ecuaciones (2.40) y (2.42). Los resultados se presentan en el resumen siguiente.

Resumen

Si X es una variable aleatoria y a y b son constantes, entonces

$$\mu_{aX+b} = a\mu_X + b \quad (2.44)$$

$$\sigma_{aX+b}^2 = a^2\sigma_X^2 \quad (2.45)$$

$$\sigma_{aX+b} = |a|\sigma_X \quad (2.46)$$

Observe que las ecuaciones (2.44) a (2.46) son similares a los resultados para la media y la desviación estándar de una muestra que se presentaron en la sección 1.2.

Ejemplo

2.45

La molaridad de un soluto en una solución se define como el número de moles de soluto por litro de solución ($1 \text{ mol} = 6.02 \times 10^{23}$ moléculas). Si la molaridad de una solución existente de ácido sulfúrico concentrado (H_2SO_4) es X y si una parte de la solución se mezcla con N partes de agua, la molaridad Y de la solución diluida está dada por $Y = X/(N + 1)$. Suponga que la solución existente se fabricó con un proceso que produce una molaridad con una media de 18 y con una desviación estándar de 0.1. Si a 100 ml de la solución existente se le agregan 300 ml de agua, determine la media y la desviación estándar de la molaridad de la solución diluida.

Solución

La molaridad de la solución diluida es $Y = 0.25X$. La media y la desviación estándar de X son $\mu_X = 18$ y $\sigma_X = 0.1$, respectivamente. Por tanto,

$$\begin{aligned} \mu_Y &= \mu_{0.25X} \\ &= 0.25\mu_X \quad (\text{utilizando la ecuación 2.41}) \\ &= 0.25(18.0) \\ &= 4.5 \end{aligned}$$

También,

$$\begin{aligned} \sigma_Y &= \sigma_{0.25X} \\ &= 0.25\sigma_X \quad (\text{utilizando la ecuación 2.43}) \\ &= 0.25(0.1) \\ &= 0.025 \end{aligned}$$

Medias de combinaciones lineales de variables aleatorias

Considere el caso de sumar dos variables aleatorias. Por ejemplo, suponga que hay dos máquinas que fabrican cierta pieza de metal. La media de la producción diaria de la máquina A

es de 100 piezas y la media de la producción diaria de la máquina B es de 150 piezas. Evidentemente la media de la producción diaria de las dos máquinas es de 250 piezas. Poniendo esto en notación matemática, sea X el número de piezas producidas en un día dado por la máquina A y sea Y el número de piezas producidas en el mismo día por la máquina B. El número total de piezas es $X + Y$ y se tiene que $\mu_{X+Y} = \mu_X + \mu_Y$.

Esta idea se generaliza para cualquier número de variables aleatorias.

Si $X_1, X_2, \dots, X_n$ son variables aleatorias, entonces la media de la suma $X_1 + X_2 + \dots + X_n$ está dada por

$$\mu_{X_1+X_2+\dots+X_n} = \mu_{X_1} + \mu_{X_2} + \dots + \mu_{X_n} \quad (2.47)$$

La suma $X_1 + X_2 + \dots + X_n$ es un caso especial de **combinación lineal**:

Si $X_1, \dots, X_n$ son variables aleatorias y $c_1, \dots, c_n$ son constantes, entonces la variable aleatoria

$$c_1X_1 + \dots + c_nX_n$$

se denomina **combinación lineal** de $X_1, \dots, X_n$.

Para encontrar la media de una combinación lineal de variables aleatorias, se puede combinar las ecuaciones (2.41) y (2.47):

Si X y Y son variables aleatorias y a y b son constantes, entonces

$$\mu_{aX+bY} = \mu_{aX} + \mu_{bY} = a\mu_X + b\mu_Y \quad (2.48)$$

De forma más general, si $X_1, X_2, \dots, X_n$ son variables aleatorias y $c_1, c_2, \dots, c_n$ son constantes, entonces la media de la combinación lineal $c_1X_1 + c_2X_2 + \dots + c_nX_n$ está dada por

$$\mu_{c_1X_1+c_2X_2+\dots+c_nX_n} = c_1\mu_{X_1} + c_2\mu_{X_2} + \dots + c_n\mu_{X_n} \quad (2.49)$$

Variables aleatorias independientes

El concepto de independencia para variables aleatorias es exactamente como el concepto de independencia para eventos. Dos variables aleatorias son independientes si se sabe que una de ellas no afecta las probabilidades de los demás. Cuando dos eventos son independientes, la probabilidad de que ambos ocurran se encuentra multiplicando las probabilidades para cada evento (véanse las ecuaciones 2.19 y 2.20 de la sección 2.3). Hay fórmulas similares para variables aleatorias independientes. La notación para estas fórmulas es la siguiente. Sea X una variable aleatoria y sea S un conjunto de números. La notación " $X \in S$ " significa que el valor de la variable aleatoria X está en el conjunto S .

Definición

Si X y Y son variables aleatorias **independientes** y S y T son conjuntos de números, entonces

$$P(X \in S \text{ y } Y \in T) = P(X \in S)P(Y \in T) \quad (2.50)$$

De forma más general, si $X_1, X_2, \dots, X_n$ son variables aleatorias independientes y $S_1, \dots, S_n$ son conjuntos, entonces

$$P(X_1 \in S_1 \text{ y } X_2 \in S_2 \text{ y } \dots \text{ y } X_n \in S_n) = P(X_1 \in S_1)P(X_2 \in S_2) \dots P(X_n \in S_n) \quad (2.51)$$

Ejemplo 2.46

Las bandejas de plástico rectangulares para un disco compacto (CD) tienen especificaciones de longitud y de ancho, sea X la longitud y Y el ancho, cada una se mide al milímetro más cercano, de una bandeja seleccionada aleatoriamente. La de masa de probabilidad de X está dada por $P(X = 129) = 0.2$, $P(X = 130) = 0.7$ y $P(X = 131) = 0.1$. La de masa de probabilidad de Y está dada por $P(Y = 15) = 0.6$ y $P(Y = 16) = 0.4$. El área de una bandeja está dada por $A = XY$. Suponga que X y Y son independientes. Determine la probabilidad de que el área sea 1 935 mm².

Solución

El área será igual a 1 935 si $X = 129$ y $Y = 15$. Por tanto,

$$\begin{aligned} P(A = 1\,935) &= P(X = 129 \text{ y } Y = 15) \\ &= P(X = 129)P(Y = 15) \quad \text{ya que } X \text{ y } Y \text{ son independientes} \\ &= (0.2)(0.6) \\ &= 0.12 \end{aligned}$$

Las ecuaciones (2.50) y (2.51) indican cómo calcular probabilidades para variables aleatorias independientes, pero no son generalmente de mucha ayuda para determinar si las variables aleatorias son en realidad independientes. En general, la mejor manera de determinar si las variables aleatorias son independientes es entendiendo el proceso que las generó.

Varianzas de combinaciones lineales de variables aleatorias independientes

Se ha visto que la media de una suma de variables aleatorias es siempre igual a la suma de las medias (ecuación 2.47). En general, la fórmula para la varianza de una suma de variables aleatorias es un poco más complicada que esto último. Pero cuando las variables aleatorias son *independientes*, el resultado es simple: la varianza de la suma es la suma de las varianzas.

Si $X_1, X_2, \dots, X_n$ son variables aleatorias *independientes*, entonces la varianza de la suma $X_1 + X_2 + \dots + X_n$ está dada por

$$\sigma_{X_1+X_2+\dots+X_n}^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \dots + \sigma_{X_n}^2 \quad (2.52)$$

Para encontrar la varianza de una combinación lineal de variables aleatorias, se puede combinar las ecuaciones (2.52) y (2.42):

Si $X_1, X_2, \dots, X_n$ son variables aleatorias *independientes* y $c_1, c_2, \dots, c_n$ son constantes, entonces la varianza de la combinación lineal $c_1X_1 + c_2X_2 + \dots + c_nX_n$ está dada por

$$\sigma_{c_1X_1+c_2X_2+\dots+c_nX_n}^2 = c_1^2\sigma_{X_1}^2 + c_2^2\sigma_{X_2}^2 + \dots + c_n^2\sigma_{X_n}^2 \quad (2.53)$$

Dos combinaciones lineales frecuentemente encontradas son la suma y la diferencia de dos variables aleatorias. Curiosamente, cuando las variables aleatorias son independientes, la varianza de la suma es igual a la varianza de la diferencia.

Si X y Y son variables aleatorias *independientes* con las varianzas σ_X^2 y σ_Y^2 , entonces la varianza de la suma $X + Y$ es

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 \quad (2.54)$$

La varianza de la diferencia $X - Y$ es

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 \quad (2.55)$$

El hecho de que la varianza de la diferencia es la *suma* de las varianzas podría parecer contradictorio. Sin embargo, se obtiene de la ecuación (2.53) al hacer $c_1 = 1$ y $c_2 = -1$.

Ejemplo

2.47

Un pistón se coloca dentro de un cilindro. La holgura es la distancia entre el borde del pistón y la pared del cilindro y es igual a la mitad de la diferencia entre el diámetro del cilindro y el diámetro del pistón. Suponga que el diámetro del pistón tiene una media de 80.85 cm con una desviación estándar de 0.02 cm. Suponga que el diámetro del cilindro tiene una media de 80.95 cm con una desviación estándar de 0.03 cm. Determine la media de la holgura. Suponiendo que el pistón y el cilindro se eligen independientemente, determine la desviación estándar de la holgura.

Solución

Sea X_1 el diámetro del cilindro y sea X_2 el diámetro del pistón. La holgura está dada por $C = 0.5X_1 - 0.5X_2$. Usando la ecuación (2.49), la media del perímetro es

$$\begin{aligned}
 \mu_C &= \mu_{0.5X_1 - 0.5X_2} \\
 &= 0.5\mu_{X_1} - 0.5\mu_{X_2} \\
 &= 0.5(80.95) - 0.5(80.85) \\
 &= 0.050
 \end{aligned}$$

Puesto que X_1 y X_2 son independientes, se puede usar la ecuación (2.53) para encontrar la desviación estándar σ_C :

$$\begin{aligned}
 \sigma_C &= \sqrt{\sigma_{0.5X_1 - 0.5X_2}^2} \\
 &= \sqrt{(0.5)^2\sigma_{X_1}^2 + (-0.5)^2\sigma_{X_2}^2} \\
 &= \sqrt{0.25(0.02)^2 + 0.25(0.03)^2} \\
 &= 0.018
 \end{aligned}$$

Independencia y muestras aleatorias simples

Cuando se toma una muestra aleatoria simple de valores numéricos de una población, a cada elemento de la muestra se le puede pensar como una variable aleatoria. Los elementos en una muestra aleatoria simple se pueden tratar como independientes, excepto cuando la muestra es una proporción grande (mayor a 5%) de una población finita (véase el análisis de independencia en la sección 1.1). De ahora en adelante, a menos que se indique explícitamente lo contrario, se supondrá que no ocurre esta excepción, por lo que los valores en una muestra aleatoria simple se pueden tratar como variables aleatorias independientes.

Resumen

Si $X_1, X_2, \dots, X_n$ es una muestra aleatoria simple, entonces $X_1, X_2, \dots, X_n$ se puede tratar como variables aleatorias independientes, todas con la misma distribución.

A veces cuando $X_1, \dots, X_n$ son variables aleatorias independientes, todas con la misma distribución, se dice que $X_1, \dots, X_n$ son **independientes e idénticamente distribuidas (i.i.d.)**.

Media y varianza de una media de la muestra

La combinación lineal frecuentemente encontrada es la media de muestra. Específicamente, si $X_1, \dots, X_n$ es una muestra aleatoria simple de una población con media μ y varianza σ^2 , entonces la media muestral $\bar{X}$ es la combinación lineal

$$\bar{X} = \frac{1}{n}X_1 + \dots + \frac{1}{n}X_n$$

A partir de lo anterior se puede calcular la media y la varianza de $\bar{X}$.

$$\begin{aligned}
 \mu_{\bar{X}} &= \mu_{\frac{1}{n}X_1 + \dots + \frac{1}{n}X_n} \\
 &= \frac{1}{n}\mu_{X_1} + \dots + \frac{1}{n}\mu_{X_n} \quad (\text{usando la ecuación 2.49})
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{n}\mu + \cdots + \frac{1}{n}\mu \\
 &= (n) \left(\frac{1}{n} \right) \mu \\
 &= \mu
 \end{aligned}$$

Como se analizó anteriormente, los elementos de una muestra aleatoria simple se pueden tratar como variables independientes aleatorias. Por tanto

$$\begin{aligned}
 \sigma_{\bar{X}}^2 &= \sigma_{\frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n}^2 \\
 &= \frac{1}{n^2}\sigma_{X_1}^2 + \cdots + \frac{1}{n^2}\sigma_{X_n}^2 \quad (\text{utilizando la ecuación 2.53}) \\
 &= \frac{1}{n^2}\sigma^2 + \cdots + \frac{1}{n^2}\sigma^2 \\
 &= (n) \left(\frac{1}{n^2} \right) \sigma^2 \\
 &= \frac{\sigma^2}{n}
 \end{aligned}$$

Resumen

Si $X_1, \dots, X_n$ es una muestra aleatoria simple de una población con media μ y varianza σ^2 , entonces la media muestral $\bar{X}$ es una variable aleatoria con

$$\mu_{\bar{X}} = \mu \quad (2.56)$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \quad (2.57)$$

La desviación estándar de $\bar{X}$ es

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad (2.58)$$

Ejemplo

2.48

Un proceso que llena botellas de plástico con una bebida tiene una media de volumen de llenado de 2.013 L y una desviación estándar de 0.005 L. Una caja contiene 24 botellas. Suponiendo que las botellas en una caja son una muestra aleatoria simple de botellas llenadas con este método, determine la media y la desviación estándar del volumen promedio por botella en una caja.

Solución

Sea que $V_1, \dots, V_{24}$ representen los volúmenes en las 24 botellas en una caja. Ésta es una muestra aleatoria simple de una población con media $\mu = 2.013$ y desviación estándar $\sigma = 0.005$. El volumen promedio es $\bar{V} = (V_1 + \cdots + V_{24})/24$. Utilizando la ecuación (2.56),

$$\mu_{\bar{V}} = \mu = 2.013$$

Usando la ecuación (2.58),

$$\sigma_{\bar{V}} = \frac{\sigma}{\sqrt{24}} = 0.001$$

Ejercicios para la sección 2.5

- Si X y Y son variables aleatorias independientes con medias $\mu_X = 10.5$ y $\mu_Y = 5.7$ y desviaciones estándares $\sigma_X = 0.5$ y $\sigma_Y = 0.3$, determine las medias y las desviaciones estándares de las cantidades siguientes:
 - $2X$
 - $X - Y$
 - $3X + 2Y$
- Si un resistor con resistencia R ohms lleva una corriente de I amperes, la diferencia de potencial a través del resistor, en volts, está dada por $V = IR$. La resistencia de un resistor elegido aleatoriamente que está etiquetado con 100Ω tiene una media de 100Ω y una desviación estándar de 10Ω . En un resistor seleccionado aleatoriamente se establece una corriente de 3 A.
 - Determine μ_V .
 - Determine σ_V .
- El tiempo de vida de cierto foco en una aplicación específica tiene una media de 700 horas y una desviación estándar de 20 horas. Cuando se funde cada foco, se reemplaza con un foco nuevo. Determine la media y la desviación estándar del tiempo que durarán cinco focos.
- Dos resistores, con resistencias R_1 y R_2 , se conectan en serie. La resistencia combinada R está dada por $R = R_1 + R_2$. Suponiendo que R_1 tiene una media de 50Ω y una desviación estándar de 5Ω y que R_2 tiene una media de 100Ω y una desviación estándar de 10Ω .
 - Determine μ_R .
 - Suponiendo que R_1 y R_2 son independientes, determine σ_R .
- Una pieza de madera contrachapada está compuesta de cinco capas. Las capas son una muestra aleatoria simple de una población cuyos espesores tienen una media de 0.125 pulg y una desviación estándar de 0.005 pulg.
 - Determine la media del espesor de una pieza de madera contrachapada.
 - Determine la desviación estándar del espesor de una pieza de madera contrachapada.
- La fabricación de cierto artículo requiere que dos máquinas funcionen consecutivamente. El tiempo en la primera máquina tiene una media de 10 min y desviación estándar de 2 min. El tiempo en la segunda máquina tiene una media de 15 min y una desviación estándar de 3 min. Suponga que los tiempos invertidos en las dos máquinas son independientes.
 - Determine la media del tiempo total invertido en las dos máquinas.
 - Determine la desviación estándar del tiempo total invertido en las dos máquinas.
- La molaridad de un soluto en una solución está definido como el número de moles de soluto por litro de solución ($1 \text{ mol} = 6.02 \times 10^{23}$ moléculas). Si X es la molaridad de una solución de cloruro de magnesio (MgCl_2) y Y es la molaridad de una solución de cloruro férrico (FeCl_3), la molaridad del ion de cloro (Cl^-) en una solución hecha a partes iguales de la solución, de MgCl_2 y FeCl_3 está dada por $M = X + 1.5Y$. Suponga que X tiene una media de 0.125 y una desviación estándar 0.05 y que Y tiene una media de 0.350 y una desviación estándar de 0.10 .
 - Determine μ_M .
 - Suponiendo que X y Y son independientes, determine σ_M .
- Una máquina que llena cajas de cartón con cereal tiene un peso de llenado cuya media es 12.02 oz, con una desviación estándar de 0.03 oz. Una caja consta de 12 cajas seleccionadas aleatoriamente del producto de la máquina.
 - Determine la media del peso total de cereal en la caja.
 - Determine la desviación estándar del peso total del cereal en la caja.
 - Determine la media del peso promedio por caja del cereal en la caja.
 - Determine la desviación estándar del peso promedio por caja del cereal en la caja.

- e) ¿Cuántas cajas se deben incluir en una caja para que la desviación estándar del peso promedio de la caja sea 0.005 oz?
9. Los cuatro lados de un marco de fotografía constan de dos piezas seleccionadas de una población cuya media de la longitud es de 30 cm con desviación estándar de 0.1 cm y dos piezas seleccionadas de una población cuya media de la longitud es de 45 cm con desviación estándar de 0.3 cm.
- a) Determine la media del perímetro.
- b) Suponga que las cuatro piezas se eligen independientemente. Determine la desviación estándar del perímetro.
10. Una estación de gasolina percibe \$1.60 en ingreso por cada galón de gasolina regular que vende, \$1.75 para cada galón de gasolina de medio grado y \$1.90 para cada galón de gasolina súper. Sea que X_1 , X_2 y X_3 indiquen los números de galones de gasolina regular, de medio grado y súper que se vendieron en un día. Suponga que X_1 , X_2 y X_3 tienen medias $\mu_1 = 1\,500$, $\mu_2 = 500$ y $\mu_3 = 300$ y desviaciones estándares $\sigma_1 = 180$, $\sigma_2 = 90$ y $\sigma_3 = 40$, respectivamente.
- a) Determine la media del ingreso diario.
- b) Suponiendo que X_1 , X_2 y X_3 son independientes. Determine la desviación estándar del ingreso diario.
11. En el artículo “An Investigation of the Ca–CO₃–CaF₂–K₂SiO₃–SiO₂–Fe Flux System Using the Submerged Arc Welding Process on HSLA-100 and AISI-1018 Steels” (G. Fredrickson, tesis de maestría, Escuela de Minas, de Colorado, 1992), el carbono equivalente p de una soldadura de metal se definió como una combinación lineal de los por-

centajes de los pesos de carbono (C), manganeso (Mn), cobre (Cu), cromo (Cr), silicio (Si), níquel (Ni), molibdeno (Mo), vanadio (V) y boro (B). El carbono equivalente está dado por

$$P = C + \frac{Mn + Cu + Cr}{20} + \frac{Si}{30} + \frac{Ni}{60} + \frac{Mo}{15} + \frac{V}{10} + 5B$$

Las medias y las desviaciones estándares de los porcentajes de los pesos de estos elementos químicos se calcularon de las mediciones en 45 soldaduras de metales producidas con el metal base de acero HSLA-100. Suponga que las medias y las desviaciones estándares (DE) son como se muestran en la tabla siguiente.

	Media	DE
C	0.0695	0.0018
Mn	1.0477	0.0269
Cu	0.8649	0.0225
Cr	0.7356	0.0113
Si	0.2171	0.0185
Ni	2.8146	0.0284
Mo	0.5913	0.0031
V	0.0079	0.0006
B	0.0006	0.0002

- a) Determine la media del equivalente de carbono de metales de soldadura fabricados con el metal base de acero HSLA-1000.
- b) Suponiendo que los porcentajes de los peso sean independientes, determine la desviación estándar del equivalente de carbono de metales de soldadura fabricados con el metal base de acero HSLA-1000.

2.6 Variables aleatorias con distribución conjunta*

En esta sección se presentan resultados de algunas variables aleatorias en un modo más riguroso que en la sección 2.5. Para las personas que deseen este tipo de presentación pueden cubrir esta sección además, o en lugar, de la sección 2.5.

Se ha dicho que observar un valor de una variable aleatoria es como seleccionar un valor de una población. En algunos casos, los elementos en la población podrían tener algunas variables aleatorias relacionadas con cada una de ellos. Por ejemplo, imagine elegir aleatoriamente un estudiante de una lista de todos los estudiantes inscritos en una universidad y tomar datos de su estatura y peso. A cada persona de la población de estudiantes le corresponden dos variables aleatorias, la altura y el peso. Si también se determinara la edad del estudiante,

* Esta sección es opcional.

a cada persona le corresponderían tres variables. En principio, cualquier número de variables aleatorias se puede relacionar con cada elemento en una población.

Cuando dos o más variables aleatorias están relacionadas con cada elemento de una población, se dice que aquéllas están con **distribución conjunta**. Si todas son discretas, se dice que son **conjuntamente discretas**. Si todas son continuas, entonces son **conjuntamente continuas**. Se estudiarán estos dos casos por separado.

Variables aleatorias conjuntamente discretas

En el ejemplo 2.46 (de la sección 2.5) se analizaron las longitudes y los anchos de la bandeja de plástico rectangular para un CD que está instalada en una computadora personal. Las mediciones se redondearon al milímetro más cercano. Sean X la longitud medida y Y el ancho medido. Los valores posibles de X son 129, 130 y 131 y los valores posibles para Y son 15 y 16. Tanto X como Y son discretas, por lo que X y Y son **conjuntamente discretas**. Hay seis valores posibles para el par ordenado (X, Y) : (129, 15), (129, 16), (130, 15), (130, 16), (131, 15) y (131, 16). Suponga que las probabilidades de cada uno de estos pares arreglados son como se muestra en la tabla siguiente.

x	y	$P(X = x \text{ y } Y = y)$
129	15	0.12
129	16	0.08
130	15	0.42
130	16	0.28
131	15	0.06
131	16	0.04

La **función de masa de probabilidad conjunta** representa la función $p(x, y) = P(X = x \text{ y } Y = y)$. Por lo que, por ejemplo, se tiene $p(129, 15) = 0.12$ y $p(130, 16) = 0.28$.

A veces se da una función de masa de probabilidad conjunta de dos variables aleatorias, pero se tiene interés en sólo una de ellas. Por ejemplo, se podría tener interés en la función de masa de probabilidad de X , la longitud de la bandeja del CD, pero no se tiene interés en el ancho Y . Se puede obtener la función de masa de probabilidad de una de las variables X o Y por separado sumando los valores adecuados de la función de masa de probabilidad conjunta. Los ejemplos 2.49 y 2.50 ilustran el método.

Ejemplo 2.49

Determine la probabilidad de que la cubierta del CD tenga una longitud de 129 mm.

Solución

Está claro de la tabla anterior que 12% de las cubiertas del CD en la población tiene una longitud de 129 y un ancho de 15 y que 8% tiene longitud de 129 y ancho de 16.

Por tanto, 20% de los elementos en la población tiene longitud de 129. La probabilidad de que una cubierta de CD tenga longitud de 129 mm es de 0.20. Simbólicamente, se tiene que

$$\begin{aligned} P(X = 129) &= P(X = 129 \text{ y } Y = 15) + P(X = 129 \text{ y } Y = 16) \\ &= 0.12 + 0.08 \\ &= 0.20 \end{aligned}$$

Ejemplo
2.50

Determine la probabilidad de que una cubierta de CD tenga ancho de 16 mm.

Solución

Se necesita encontrar $P(Y = 16)$. Se puede encontrar esta cantidad sumando las probabilidades de todos los pares (x, y) para los cuales $y = 16$. Se obtiene que

$$\begin{aligned} P(Y = 16) &= P(X = 129 \text{ y } Y = 16) + P(X = 130 \text{ y } Y = 16) \\ &\quad + P(X = 131 \text{ y } Y = 16) \\ &= 0.08 + 0.28 + 0.04 \\ &= 0.40 \end{aligned}$$

Los ejemplos 2.49 y 2.50 muestran que se puede encontrar la función de masa de probabilidad de X (o Y) sumando la función de masa de probabilidad conjunta sobre todos los valores de Y (o X). La tabla 2.3 presenta la función de masa de probabilidad conjunta de X y Y . La función de masa de probabilidad de X se presenta en la columna más a la derecha y se obtiene sumando a lo largo de los renglones. La función de masa de probabilidad de Y se presenta en el renglón inferior y se obtiene sumando abajo las columnas. Observe que las funciones de masa de probabilidad de X y de Y se muestran en los márgenes de la tabla. Por esta razón con frecuencia son llamadas funciones de masa de probabilidad **marginal**.

TABLA 2.3 Funciones de masa de probabilidad conjunta y marginal para la longitud X y el ancho Y de una cubierta de CD

x	y		$p_x(x)$
	15	16	
129	0.12	0.08	0.20
130	0.42	0.28	0.70
131	0.06	0.04	0.10
$p_y(y)$	0.60	0.40	

Por último, si se suman las funciones de densidad de probabilidad conjunta sobre todos valores posibles de x y y , se obtiene la probabilidad de que X y Y tomen valores en algún lugar dentro de sus rangos posibles y esta probabilidad es igual a 1.

Resumen

Si X y Y son variables aleatorias conjuntamente discretas:

- La función de masa de probabilidad conjunta de X y Y es la función

$$p(x,y) = P(X = x \text{ y } Y = y)$$

- Las funciones de masa de probabilidad marginal de X y Y se pueden obtener a partir de la función de masa de probabilidad conjunta como se muestra a continuación:

$$p_X(x) = P(X = x) = \sum_y p(x,y) \quad p_Y(y) = P(Y = y) = \sum_x p(x,y)$$

donde las sumatorias se toman sobre todos los valores posibles de Y y de X , respectivamente.

- La función de masa de probabilidad conjunta tiene la propiedad de que

$$\sum_x \sum_y p(x,y) = 1$$

donde la suma incluye a todos los valores posibles de X y Y .

Variables aleatorias conjuntamente continuas

Se ha visto que si X es una variable aleatoria continua, se pueden encontrar sus probabilidades al integrar su función de densidad de probabilidad. Se dice que las variables aleatorias X y Y son **conjuntamente continuas** si se determinan sus probabilidades al integrar una función con dos variables, denominada **función de densidad de probabilidad conjunta** de X y Y . Para determinar la probabilidad de que X y Y tengan ciertos valores en una región, integramos la función de densidad de probabilidad conjunta en esa región tal y como lo muestra el ejemplo 2.51.

Ejemplo

2.51

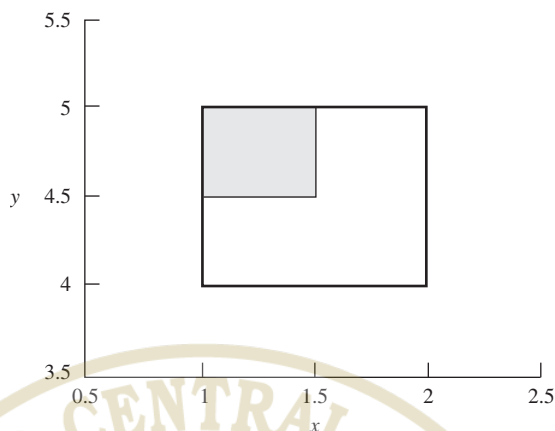
Suponga que en cierto tipo de lavadora, tanto el espesor como el diámetro de la cavidad son diferentes en cada unidad. X representa el espesor en milímetros y Y denota el diámetro de la cavidad en milímetros, de una lavadora seleccionada aleatoriamente. Suponga que la función de densidad de la probabilidad conjunta de X y Y está dada por

$$f(x,y) = \begin{cases} \frac{1}{6}(x+y) & \text{si } 1 \leq x \leq 2 \text{ y } 4 \leq y \leq 5 \\ 0 & \text{de otro modo} \end{cases}$$

Determine la probabilidad de que una lavadora escogida aleatoriamente tenga un espesor entre 1.0 y 1.5 mm y una cavidad con un diámetro entre 4.5 y 5 mm.

Solución

Se necesita encontrar $P(1 \leq X \leq 1.5 \text{ y } 4.5 \leq Y \leq 5)$. El rectángulo grande en la figura muestra la región en la cual la densidad conjunta es positiva. El rectángulo sombreado indica la región en la que $1 \leq x \leq 1.5$ y $4.5 \leq y \leq 5$, en la cual se integrará a la densidad conjunta.



Se integra a la función de densidad de probabilidad conjunta en la región indicada:

$$\begin{aligned}
 P(1 \leq X \leq 1.5 \text{ y } 4.5 \leq Y \leq 5) &= \int_1^{1.5} \int_{4.5}^5 \frac{1}{6}(x+y) dy dx \\
 &= \int_1^{1.5} \left\{ \frac{xy}{6} + \frac{y^2}{12} \right\} \bigg|_{y=4.5}^{y=5} dx \\
 &= \int_1^{1.5} \left(\frac{x}{12} + \frac{19}{48} \right) dx \\
 &= \frac{1}{4}
 \end{aligned}$$

Observe que si se integrara una función de densidad de probabilidad conjunta en todo el plano, es decir, si los límites fueran de $-\infty$ a ∞ tanto para x como para y , se obtendría la probabilidad de que X y Y tuvieran valores entre $-\infty$ y ∞ , lo que es igual a 1.

Resumen

Si X y Y son variables aleatorias conjuntamente continuas, con una función de densidad de probabilidad conjunta $f(x, y)$ y $a < b$, $c < d$, entonces

$$P(a \leq X \leq b \text{ y } c \leq Y \leq d) = \int_a^b \int_c^d f(x, y) dy dx$$

La función de densidad de probabilidad conjunta tiene la propiedad de que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1$$

Se ha visto que si X y Y son variables conjuntamente discretas, la función de masa de probabilidad de cualquier variable se puede encontrar al sumar la función de masa de probabilidad conjunta en todos los valores de la otra variable. Cuando se calcula de esta forma, a la función de masa de probabilidad se le llama función de masa de probabilidad marginal. Por analogía, si X y Y son variables conjuntamente continuas, la función de densidad de probabilidad de cualquier variable se puede determinar al integrar la función de densidad de probabilidad conjunta con respecto a la otra variable. Cuando se calcula de esta manera, a la función de densidad de probabilidad se le denomina **función de densidad de probabilidad marginal**. El ejemplo 2.52 ilustra esta idea.

Ejemplo

2.52

Con referencia al ejemplo 2.51, encuentre la función de densidad de probabilidad marginal del espesor X de una lavadora. Determine la función de densidad de probabilidad marginal del diámetro de la cavidad Y de una lavadora.

Solución

Denote a la función de densidad de probabilidad marginal de X por medio de $f_X(x)$ y a la función de densidad de probabilidad marginal de Y por $f_Y(y)$. Entonces

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
 &= \int_4^5 \frac{1}{6}(x + y) dy \\
 &= \frac{1}{6} \left(x + \frac{9}{2} \right) \quad \text{para } 1 \leq x \leq 2
 \end{aligned}$$

y

$$\begin{aligned}
 f_Y(y) &= \int_{-\infty}^{\infty} f(x, y) dx \\
 &= \int_1^2 \frac{1}{6}(x + y) dx \\
 &= \frac{1}{6} \left(y + \frac{3}{2} \right) \quad \text{para } 4 \leq y \leq 5
 \end{aligned}$$

Resumen

Si X y Y son variables conjuntamente continuas con una función de densidad de probabilidad conjunta $f(x, y)$, entonces las funciones de densidad de probabilidad marginal de X y Y están dadas, respectivamente, por

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Ejemplo

2.53

El artículo “Performance Comparison of Two Location Based Routing Protocols for ad Hoc Networks” (T. Camp, J. Boleng y colaboradores, en *Proceedings of the Twenty-first annual Joint Conference of IEEE Computer and Communications Societies*, 2002:1678-1687) describe un modelo del movimiento de una computadora móvil. Suponga que ésta se mueve dentro de la región A delimitada por el eje x , la recta $x = 1$ y la recta $y = x$ de tal forma que si (X, Y) representa la posición de la computadora en determinado momento, la función de densidad conjunta de X y Y está dada por

$$f(x, y) = \begin{cases} 8xy & (x, y) \in A \\ 0 & (x, y) \notin A \end{cases}$$

Encuentre a $P(X > 0.5 \text{ y } Y < 0.5)$.

Solución

La región A es el triángulo que se muestra en la figura 2.13, donde la región $X > 0.5$ y $Y < 0.5$ aparece sombreada. Para determinar a $P(X > 0.5 \text{ y } Y < 0.5)$, se integra a la densidad conjunta en la región sombreada.

$$\begin{aligned} P(X > 0.5 \text{ y } Y < 0.5) &= \int_{0.5}^1 \int_0^{0.5} 8xy \, dy \, dx \\ &= \int_{0.5}^1 \left\{ 4xy^2 \right|_{y=0}^{y=0.5} \right\} dx \\ &= \int_{0.5}^1 x \, dx \\ &= 0.375 \end{aligned}$$

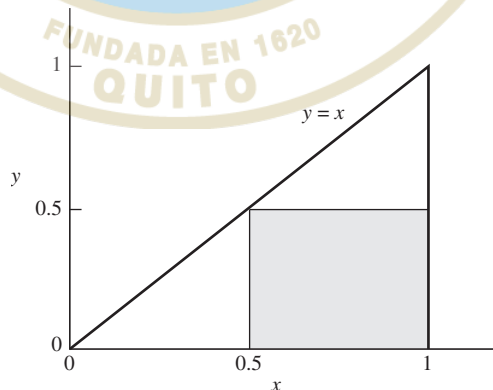


FIGURA 2.13 El triángulo representa la región en la cual la densidad conjunta de X y Y es positiva. Al integrar la densidad conjunta en el cuadrado sombreado, se encuentra la probabilidad de que el punto (X, Y) se encuentre en el cuadrado sombreado.

Ejemplo

2.54

Con referencia al ejemplo 2.53, encuentre las densidades marginales de X y de Y .

Solución

Para calcular a $f_X(x)$, la densidad marginal de X , se mantiene constante a x y se integra a la densidad conjunta a lo largo de la recta vertical que pasa por x , como se muestra en la figura 2.14. La integración se realiza con respecto a y , y los límites de integración son $y = 0$ a $y = x$.

$$\begin{aligned} f_X(x) &= \int_0^x 8xy \, dy \\ &= 4xy^2 \Big|_{y=0}^{y=x} \\ &= 4x^3 \quad \text{para } 0 < x < 1 \end{aligned}$$

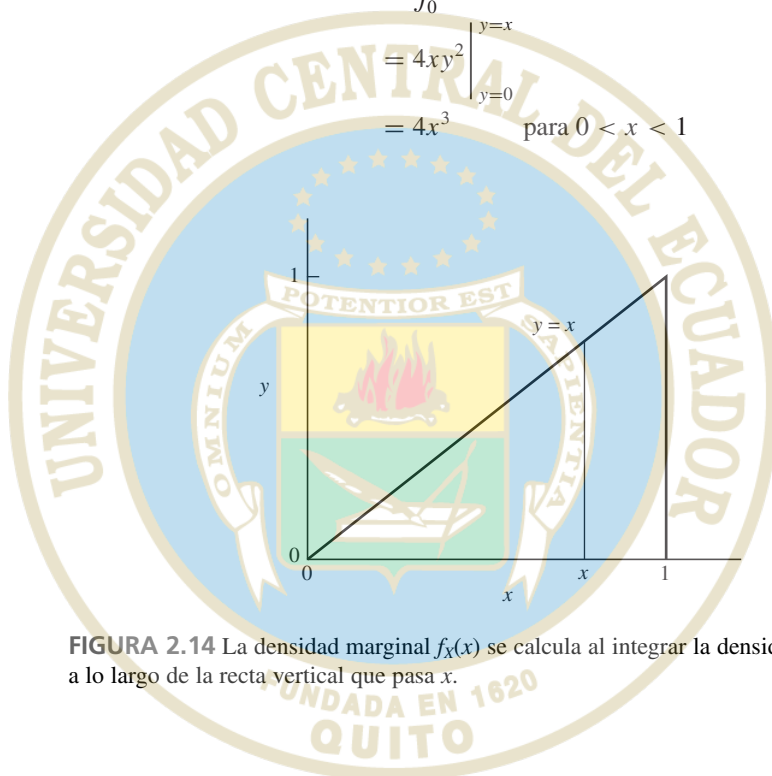


FIGURA 2.14 La densidad marginal $f_X(x)$ se calcula al integrar la densidad conjunta a lo largo de la recta vertical que pasa x .

Para calcular $f_Y(y)$, la densidad marginal de Y se mantiene constante a y y se integra la densidad conjunta a lo largo de la recta horizontal que pasa por y , como se muestra la figura 2.15. La integración es con respecto a x y los límites de integración van de $x = y$ a $x = 1$.

$$\begin{aligned} f_Y(y) &= \int_y^1 8xy \, dx \\ &= 4x^2y \Big|_{x=y}^{x=1} \\ &= 4y - 4y^3 \quad \text{para } 0 < y < 1 \end{aligned}$$

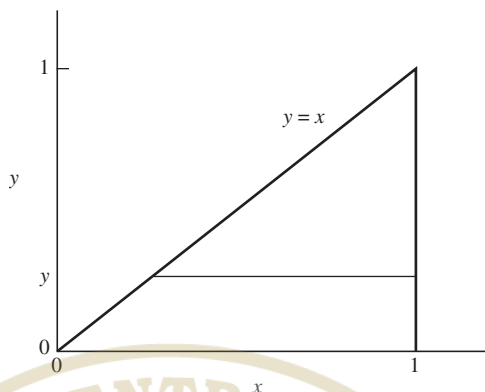


FIGURA 2.15 La densidad marginal $f_y(y)$ se calcula al integrar la densidad conjunta a lo largo de la recta horizontal que pasa por y .

Más de dos variables aleatorias

Los conceptos de funciones de masa de probabilidad conjunta y funciones de densidad de probabilidad conjunta se amplían fácilmente a más de dos variables aleatorias. Aquí se presentan las definiciones.

Definición

- Si las variables aleatorias $X_1, \dots, X_n$ son conjuntamente discretas, la función de masa de probabilidad es

$$p(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)$$

- Si las variables aleatorias $X_1, \dots, X_n$ son conjuntamente continuas, tienen una función de densidad de probabilidad conjunta $f(x_1, \dots, x_n)$, donde

$$P(a_1 \leq X_1 \leq b_1, \dots, a_n \leq X_n \leq b_n) = \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

para cualesquiera constantes $a_1 \leq b_1, \dots, a_n \leq b_n$.

Medias de funciones de variables aleatorias

Algunas veces se da una variable aleatoria X y se necesita trabajar con una función de X . Si X es una variable aleatoria y $h(X)$ es una función de X , entonces $h(X)$ es también una variable aleatoria. Si se desea calcular la media de $h(X)$, se puede hacer utilizando la función de masa de probabilidad o la función de densidad de probabilidad de X . No es necesario conocer la función de masa de probabilidad o la función de densidad de probabilidad de $h(X)$.

Sea X una variable aleatoria y sea $h(X)$ una función de X , entonces

- Si X es discreta con una función de masa de probabilidad $p(x)$, la media de $h(X)$ está dada por

$$\mu_{h(X)} = \sum_x h(x)p(x) \quad (2.59)$$

donde la suma incluye a todos los valores posibles de X .

- Si X es continua con una función de densidad de probabilidad $f(x)$, la media de $h(X)$ está dada por

$$\mu_{h(X)} = \int_{-\infty}^{\infty} h(x)f(x) dx \quad (2.60)$$

Observe que si se sustituye $h(X) = (X - \mu_X)^2$ ya sea en la ecuación (2.59) o en la (2.60), la parte derecha de la ecuación se convierte en una expresión de la varianza de X . Por lo que $\sigma_X^2 = \mu_{(X - \mu_X)^2}$. Se puede obtener otra expresión de la varianza de X al sustituir $h(X) = X^2$ y restar μ_X^2 en ambos lados de la ecuación. Se concluye que $\sigma_X^2 = \mu_{X^2} - \mu_X^2$.

Ejemplo

2.55

Un motor de combustión interna contiene varios cilindros calibrados en el bloque del motor. X representa el diámetro del calibre de un cilindro, en milímetros. Suponga que la función de densidad de probabilidad de X es

$$f(x) = \begin{cases} 10 & 80.5 < x < 80.6 \\ 0 & \text{de otro modo} \end{cases}$$

$A = \pi X^2/4$ representa el área del calibre. Encuentre la media de A .

Solución

$$\begin{aligned} \mu_A &= \int_{-\infty}^{\infty} \frac{\pi x^2}{4} f(x) dx \\ &= \int_{80.5}^{80.6} \frac{\pi x^2}{4} (10) dx \\ &= 5\,096 \end{aligned}$$

La media del área es de 5 096 mm².

Si $h(X) = aX + b$ es una función lineal de X , entonces la media μ_{aX+b} y la varianza σ_{aX+b}^2 se pueden expresar en términos de μ_X y σ_X^2 . Estos resultados se presentaron en la sección 2.5 de la ecuación (2.44) a la (2.46); aquí se repiten nuevamente.

Si X es una variable aleatoria, y a y b son constantes, entonces

$$\mu_{aX+b} = a\mu_X + b \quad (2.61)$$

$$\sigma_{aX+b}^2 = a^2\sigma_X^2 \quad (2.62)$$

$$\sigma_{aX+b} = |a|\sigma_X \quad (2.63)$$

Las demostraciones de estos resultados se encuentran al final de esta sección.

Si X y Y son variables aleatorias con distribución conjunta y $h(X,Y)$ es una función de X y Y , entonces se puede calcular la media de $h(X,Y)$ a partir de la función de masa de probabilidad conjunta o la función de densidad de probabilidad conjunta de X y Y .

Si X y Y son variables aleatorias con distribución conjunta y $h(X,Y)$ es una función de X y Y , entonces

- Si X y Y son conjuntamente discretas con una función de masa de probabilidad conjunta $p(x,y)$

$$\mu_{h(X,Y)} = \sum_x \sum_y h(x,y) p(x,y) \quad (2.64)$$

donde la sumatoria se realiza sobre todos los valores posibles de X y Y .

- Si X y Y son conjuntamente continuas con una función de densidad de probabilidad conjunta $f(x,y)$

$$\mu_{h(X,Y)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) f(x,y) dx dy \quad (2.65)$$

Ejemplo

2.56

El desplazamiento de un pistón en un motor de combustión interna se define como el volumen que la parte superior del pistón mueve desde la parte superior hasta el fondo de su golpe. X representa el diámetro del calibre del cilindro y Y representa la longitud del golpe del pistón en milímetros. El desplazamiento está dado por $D = \pi X^2 Y / 4$. Suponga que X y Y se distribuyen conjuntamente con una función de masa de probabilidad conjunta

$$f(x,y) = \begin{cases} 100 & 80.5 < x < 80.6 \text{ y } 65.1 < y < 65.2 \\ 0 & \text{de otro modo} \end{cases}$$

Determine la media de D .

Solución

$$\mu_D = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\pi x^2 y}{4} f(x,y) dx dy$$

$$\begin{aligned}
 &= \int_{65.1}^{65.2} \int_{80.5}^{80.6} \frac{\pi x^2 y}{4} (100) dx dy \\
 &= 331\,998
 \end{aligned}$$

La media del desplazamiento es de $331\,998 \text{ mm}^3$, o $\approx 332 \text{ mL}$.

Distribuciones condicionales

Si X y Y son variables aleatorias con distribución conjunta, entonces conocer el valor de X puede cambiar las probabilidades de la variable aleatoria Y . Por ejemplo, X representa la altura en pulgadas y Y representa el peso en libras de un estudiante elegido aleatoriamente. Se tiene interés en la probabilidad $P(Y \geq 200)$. Si se conoce la densidad conjunta de X y Y , se puede determinar esta probabilidad al calcular la densidad marginal de Y . Ahora se ha aprendido que la estatura del estudiante es $X = 78$. Evidentemente, saber lo anterior cambia la probabilidad de que $Y \geq 200$. Para calcular esta nueva probabilidad, se necesita el concepto de distribución **condicional**.

Primero, se analiza el caso en el que X y Y son variables conjuntamente discretas. Sea x cualquier valor para el cual $P(X = x) > 0$. Entonces, la probabilidad condicional de que $Y = y$ dado $X = x$ es $P(Y = y | X = x)$. Se expresa esta probabilidad condicional en términos de las funciones de masa de probabilidad conjunta y marginal. Si $p(x, y)$ representa la función de masa de probabilidad conjunta de X y Y , y $p_X(x)$ denota la función de masa de probabilidad marginal de X , entonces la probabilidad condicional

$$P(Y = y | X = x) = \frac{P(X = x \text{ y } Y = y)}{P(X = x)} = \frac{p(x, y)}{p_X(x)}$$

La **función de masa de probabilidad condicional** de Y dado $X = x$ es la probabilidad condicional $P(Y = y | X = x)$, considerada como una función de y y x .

Definición

Sean X y Y variables aleatorias conjuntamente discretas, con una función de masa de probabilidad conjunta $p(x, y)$. Sea $p_X(x)$ la de masa de probabilidad marginal de X y sea x cualquier número para el cual $p_X(x) > 0$.

La función de masa de probabilidad condicional de Y dado $X = x$ es

$$p_{Y|X}(y | x) = \frac{p(x, y)}{p_X(x)} \quad (2.66)$$

Observe que para cualesquiera valores particulares de x y y , el valor de $p_{Y|X}(y|x)$ es, de hecho, la probabilidad condicional $P(Y = y | X = x)$.

Ejemplo

2.57

La tabla 2.3 presenta la función de masa de probabilidad conjunta de la longitud X y ancho Y de una cubierta de CD. Calcule la función de masa de probabilidad condicional $p_{Y|X}(y | 130)$.

Solución

Los posibles valores de Y son $y = 15$ y $y = 16$. De la tabla 2.3, $P(Y = 15 \text{ y } X = 130) = 0.42$, y $P(X = 130) = 0.70$. Por tanto,

$$\begin{aligned} p_{Y|X}(15 | 130) &= P(Y = 15 | X = 130) \\ &= \frac{P(Y = 15 \text{ y } X = 130)}{P(X = 130)} \\ &= \frac{0.42}{0.70} \\ &= 0.60 \end{aligned}$$

El valor de $p_{Y|X}(16 | 130)$ se puede calcular con un cálculo similar. Alternativamente observe que $p_{Y|X}(16 | 130) = 1 - p_{Y|X}(15 | 130)$, puesto que $y = 15$ y $y = 16$ son los dos únicos valores posibles de Y . Por tanto, $p_{Y|X}(16 | 130) = 0.4$. La función de masa de probabilidad condicional de Y dado $X = 130$ es, por consiguiente, $p_{Y|X}(15 | 130) = 0.60$, $p_{Y|X}(16 | 130) = 0.40$, y $p_{Y|X}(y | 130) = 0$ para cualquier valor de y diferente a 15 o 16.

La analogía de la función de masa de probabilidad condicional para variables aleatorias conjuntamente continuas es la **función de densidad de probabilidad condicional**. La definición de la función de densidad de probabilidad condicional es igual a la de la función de masa de probabilidad condicional, sólo que las funciones de masa se reemplazan por funciones de densidad.

Definición

Sean X y Y variables aleatorias conjuntamente continuas, con una función de densidad de probabilidad conjunta $f(x, y)$. Sea $f_X(x)$ la función de densidad de probabilidad marginal de X y x cualquier número para el que $f_X(x) > 0$.

La función de densidad de probabilidad condicional de Y dado $X = x$ es

$$f_{Y|X}(y | x) = \frac{f(x, y)}{f_X(x)} \quad (2.67)$$

Ejemplo

2.58

(Continuación del ejemplo 2.51.) La función de densidad de probabilidad conjunta del espesor X y del diámetro de la cavidad Y (ambos en milímetros) de una lavadora elegida aleatoriamente es $f(x, y) = (1/6)(x + y)$ para $1 \leq x \leq 2$ y $4 \leq y \leq 5$. Encuentre la función de densidad de probabilidad condicional de Y dado que $X = 1.2$. Determine la probabilidad de que el diámetro de la cavidad sea menor o igual a 4.8 mm dado que el espesor mida 1.2 mm.

Solución

En el ejemplo 2.52 se calculan las funciones de densidad de probabilidad marginal

$$f_X(x) = \frac{1}{6}(x + 4.5) \quad \text{para } 1 \leq x \leq 2 \qquad f_Y(y) = \frac{1}{6}(y + 1.5) \quad \text{para } 4 \leq y \leq 5$$

La función de densidad de probabilidad condicional de Y dado que $X = 1.2$ es

$$\begin{aligned} f_{Y|X}(y | 1.2) &= \frac{f(1.2, y)}{f_X(1.2)} \\ &= \begin{cases} \frac{(1/6)(1.2 + y)}{(1/6)(1.2 + 4.5)} & \text{si } 4 \leq y \leq 5 \\ 0 & \text{de otro modo} \end{cases} \\ &= \begin{cases} \frac{1.2 + y}{5.7} & \text{si } 4 \leq y \leq 5 \\ 0 & \text{de otro modo} \end{cases} \end{aligned}$$

La probabilidad de que el diámetro de la cavidad sea menor o igual a 4.8 mm dado que el espesor mide 1.2 mm es $P(Y \leq 4.8 | X = 1.2)$. Esto se determina al integrar $f_{Y|X}(y | 1.2)$ en la región $y \leq 4.8$:

$$\begin{aligned} P(Y \leq 4.8 | X = 1.2) &= \int_{-\infty}^{4.8} f_{Y|X}(y | 1.2) dy \\ &= \int_4^{4.8} \frac{1.2 + y}{5.7} dy \\ &= 0.786 \end{aligned}$$

Esperanza condicional

La esperanza es otro término para la media. Una **esperanza condicional** es una esperanza, o media, que se calcula utilizando una función de masa de probabilidad condicional o función de densidad de probabilidad condicional. La esperanza condicional de Y dado $X = x$ se representa por $E(Y | X = x)$ o $\mu_{Y|X}$. Lo anterior se ilustra en los ejemplos 2.59 a 2.61.

Ejemplo

2.59

La tabla 2.3 presenta la función de masa de probabilidad conjunta de la longitud X y ancho Y de una cubierta de CD. Calcule la esperanza condicional $E(Y | X = 130)$.

Solución

En el ejemplo 2.57 se calculó la función de masa de probabilidad condicional $p_{Y|X}(y | 130)$. La esperanza condicional $E(Y | X = 130)$ se calcula utilizando la definición de la media de una variable aleatoria discreta y la función de masa de probabilidad condicional. Específicamente,

$$\begin{aligned}
 E(Y | X = 130) &= \sum_y y p_{Y|X}(y | 130) \\
 &= 15 p_{Y|X}(15 | 130) + 16 p_{Y|X}(16 | 130) \\
 &= 15(0.60) + 16(0.40) \\
 &= 15.4
 \end{aligned}$$

Ejemplo 2.60

Con referencia al ejemplo 2.58, determine la esperanza condicional de Y dado que $X = 1.2$.

Solución

Ya que X y Y son conjuntamente continuas, utilizamos la definición de la media de una variable aleatoria continua para calcular la esperanza condicional.

$$\begin{aligned}
 E(Y | X = 1.2) &= \int_{-\infty}^{\infty} y f_{Y|X}(y | 1.2) dy \\
 &= \int_4^5 y \frac{1.2 + y}{5.7} dy \\
 &= 4.5146
 \end{aligned}$$

Ejemplo 2.61

Con referencia al ejemplo 2.58, determine el valor de μ_Y (a la que se puede nombrar como media incondicional de Y). ¿Difiere ésta de $E(Y | X = 1.2)$?

Solución

El valor de μ_Y se calcula por medio de la función de masa de probabilidad marginal de Y . Por tanto,

$$\begin{aligned}
 \mu_Y &= \int_{-\infty}^{\infty} y f_Y(y) dy \\
 &= \int_4^5 y \frac{1}{6} (y + 1.5) dy \\
 &= 4.5139
 \end{aligned}$$

En este caso, la esperanza condicional difiere ligeramente de la esperanza incondicional.

Variables aleatorias independientes

El concepto de independencia de variables aleatorias es muy parecido al de independencia de eventos. Dos variables aleatorias son independientes si el conocimiento de una de ellas no afecta las probabilidades de la otra. Aquí se presenta un concepto de independencia de variables aleatorias en términos de su función de masa de probabilidad conjunta o de su función de densidad de probabilidad conjunta. Una definición diferente, pero lógicamente equivalente, se presentó en la sección 2.5.

Definición

Dos variables aleatorias X y Y son independientes, si se satisface que

- Si X y Y son conjuntamente discretas, la función de masa de probabilidad conjunta es igual al producto de las probabilidades marginales:

$$p(x, y) = p_X(x)p_Y(y)$$

- Si X y Y son conjuntamente continuas, la función de densidad de probabilidad conjunta es igual al producto de las probabilidades marginales:

$$f(x, y) = f_X(x)f_Y(y)$$

Las variables aleatorias $X_1, \dots, X_n$ son independientes, si se satisface que

- Si $X_1, \dots, X_n$ son conjuntamente discretas, la función de masa de probabilidad conjunta es igual al producto de las probabilidades marginales:

$$p(x_1, \dots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n)$$

- Si $X_1, \dots, X_n$ son conjuntamente continuas, la función de densidad de probabilidad conjunta es igual al producto de las probabilidades marginales:

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

Intuitivamente, cuando dos variables aleatorias son independientes, el conocer el valor de una de ellas no afecta a la distribución de probabilidad de la otra. En otras palabras, la distribución condicional de Y dado X es la misma que la distribución marginal de Y .

Si X y Y son variables aleatorias independientes, entonces

- Si X y Y son conjuntamente discretas y x es un valor para el cual $p_X(x) > 0$, entonces

$$p_{Y|X}(y|x) = p_Y(y)$$

- Si X y Y son conjuntamente continuas y x es un valor para el cual $f_X(x) > 0$, entonces

$$f_{Y|X}(y|x) = f_Y(y)$$

Ejemplo

2.62

La función de masa de probabilidad conjunta de la longitud X y el ancho Y de una bandeja de CD se encuentran en la tabla 2.3. ¿ X y Y son independientes?

Solución

Se debe comprobar si $P(X = x \text{ y } Y = y) = P(X = x)P(Y = y)$ en cada valor de x y y . Se empieza por verificar si $x = 129$, y $y = 15$:

$$P(X = 129 \text{ y } Y = 15) = 0.12 = (0.20)(0.60) = P(X = 129)P(Y = 15)$$

Si se continúa de esta forma, se puede comprobar que $P(X = x \text{ y } Y = y) = P(X = x)P(Y = y)$ para cada valor de x y y . Por consecuencia, X y Y son independientes.

Ejemplo

2.63

(Continuación del ejemplo 2.51.) La función de densidad de probabilidad conjunta del espesor X y del diámetro de la cavidad Y de una lavadora elegida aleatoriamente es $f(x,y) = (1/6)(x + y)$ para $1 \leq x \leq 2$ y $4 \leq y \leq 5$. ¿ X y Y son independientes?

Solución

En el ejemplo 2.52, se calculan las funciones de masa de probabilidad marginal

$$f_X(x) = \frac{1}{6} \left(x + \frac{9}{2} \right) \quad f_Y(y) = \frac{1}{6} \left(y + \frac{3}{2} \right)$$

Evidentemente $f(x,y) \neq f_X(x)f_Y(y)$. Por tanto, X y Y no son independientes.

Covarianza

Cuando dos variables aleatorias no son independientes, es útil tener una medida de la intensidad de la relación entre ellas. La **covarianza poblacional** representa una medida de un cierto tipo de relación conocida como **relación lineal**. Usualmente se quita el término “poblacional” y se le llama simplemente covarianza.

Definición

Sean X y Y variables aleatorias con medias μ_X y μ_Y . La covarianza de X y Y es

$$\text{Cov}(X,Y) = \mu_{(X-\mu_X)(Y-\mu_Y)} \quad (2.68)$$

Una fórmula alternativa es

$$\text{Cov}(X,Y) = \mu_{XY} - \mu_X \mu_Y \quad (2.69)$$

La demostración de la equivalencia de estas dos fórmulas se encuentra al final de la sección. Es importante observar que las unidades de $\text{Cov}(X,Y)$ son las unidades de X multiplicadas por las unidades de Y .

¿Cómo mide la covarianza la intensidad de la relación lineal entre X y Y ? La covarianza es la media del producto de las desviaciones $(X - \mu_X)(Y - \mu_Y)$. Si se construyera un sistema cartesiano con el origen en (μ_X, μ_Y) , este producto sería positivo en el primero y tercer cuadrantes y negativo en el segundo y cuarto cuadrantes (véase la figura 2.16). De ahí que si $\text{Cov}(X,Y)$ es positiva, entonces los valores de (X,Y) en el primero y tercer cuadrantes serán más frecuentes que los valores en el segundo y cuarto cuadrantes. Por consiguiente, en una muestra aleatoria de puntos, los grandes valores de X tienden a hacer pareja con valores grandes de Y , mientras que los valores pequeños de X se equiparan con valores pequeños de Y (véase la figura 2.16a). De forma similar, si $\text{Cov}(X,Y)$ es negativa, es más probable que los

puntos en una muestra aleatoria se encuentren en el segundo y cuarto cuadrantes, de tal forma que los valores grandes de X tenderán a hacer pareja con los valores pequeños de Y (véase la figura 2.16b). Por último, si $\text{Cov}(X,Y)$ se aproxima a 0, hay poca tendencia a que valores grandes de X se equiparen con valores pequeños o grandes de Y (véase la figura 2.16c).

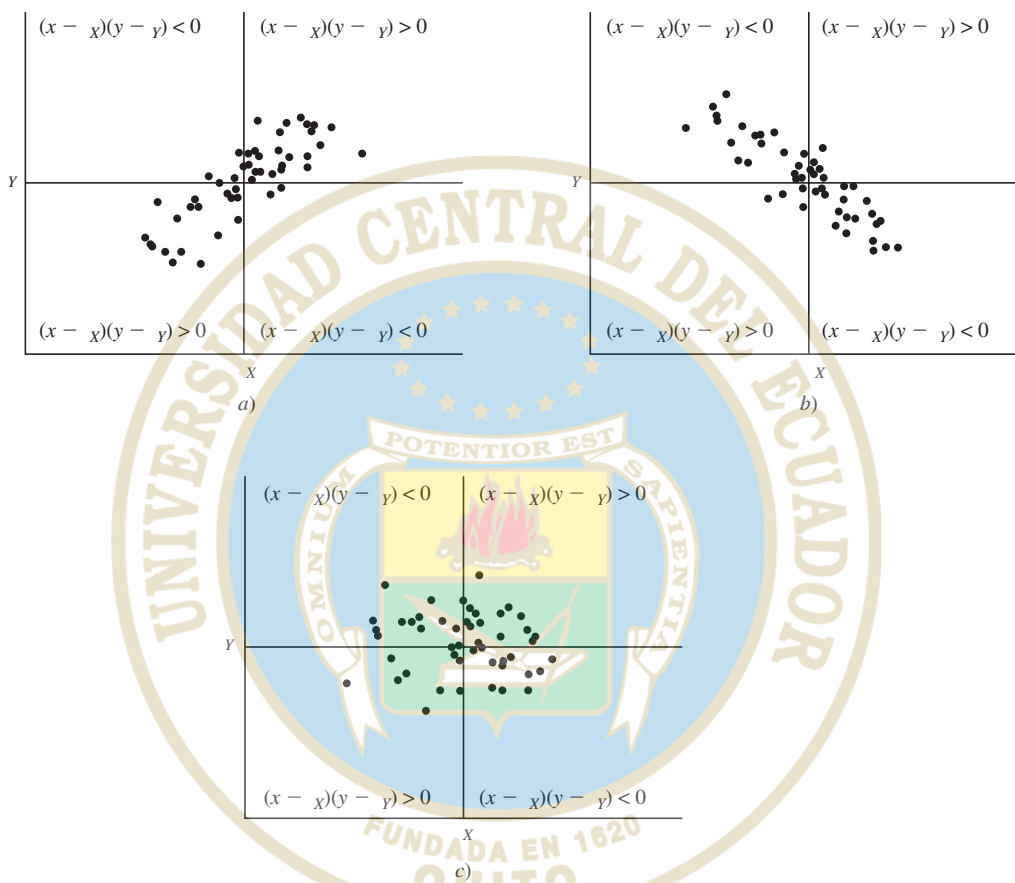


FIGURA 2.16 a) Muestra aleatoria de puntos de una población con covarianza positiva. b) Muestra aleatoria de puntos de una población con covarianza negativa. c) Muestra aleatoria de puntos de una población con covarianza cercana a 0.

Ejemplo

2.64

Continuando con el ejemplo 2.53, una computadora móvil se mueve en la región A delimitada por el eje x , la recta $x = 1$ y la recta $y = x$ (véase la figura 2.13). Si (X,Y) denota la posición de la computadora en cierto momento, la densidad conjunta de X y Y está dada por

$$f(x,y) = \begin{cases} 8xy & (x,y) \in A \\ 0 & (x,y) \notin A \end{cases}$$

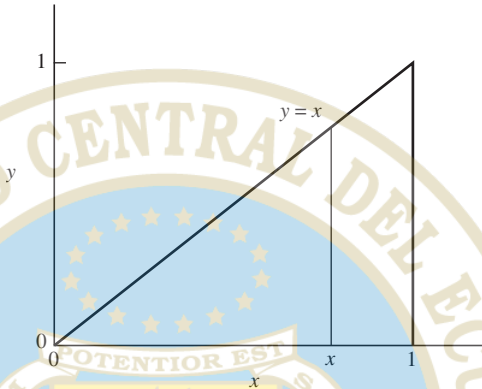
Determine a $\text{Cov}(X,Y)$.

Solución

Se utilizará la fórmula $\text{Cov}(X,Y) = \mu_{XY} - \mu_X\mu_Y$ (ecuación 2.69). Primero se calcula μ_{XY} :

$$\mu_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y) dy dx$$

Ahora la densidad conjunta es positiva en el triángulo que se muestra



Para calcular la integral en esta región, se fija un valor de x , como se muestra a continuación. Se calcula la integral interior integrando con respecto a y a lo largo de la recta vertical que pasa por x . Los límites de integración a lo largo de esta recta son de $y = 0$ a $y = x$. Posteriormente se calcula la integral exterior integrando con respecto a x todos los valores posibles de x , de manera que los límites de integración en la integral exterior son de $x = 0$ a $x = 1$.

Por tanto,

$$\begin{aligned} \mu_{XY} &= \int_0^1 \int_0^x xy(8xy) dy dx \\ &= \int_0^1 \left(\int_0^x 8x^2 y^2 dy \right) dx \\ &= \int_0^1 \frac{8x^5}{3} dx \\ &= \frac{4}{9} \end{aligned}$$

Para encontrar a μ_X y μ_Y se utilizan las densidades marginales calculadas en el ejemplo 2.54. Éstas son

$$\begin{aligned} f_X(x) &= \begin{cases} 4x^3 & 0 < x < 1 \\ 0 & \text{de otro modo} \end{cases} \\ f_Y(y) &= \begin{cases} 4y - 4y^3 & 0 < y < 1 \\ 0 & \text{de otro modo} \end{cases} \end{aligned}$$

Ahora se calculará a μ_X y μ_Y :

$$\begin{aligned}\mu_X &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_0^1 4x^4 dx \\ &= \frac{4}{5}\end{aligned}$$

$$\begin{aligned}\mu_Y &= \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= \int_0^1 (4y^2 - 4y^4) dy \\ &= \frac{8}{15}\end{aligned}$$

$$\text{Ahora, } \text{Cov}(X, Y) = \frac{4}{9} - \left(\frac{4}{5}\right) \left(\frac{8}{15}\right) = \frac{4}{225} = 0.01778.$$

Ejemplo 2.65

Las inspecciones de control de calidad de entrepaños de madera consisten en contar el número de imperfecciones en la superficie de cada entrepaño. En uno de éstos, de 2×8 pies, X es el número de imperfecciones en la superficie provocadas por una aplicación desigual de la última capa del material de acabado y Y representa el número de imperfecciones en la superficie debidas a la inclusión de partículas externas en el acabado. La función de masa de probabilidad conjunta $p(x, y)$ de X y Y se presenta en la siguiente tabla. Las funciones de masa de probabilidad marginal están también en los márgenes de la tabla. Determine la covarianza de X y Y .

		y			
		0	1	2	$p_X(x)$
x	0	0.05	0.10	0.20	0.35
	1	0.05	0.15	0.05	0.25
	2	0.25	0.10	0.05	0.40
$p_Y(y)$		0.35	0.35	0.30	

Solución

Se utilizará la fórmula $\text{Cov}(X, Y) = \mu_{XY} - \mu_X \mu_Y$ (ecuación 2.69). Primero se calculará μ_{XY} .

$$\begin{aligned}\mu_{XY} &= \sum_{x=0}^2 \sum_{y=0}^2 xy p(x, y) \\ &= (1)(1)(0.15) + (1)(2)(0.05) + (2)(1)(0.10) + (2)(2)(0.05) \\ &= 0.65 \quad (\text{eliminando los términos igual a } 0)\end{aligned}$$

Se utilizan las marginales para calcular μ_X y μ_Y :

$$\mu_X = (0)(0.35) + (1)(0.25) + (2)(0.40) = 1.05$$

$$\mu_Y = (0)(0.35) + (1)(0.35) + (2)(0.30) = 0.95$$

Por lo que $\text{Cov}(X,Y) = 0.65 - (1.05)(0.95) = -0.3475$.

Correlación

Si X y Y son variables aleatorias conjuntamente distribuidas, la $\text{Cov}(X,Y)$ mide la intensidad de la relación lineal entre ellas. Como se mencionó anteriormente, la covarianza tiene unidades, que son las unidades del producto de las unidades de X por las unidades de Y . Esto es una seria desventaja en la práctica, ya que no se puede utilizar la covarianza para determinar cuál de los dos pares de variables aleatorias está más relacionado, dado que las dos covarianzas tienen diferentes unidades. Lo que se necesita es una medida de la intensidad de una relación lineal que sea sólo un número. La **correlación poblacional** representa dicha medida. Usualmente no se emplea el término “poblacional” y se llama simplemente correlación. Representaremos la correlación entre las variables X y Y mediante $\rho_{X,Y}$.

La correlación es una versión ajustada de la covarianza. Específicamente, para calcular la correlación entre X y Y , primero se debe calcular la covarianza para después deshacerse de las unidades al dividir entre el producto de la desviación estándar de X y Y . Se puede comprobar mediante métodos avanzados que la correlación se encuentra siempre entre -1 y 1 .

Resumen

Sean X y Y variables aleatorias conjuntamente distribuidas con desviaciones estándar σ_X y σ_Y . La correlación entre X y Y se denota como $\rho_{X,Y}$ y está dada por

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} \quad (2.70)$$

Para cualesquiera dos variables aleatorias X y Y :

$$-1 \leq \rho_{X,Y} \leq 1$$

Ejemplo

2.66

Con referencia al ejemplo 2.64, encuentre $\rho_{X,Y}$.

Solución

En el ejemplo 2.64 se calcula que $\text{Cov}(X,Y) = 0.01778$, $\mu_X = 4/5$ y $\mu_Y = 8/15$. Ahora se debe calcular a σ_X y σ_Y . Con este propósito se utilizan las densidades marginales de X y Y , que se determinaron en el ejemplo 2.54. Éstas son: $f_X(x) = 4x^3$ para $0 < x < 1$ y $f_Y(y) = 4y - 4y^3$ para $0 < y < 1$. Se obtiene

$$\begin{aligned}
 \sigma_X^2 &= \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2 \\
 &= \int_0^1 4x^5 dx - \left(\frac{4}{5}\right)^2 \\
 &= 0.02667 \\
 \sigma_Y^2 &= \int_{-\infty}^{\infty} y^2 f_Y(y) dy - \mu_Y^2 \\
 &= \int_0^1 (4y^3 - 4y^5) dy - \left(\frac{8}{15}\right)^2 \\
 &= 0.04889
 \end{aligned}$$

Por lo que $\rho_{X,Y} = \frac{0.01778}{\sqrt{(0.02667)(0.04889)}} = 0.492$.

Ejemplo

2.67

Con referencia al ejemplo 2.65, determine $\rho_{X,Y}$.

Solución

En el ejemplo 2.65 se calcula que $\text{Cov}(X,Y) = -0.3475$, $\mu_X = 1.05$ y $\mu_Y = 0.95$. Ahora se debe calcular σ_X y σ_Y . Con ese fin se utilizan las densidades marginales de X y Y , que se encuentran en la tabla del ejemplo 2.65. Se obtiene

$$\begin{aligned}
 \sigma_X^2 &= \sum_{x=0}^2 x^2 p_X(x) - \mu_X^2 \\
 &= (0^2)(0.35) + (1^2)(0.25) + (2^2)(0.40) - 1.05^2 \\
 &= 0.7475
 \end{aligned}$$

$$\begin{aligned}
 \sigma_Y^2 &= \sum_{y=0}^2 y^2 p_Y(y) - \mu_Y^2 \\
 &= (0^2)(0.35) + (1^2)(0.35) + (2^2)(0.30) - 0.95^2 \\
 &= 0.6475
 \end{aligned}$$

Por lo que

$$\rho_{X,Y} = \frac{-0.3475}{\sqrt{(0.7475)(0.6475)}} = -0.499$$

Covarianza, correlación e independencia

Cuando $\text{Cov}(X,Y) = \rho_{X,Y} = 0$ no hay una relación lineal entre X y Y . En este caso se dice que X y Y **no están correlacionadas**. Advierta que si $\text{Cov}(X,Y) = 0$, entonces éste siempre será el caso cuando $\rho_{X,Y} = 0$ y viceversa. Si X y Y son variables aleatorias independientes, enton-

ces X y Y nunca estarán correlacionadas, ya que no existe ninguna relación lineal o de otro tipo entre ellos. Matemáticamente es posible construir variables aleatorias que no estén correlacionadas ni sean independientes. Sin embargo, este fenómeno se ve pocas veces en la práctica.

Resumen

- Si $\text{Cov}(X, Y) = \rho_{X, Y} = 0$, entonces se dice que X y Y no están correlacionadas.
- Si X y Y son independientes, entonces X y Y no están correlacionadas.
- Matemáticamente es posible que X y Y no estén correlacionadas ni sean independientes. Esto es poco común en la práctica.

Al final de esta sección se encuentra una demostración del hecho de que las variables aleatorias independientes nunca están correlacionadas. El ejercicio 26 muestra un ejemplo de variables aleatorias no correlacionadas ni independientes.

Combinaciones lineales de variables aleatorias

En la sección 2.5 se analizó las combinaciones lineales de variables aleatorias. Aquí vemos de nuevo los resultados e incluimos resultados adicionales de la varianza de una combinación lineal de variables aleatorias dependientes.

Si $X_1, \dots, X_n$ son variables aleatorias y $c_1, \dots, c_n$ son constantes, entonces a la variable aleatoria

$$c_1X_1 + \dots + c_nX_n$$

se le denomina **combinación lineal** de $X_1, \dots, X_n$.

Si $X_1, \dots, X_n$ son variables aleatorias y $c_1, \dots, c_n$ son constantes, entonces

$$\mu_{c_1X_1 + \dots + c_nX_n} = c_1\mu_{X_1} + \dots + c_n\mu_{X_n} \quad (2.71)$$

$$\sigma_{c_1X_1 + \dots + c_nX_n}^2 = c_1^2\sigma_{X_1}^2 + \dots + c_n^2\sigma_{X_n}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_i c_j \text{Cov}(X_i, X_j) \quad (2.72)$$

Al final de esta sección se encuentran las demostraciones de estos resultados para el caso de $n = 2$. La ecuación (2.72) es el resultado más general referente a la varianza de una combinación lineal de variables aleatorias. Como un caso especial, observe que si $X_1, \dots, X_n$ son independientes, entonces todas las covarianzas son iguales a 0, por lo que el resultado se simplifica en:

Si $X_1, \dots, X_n$ son variables aleatorias *independientes* y $c_1, \dots, c_n$ son constantes, entonces

$$\sigma_{c_1 X_1 + \dots + c_n X_n}^2 = c_1^2 \sigma_{X_1}^2 + \dots + c_n^2 \sigma_{X_n}^2 \quad (2.73)$$

En particular,

$$\sigma_{X_1 + \dots + X_n}^2 = \sigma_{X_1}^2 + \dots + \sigma_{X_n}^2 \quad (2.74)$$

Finalmente, se presentan algunos casos especiales de las ecuaciones (2.72) y (2.74) en las cuales sólo hay dos variables aleatorias:

Si X y Y son variables aleatorias, entonces

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2 \text{Cov}(X, Y) \quad (2.75)$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2 \text{Cov}(X, Y) \quad (2.76)$$

Si X y Y son variables aleatorias *independientes*, entonces

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 \quad (2.77)$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 \quad (2.78)$$

Observe que la varianza de la diferencia $X - Y$ de dos variables aleatorias independientes es la *suma* de las varianzas.

Ejemplo

2.68

(Continuación del ejemplo 2.53.) Suponga que la computadora móvil se mueve verticalmente desde una posición aleatoria (X, Y) al punto $(X, 0)$ y posteriormente a lo largo del eje x hacia el origen. Encuentre la media y la varianza de la distancia recorrida.

Solución

La distancia recorrida es la suma $X + Y$. Las medias de X y Y se calcularon en el ejemplo 2.64. Éstas son $\mu_X = 4/5 = 0.800$ y $\mu_Y = 8/15 = 0.533$. Se calcula

$$\begin{aligned} \mu_{X+Y} &= \mu_X + \mu_Y \\ &= 0.800 + 0.533 \\ &= 1.333 \end{aligned}$$

Para calcular σ_{X+Y}^2 se emplea la ecuación (2.75). En el ejemplo 2.64 se calculó que la $\text{Cov}(X, Y) = 0.01778$. En el ejemplo 2.66 se calculó $\sigma_X^2 = 0.02667$ y $\sigma_Y^2 = 0.04889$. Por consecuencia,

$$\begin{aligned}
 \sigma_{X+Y}^2 &= \sigma_X^2 + \sigma_Y^2 + 2 \operatorname{Cov}(X, Y) \\
 &= 0.02667 + 0.04889 + 2(0.01778) \\
 &= 0.1111
 \end{aligned}$$

Media y varianza de una media muestral

Se repasarán los procedimientos para calcular la media y la varianza de una media muestral que se presentaron en la sección 2.5. Cuando de una población se extrae una muestra aleatoria simple de valores numéricos, se puede considerar que cada unidad en la muestra constituye una variable aleatoria. A menos que la muestra represente una proporción grande (mayor que 5%) de la población, se puede tratar a las unidades en la muestra como independientes (véase en la sección 1.1 el análisis acerca de independencia). A partir de ahora, a menos que se indique lo contrario, se supondrá que los valores de una muestra aleatoria simple son variables aleatorias independientes

Si $X_1, \dots, X_n$ es una muestra aleatoria simple, entonces se puede considerar que $X_1, \dots, X_n$ son variables aleatorias independientes, todas con la misma distribución.

La combinación lineal más frecuente es la media muestral. Específicamente, si $X_1, \dots, X_n$ es una muestra aleatoria simple, entonces $X_1, \dots, X_n$ son independientes y la media muestral $\bar{X}$ es la combinación lineal

$$\bar{X} = \frac{1}{n}X_1 + \dots + \frac{1}{n}X_n$$

Por tanto, las fórmulas de la media y varianza de $\bar{X}$ se deducen de las ecuaciones (2.71) y (2.73), respectivamente, al establecer que $c_1 = c_2 = \dots = c_n = 1/n$.

Si $X_1, \dots, X_n$ es una muestra aleatoria simple de una población con media μ y varianza σ^2 , entonces la media muestral $\bar{X}$ es una variable aleatoria con

$$\mu_{\bar{X}} = \mu \quad (2.79)$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \quad (2.80)$$

La desviación estándar de $\bar{X}$ es

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad (2.81)$$

Ejemplo 2.69

El artículo “Water Price Influence on Apartment Complex Water Use” (D. Aghte y R. Billings, en *Journal of Water Resources Planning and Management*, 2002:366-369) analiza el volumen de agua consumido por los departamentos de 308 complejos habitacionales en Tucson, Arizona.

El volumen gastado por departamento durante el verano tuvo una media de 20.4 m^3 y una desviación estándar de 11.1 m^3 . Encuentre la media y la desviación estándar de la media muestral del uso de agua en una muestra de 100 departamentos. ¿Cuántos de éstos se deben incluir en la muestra para que la media muestral del uso de agua tenga una desviación estándar igual a 0.5 m^3 ?

Solución

Sean $X_1, \dots, X_{100}$ las cantidades de agua utilizadas en una muestra de 100 departamentos. Por tanto, $X_1, \dots, X_{100}$ provienen de una población con una media $\mu = 20.4$ y una desviación estándar $\sigma = 11.1$. Se concluye que la media muestral $\bar{X}$ tiene una media $\mu_{\bar{X}} = \mu = 20.4$ y una desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{100} = 1.11$. Sea n el tamaño de la muestra requerido para que $\sigma_{\bar{X}} = 0.5$. Entonces $\sigma/\sqrt{n} = 11.1/\sqrt{n} = 0.5$. Al despejar n , se obtiene que $n \approx 493$.

Aplicación al manejo de carteras de inversiones

La ecuación (2.72) y sus variantes son determinantes en las finanzas. Suponga que una persona tiene un número fijo de dólares para invertir. Ella puede elegir entre una variedad de inversiones, por ejemplo, acciones, bonos y bienes raíces. Después de un año ella venderá su inversión; X representa su ganancia (o pérdida). Como no se puede predecir con certeza el valor de X , los economistas lo tratan como una variable aleatoria. La media μ_X indica la cantidad que se espera que la inversión genere en promedio. La desviación estándar σ_X refleja la *volatilidad*, o *riesgo*, de la inversión. Si σ_X es muy pequeña, entonces es casi seguro que la inversión generará una cantidad cercana a la ganancia promedio μ_X , por lo que el riesgo es bajo. Si σ_X es grande, la ganancia puede variar ampliamente, por lo que el riesgo es alto. En general, si dos inversiones tienen la misma ganancia promedio, se prefiere a la que tenga la menor desviación estándar, puesto que en promedio genera la misma ganancia con un riesgo menor.

Ejemplo 2.70

Un individuo tiene \$200 para invertir. Él invertirá \$100 en cada una de dos inversiones. X y Y denotan las ganancias de las dos inversiones. Suponga que $\mu_X = \mu_Y = \$5$, $\sigma_X = \sigma_Y = \$2$, y $\rho_{X,Y} = 0.5$. Encuentre la media y la desviación estándar de la ganancia total de las dos inversiones.

Solución

La ganancia total es $X + Y$. La media es

$$\begin{aligned}\mu_{X+Y} &= \mu_X + \mu_Y \\ &= \$5 + \$5 \\ &= \$10\end{aligned}$$

Utilizando la ecuación (2.75), la desviación estándar es $\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2 + 2 \text{Cov}(X,Y)}$. Ahora la $\text{Cov}(X,Y) = \rho_{X,Y}\sigma_X\sigma_Y = (0.5)(2)(2) = 2$. Por consiguiente,

$$\begin{aligned}\sigma_{X+Y} &= \sqrt{2^2 + 2^2 + 2(2)} \\ &= \$3.46\end{aligned}$$

Es prudente comparar el resultado del ejemplo 2.70 con el resultado que se obtendría si los \$200 se hubieran aplicado en una sola inversión. El ejemplo 2.71 analiza esta posibilidad.

Ejemplo 2.71

Si el individuo del ejemplo 2.70 invirtió los \$200 en una de las dos inversiones, encuentre la media y la desviación estándar de las ganancias.

Solución

Suponga que ese tipo elige un plan de inversión, cuya ganancia de \$100 es X (el resultado sería el mismo si se eligiera Y). Como consecuencia de que se invierten \$200, en vez de \$100, la ganancia será $2X$. La media de la ganancia es

$$\mu_{2X} = 2\mu_X = 2(5) = \$10$$

La desviación estándar es

$$\sigma_{2X} = 2\sigma_X = 2(2) = \$4$$

Al comparar los resultados de los ejemplos 2.70 y 2.71 se demuestra que las medias de las ganancias de las dos estrategias de inversión son iguales, pero la desviación estándar (es decir, el riesgo) es menor cuando el capital se divide entre dos inversiones. Éste es el principio de *diversificación*. Cuando están disponibles dos inversiones, cuyas ganancias tienen la misma media y el mismo riesgo, siempre es ventajoso dividir el capital entre ambas, en vez de invertirlo en sólo una de ellas.

Demostración de que $\mu_{aX+b} = a\mu_X + b$

Se supone que X es una variable aleatoria continua con función de densidad $f(x)$. Entonces

$$\begin{aligned}\mu_{aX+b} &= \int_{-\infty}^{\infty} (ax + b)f(x) dx \quad (\text{ecuación 2.60}) \\ &= \int_{-\infty}^{\infty} ax f(x) dx + \int_{-\infty}^{\infty} b f(x) dx \\ &= a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx \\ &= a\mu_X + b(1) \\ &= a\mu_X + b\end{aligned}$$

La demostración en el caso que X sea una variable aleatoria discreta es similar, sólo que las integrales son sustituidas por sumatorias.

Demostración de que $\mu_{aX+bY} = a\mu_X + b\mu_Y$

Sean X y Y variables conjuntamente continuas con una densidad conjunta $f(x,y)$ y densidades marginales $f_X(x)$ y $f_Y(y)$. Sean a y b constantes, entonces

$$\begin{aligned}
\mu_{aX+bY} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ax + by) f(x, y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ax f(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} by f(x, y) dx dy \\
&= a \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dy dx + b \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy \\
&= a \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx + b \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x, y) dx \right] dy \\
&= a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= a\mu_X + b\mu_Y
\end{aligned}$$

La demostración en el caso de que X y Y sean variables conjuntamente discretas es similar, sólo que las integrales son sustituidas por sumatorias.

Demostración de que $\sigma_{aX+b}^2 = a^2 \sigma_X^2$

Se utiliza la notación $E(X)$ alternativamente con μ_X , $E(Y)$ alternativamente con μ_Y y así sucesivamente. Sea $Y = aX + b$, entonces

$$\begin{aligned}
\sigma_{aX+b}^2 &= \sigma_Y^2 \\
&= E(Y^2) - \mu_Y^2 \\
&= E[(aX + b)^2] - \mu_{aX+b}^2 \\
&= E(a^2 X^2 + 2abX + b^2) - (a\mu_X + b)^2 \\
&= E(a^2 X^2) + E(2abX) + E(b^2) - (a\mu_X + b)^2 \\
&= a^2 E(X^2) + 2abE(X) + b^2 - a^2 \mu_X^2 - 2ab\mu_X - b^2 \\
&= a^2 [E(X^2) - \mu_X^2] \\
&= a^2 \sigma_X^2
\end{aligned}$$

Demostración de que $\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab \text{Cov}(X, Y)$

Se utiliza la notación $E(X)$ intercambiable con μ_X , $E(Y)$ intercambiable con μ_Y y así sucesivamente.

$$\begin{aligned}
\sigma_{aX+bY}^2 &= E[(aX + bY)^2] - \mu_{aX+bY}^2 \\
&= E(a^2 X^2 + 2abXY + b^2 Y^2) - \mu_{aX+bY}^2 \\
&= E(a^2 X^2) + E(2abXY) + E(b^2 Y^2) - (a\mu_X + b\mu_Y)^2 \\
&= a^2 E(X^2) + 2abE(XY) + b^2 E(Y^2) - a^2 \mu_X^2 - 2ab\mu_X \mu_Y - b^2 \mu_Y^2 \\
&= a^2 [E(X^2) - \mu_X^2] + b^2 [E(Y^2) - \mu_Y^2] + 2ab[E(XY) - \mu_X \mu_Y] \\
&= a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab \text{Cov}(X, Y)
\end{aligned}$$

Demostración de la equivalencia de las ecuaciones (2.68) y (2.69)

Se utiliza la notación $E(X)$ alternativamente con μ_X , $E(Y)$ alternativamente con μ_Y y así sucesivamente. Se debe demostrar que

$$E[(X - \mu_X)(Y - \mu_Y)] = \mu_{XY} - \mu_X \mu_Y$$

Ahora

$$\begin{aligned}
 E[(X - \mu_X)(Y - \mu_Y)] &= E(XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y) \\
 &= E(XY) - E(X\mu_Y) - E(Y\mu_X) + E(\mu_X\mu_Y) \\
 &= E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X\mu_Y \\
 &= \mu_{XY} - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y \\
 &= \mu_{XY} - \mu_X\mu_Y
 \end{aligned}$$

Demostración de que si X y Y son independientes entonces X y Y no están correlacionadas

Sean X y Y variables aleatorias independientes. Se demostrará que $\mu_{XY} = \mu_X\mu_Y$, por lo que $\text{Cov}(X, Y) = \rho_{X, Y} = 0$. Se supondrá que X y Y son variables conjuntamente continuas con una densidad conjunta $f(x, y)$ y densidades marginales $f_X(x)$ y $f_Y(y)$. La clave de esta demostración es el hecho de que dado que X y Y son independientes, $f(x, y) = f_X(x)f_Y(y)$.

$$\begin{aligned}
 \mu_{XY} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) \, dx \, dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) \, dx \, dy \\
 &= \int_{-\infty}^{\infty} x f_X(x) \, dx \int_{-\infty}^{\infty} y f_Y(y) \, dy \\
 &= \mu_X \mu_Y
 \end{aligned}$$

La demostración en el caso de que X y Y sean variables conjuntamente discretas es similar, sólo que las integrales son sustituidas por sumatorias.

Ejercicios para la sección 2.6

1. En una muestra de mil pernos seleccionada aleatoriamente, X es el número de pernos que no tienen la longitud específica y Y el número de pernos que no tienen el diámetro específico. Suponga que la función de masa de probabilidad conjunta de X y Y está dada por la siguiente tabla.
 - e) Determine la probabilidad de que todos los pernos de la muestra satisfagan la especificación de la longitud.
 - f) Determine la probabilidad de que todos los pernos de la muestra satisfagan la especificación del diámetro.
 - g) Determine la probabilidad de que todos los pernos de la muestra cumplan con ambas especificaciones.

x	y		
	0	1	2
0	0.40	0.12	0.08
1	0.15	0.08	0.03
2	0.10	0.03	0.01

- a) Determine $P(X = 0 \text{ y } Y = 2)$.
- b) Determine $P(X > 0 \text{ y } Y \leq 1)$.
- c) Determine $P(X \leq 1)$.
- d) Determine $P(Y > 0)$.
2. Con referencia al ejercicio 1.
 - a) Determine la función de masa de probabilidad marginal $\rho_X(x)$.
 - b) Determine la función de masa de probabilidad marginal $\rho_Y(y)$.
 - c) Determine μ_X .
 - d) Determine μ_Y .
 - e) Determine σ_X .
 - f) Determine σ_Y .

g) Determine $\text{Cov}(X, Y)$.

h) Determine $\rho_{X,Y}$.

i) ¿ X y Y son independientes? Explique.

3. Con referencia al ejercicio 1.

a) Determine la función de masa de probabilidad condicional $\rho_{Y|X}(y | 1)$.

b) Determine la función de masa de probabilidad condicional $\rho_{X|Y}(x | 1)$.

c) Determine la esperanza condicional $E(Y | X = 1)$.

d) Determine la esperanza condicional $E(X | Y = 1)$.

4. Un software puede hacer llamadas a dos subrutinas, A y B . En una ejecución elegida aleatoriamente, X es el número de llamadas hechas a la subrutina A y Y representa las llamadas a la subrutina B . La función de masa de probabilidad conjunta de X y Y está dada por la siguiente tabla.

x	y		
	1	2	3
1	0.15	0.10	0.10
2	0.10	0.20	0.15
3	0.05	0.05	0.10

a) Determine la función de masa de probabilidad marginal de X .

b) Determine la función de masa de probabilidad marginal de Y .

c) ¿ X y Y son independientes? Explique.

d) Determine μ_X y μ_Y .

e) Determine σ_X y σ_Y .

f) Determine $\text{Cov}(X, Y)$.

g) Encuentre $\rho(X, Y)$.

5. Con referencia al ejercicio 4, el número total de llamadas hechas a las dos subrutinas es $X + Y$.

a) Determine μ_{X+Y} .

b) Determine σ_{X+Y} .

c) Determine $P(X + Y = 4)$.

6. Con referencia al ejercicio 4.

a) Determine la función de masa de probabilidad condicional $\rho_{Y|X}(y | 2)$.

b) Determine la función de masa de probabilidad condicional $\rho_{X|Y}(x | 3)$.

c) Determine la esperanza condicional $E(Y | X = 2)$.

d) Determine la esperanza condicional $E(X | Y = 3)$.

7. Con referencia al ejercicio 4, suponga que cada ejecución de la subrutina A tarda 100 ms y que cada ejecución de la subrutina B tarda 200 ms.

a) Exprese el número de milisegundos de todas las llamadas hechas a las dos subrutinas en términos de X y Y .

b) Determine el número promedio de milisegundos de todas las llamadas realizadas a las dos subrutinas.

c) Encuentre la desviación estándar del número de milisegundos de todas las llamadas hechas a las dos subrutinas.

8. El número de clientes formados en una caja de supermercado es una variable aleatoria cuya función de masa de probabilidad está dada por la siguiente tabla.

x	0	1	2	3	4	5
$p(x)$	0.10	0.25	0.30	0.20	0.10	0.05

Para cada cliente, el número de productos que compra constituye una variable aleatoria con una función de masa de probabilidad

y	1	2	3	4	5	6
$p(y)$	0.05	0.15	0.25	0.30	0.15	0.10

Sea X el número de clientes formados y Y denota el número total de productos que compran todos los clientes formados. Suponga que el número de productos comprados por un cliente es independiente del número de productos comprados por cualquier otro cliente.

a) Determine $P(X = 2 \text{ y } Y = 2)$.

b) Determine $P(X = 2 \text{ y } Y = 6)$.

c) Determine $P(Y = 2)$.

9. En cierta intersección de una carretera de cuatro carriles, sólo hay dos para dar vuelta a la izquierda. El A está en el extremo izquierdo y el B está al lado de él. X representa el número de vehículos en el carril A , Y representa el número de vehículos en el carril B , cuando la señal cambia a verde. Suponga que X y Y tienen una función de masa de probabilidad conjunta $p(x, y)$ dada por la siguiente tabla.

x	y				
	0	1	2	3	4
0	0.05	0.04	0.01	0.00	0.00
1	0.05	0.10	0.03	0.02	0.00
2	0.03	0.05	0.15	0.05	0.02
3	0.00	0.02	0.08	0.10	0.05
4	0.00	0.00	0.02	0.05	0.08

- a) Determine la función de masa de probabilidad marginal de X .
- b) Determine la función de masa de probabilidad marginal de Y .
- c) ¿ X y Y son independientes? Explique.
- d) Determine μ_X y μ_Y .
- e) Determine σ_X y σ_Y .
- f) Determine $\text{Cov}(X, Y)$.
- g) Determine $\rho(X, Y)$.

10. Con referencia al ejercicio 9.

- a) Determine la media del número total de vehículos en los dos carriles.
- b) Determine la varianza del número total de vehículos en los dos carriles.
- c) Determine la probabilidad de que el número total de vehículos en los dos carriles sea exactamente igual a 6.

11. Con referencia al ejercicio 9.

- a) Determine la función de masa de probabilidad condicional $\rho_{Y|X}(y | 3)$.
- b) Determine la función de masa de probabilidad condicional $\rho_{X|Y}(x | 4)$.
- c) Determine la esperanza condicional $E(Y | X = 3)$.
- d) Determine la esperanza condicional $E(X | Y = 4)$.

- 12.** X representa el número de automóviles y Y el de camiones que pasan por cierta caseta de cobro en un minuto. La función de masa de probabilidad conjunta de X y Y está dada en la siguiente tabla:

x	y		
	0	1	2
0	0.10	0.05	0.05
1	0.10	0.10	0.05
2	0.05	0.20	0.10
3	0.05	0.05	0.10

- a) Determine la función de masa de probabilidad marginal $p_X(x)$.
- b) Determine la función de masa de probabilidad marginal $p_Y(y)$.
- c) Determine μ_X .
- d) Determine μ_Y .
- e) Determine σ_X .
- f) Determine σ_Y .
- g) Determine $\text{Cov}(X, Y)$.
- h) Determine $\rho_{X, Y}$.

- 13.** Con referencia al ejercicio 12, Z representa el número total de vehículos que pasa por la caseta de cobro en un minuto.

- a) Determine μ_Z .
- b) Determine σ_Z .
- c) Determine $P(Z = 2)$.

- 14.** Con referencia al ejercicio 12, suponga que la tarifa para autos es de \$2 y la de camiones es de \$5. T representa la cantidad total de las tarifas pagadas por los vehículos que pasan por la caseta de cobro en un minuto.

- a) Determine μ_T .
- b) Determine σ_T .
- c) Determine $P(T = 2)$.

- 15.** Con referencia al ejercicio 12.

- a) Determine la función de masa de probabilidad condicional $\rho_{Y|X}(y | 3)$.
- b) Determine la función de masa de probabilidad condicional $\rho_{X|Y}(x | 1)$.
- c) Determine la esperanza condicional $E(Y | X = 3)$.
- d) Determine la esperanza condicional $E(X | Y = 1)$.

- 16.** Para variables aleatorias continuas X y Y con una función de densidad de probabilidad conjunta

$$f(x, y) = \begin{cases} 4xy & 0 < x < 1 \text{ y } 0 < y < 1 \\ 0 & \text{de otro modo} \end{cases}$$

- a) Determine $P(X < 0.5 \text{ y } Y > 0.75)$.
- b) Determine las funciones de densidad de probabilidad marginal $f_X(x)$ y $f_Y(y)$.
- c) ¿ X y Y son independientes? Explique.

17. Con referencia al ejemplo 2.51 (p. 123).

- Determine $\text{Cov}(X, Y)$.
- Determine $\rho_{X, Y}$.

18. Para variables aleatorias continuas X y Y con una función de densidad de probabilidad conjunta

$$f(x, y) = \begin{cases} x + y & 0 < x < 1 \text{ y } 0 < y < 1 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine $P(X \leq 0.25 \text{ y } Y \leq 0.85)$.
- Determine $\text{Cov}(X, Y)$.
- Determine $\rho_{X, Y}$.
- ¿ X y Y son independientes? Explique.

19. Con referencia al ejercicio 18.

- Determine las funciones de densidad de probabilidad marginal $f_X(x)$ y $f_Y(y)$.
- Determine la función de densidad de probabilidad condicional $f_{Y|X}(y | 0.75)$.
- Determine la esperanza condicional $E(Y | X = 0.75)$.

20. Sea X la cantidad de encogimiento (en %) que experimenta una fibra de cierto tipo elegida aleatoriamente cuando se calienta a una temperatura de 120°C . Y representa el encogimiento adicional (en %) cuando la fibra se calienta a 140°C . Suponga que la función de densidad de probabilidad conjunta de X y Y está dada por

$$f(x, y) = \begin{cases} \frac{48xy^2}{49} & 3 < x < 4 \text{ y } 0.5 < y < 1 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine $P(X < 3.25 \text{ y } Y > 0.8)$.
- Determine las funciones de densidad de probabilidad marginal $f_X(x)$ y $f_Y(y)$.
- ¿ X y Y son independientes? Explique.

21. Se mide (en cm) la longitud y ancho de un componente rectangular. Debido a errores en la medición, las medidas son variables aleatorias. X denota la medida de la longitud y Y representa la medida del ancho. Suponga que la función de densidad de probabilidad de X es

$$f(x) = \begin{cases} 10 & 9.95 < x < 10.05 \\ 0 & \text{de otro modo} \end{cases}$$

y que la función de densidad de probabilidad de Y es

$$g(y) = \begin{cases} 5 & 4.9 < y < 5.1 \\ 0 & \text{de otro modo} \end{cases}$$

Suponga que las medidas de X y Y son independientes.

- Determine $P(X < 9.98)$.
- Determine $P(Y > 5.01)$.
- Determine $P(X < 9.98 \text{ y } Y > 5.01)$.
- Determine μ_X .
- Determine μ_Y .
- Sea $A = XY$ el área calculada a partir de las mediciones X y Y . Determine μ_A .

22. El espesor X de una cuña de madera (en milímetros) tiene una función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{3}{4} - \frac{3(x-5)^2}{4} & 4 \leq x \leq 6 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine μ_X .
- Determine σ_X^2 .
- Y denota el espesor de una cuña en pulgadas (1 mm = 0.0394 pulgadas). Determine μ_Y y σ_Y^2 .
- Si se seleccionan tres cuñas de manera independiente y las apilamos una encima de la otra, encuentre la media y la varianza del espesor total.

23. El tiempo de vida de cierto componente, en años, tiene una función de densidad de probabilidad

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Están disponibles dos de dichos componentes, cuyos tiempos de vida son independientes. Tan pronto como falle el primer componente, éste se reemplaza por el segundo. X representa el tiempo de vida del primer componente y Y denota el del segundo componente.

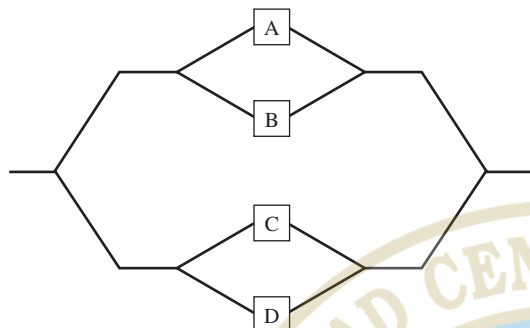
- Determine la función de densidad de probabilidad conjunta de X y Y .
- Determine $P(X \leq 1 \text{ y } Y > 1)$.
- Determine μ_X .
- Determine μ_{X+Y} .
- Determine $P(X + Y \leq 2)$. (Sugerencia: Grafique la región del plano en la que $x + y \leq 2$ y después integre la función de densidad de probabilidad conjunta en toda esa región.)

24. El número de bytes descargados por segundo en un canal de información tiene una media de 10^5 y desviación estándar de 10^4 . Entre los factores que influyen la razón está la congestión, la cual produce la alternación de periodos de transmisión más rápida o más lenta. X representa el número de bytes descargados durante un periodo elegido aleatoriamente de cinco segundos.
- a) ¿Es razonable suponer que $\mu_X = 5 \times 10^5$? Explique.
- b) ¿Es razonable suponer que $\sigma_X = \sqrt{5} \times 10^4$? Explique.
25. El artículo “Abyssal Periodities > 3 800 Ma from Southern West Greenland: Field Relationships, Petrography, Geochronology, Whole-Rock and Mineral Chemistry of Dunite and Harzburgite Inclusions in the Itsaq Gneiss Complex” (C. Friend, V. Bennett y A. Nutman, en *Contributions to Mineralogy and Petrology*, 2002:71-92) describe las composiciones químicas de ciertos minerales en el anterior manto arqueano. Para cierto tipo de ensamble olivino, el dióxido de silicio (SiO_2) contenido (en peso %) en una roca elegida de forma aleatoria tiene una media de 40.25 y desviación estándar de 0.36.
- a) Encuentre la media y la desviación estándar de la media de la muestra del SiO_2 contenido en una muestra aleatoria de diez rocas.
- b) ¿Cuántas rocas debe incluir la muestra para que la desviación estándar de la media de la muestra del SiO_2 contenido sea de 0.05?
26. Aquí hay dos variables aleatorias que no están correlacionadas ni son independientes. X y Y tienen la siguiente función de masa de probabilidad conjunta:
- a) Use la definición de independencia en la página 135 para demostrar que X y Y no son independientes (de hecho, $Y = |X|$, por lo que Y es realmente una función de X).
- b) Demuestre que X y Y no están correlacionadas.
27. Una persona tiene \$100 para invertir y dos opciones en las cuales dividirlos. Si ella aplica toda la cantidad en la primera inversión, su ganancia será X , mientras que si ella invierte la cantidad total en la segunda opción, su ganancia será Y . Tanto X como Y tienen una media de \$6 y desviación estándar (riesgo) de \$3. La correlación entre X y Y es 0.3.
- a) Expresé la cantidad en términos de X y Y si ella invierte \$30 en la primera inversión y \$70 en la segunda.
- b) Encuentre la ganancia promedio y el riesgo si invierte \$30 en la primera inversión y \$70 en la segunda.
- c) Encuentre la ganancia promedio y el riesgo, en términos de K , si ella invierte $\$K$ en la primera inversión y $\$(100 - K)$ en la segunda.
- d) Determine el valor de K que minimice el riesgo en la parte c).
- e) Demuestre que el valor de K que minimiza el riesgo en el inciso c) es el mismo para cualquier correlación $\rho_{X,Y} \neq 1$.
28. Si X es una variable aleatoria, demuestre que $\text{Cov}(X, X) = \sigma_X^2$.
29. X y Y son variables aleatorias y a y b son constantes.
- a) Demuestre que $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$.
- b) Demuestre que si $a > 0$ y $b > 0$, entonces $\rho_{aX, bY} = \rho_{X, Y}$. Concluya que el coeficiente de correlación no se ve afectado por el cambio en las unidades.

x	y	$p(x, y)$
-1	1	1/3
0	0	1/3
1	1	1/3

Ejercicios adicionales para el capítulo 2

1. Un sistema consta de cuatro componentes conectados, tal como se muestra.



Suponga que A , B , C y D funcionan independientemente. Si las probabilidades de que A , B , C y D fallen son 0.1, 0.2, 0.05 y 0.3, respectivamente, ¿cuál es la probabilidad de que el sistema funcione?

2. Se lanza un dado hasta que aparece el número 6. ¿Cuál es la probabilidad de que se necesiten más de cinco lanzamientos?
3. Las placas de silicón se utilizan en la producción de circuitos integrados. De las placas producidas por determinado proceso, 10% tiene resistencias menores a la especificación y 5% tiene resistencias superiores a la especificación.
- ¿Cuál es la probabilidad de que la resistencia de una placa elegida de forma aleatoria no satisfaga la especificación?
 - Si una placa escogida aleatoriamente tiene una resistencia que no cumple con la especificación, ¿cuál es la probabilidad de que tenga una resistencia demasiado baja?
4. Se utilizan dos líneas de producción para empaquetar azúcar en bolsas de 5 kg. La línea 1 produce el doble de bolsas que la línea 2. Uno por ciento de las bolsas de la línea 1 están defectuosas ya que no cumplen con una especificación de calidad, mientras que 3% de las bolsas de la línea 2 están defectuosas. Se elige aleatoriamente una bolsa para inspeccionarla.
- ¿Cuál es la probabilidad de que provenga de la línea 1?
 - ¿Cuál es la probabilidad de que esté defectuosa?
 - Si la bolsa está defectuosa, ¿cuál es la probabilidad de que venga de la línea 1?
 - Si la bolsa no está defectuosa, ¿cuál es la probabilidad de que venga de la línea 1?
5. Una cliente recibe un cargamento de diez mil fusibles. Ella extrae de forma aleatoria tres y los prueba. Si alguno está defectuoso, regresará el cargamento. Si, de hecho, mil están defectuosos, ¿cuál es la probabilidad de que regrese el cargamento?
6. En cierto tipo de motor de automóvil, la cabeza del cilindro está sujeta al bloque mediante diez pernos, cada uno debe experimentar un par de torsión de $60 \text{ N} \cdot \text{m}$. Suponga que los pares de torsión de los pernos son independientes.
- Se tiene una probabilidad de 0.99 si cada perno experimenta un par de torsión adecuado, ¿cuál es la probabilidad de que todos los pernos de la cabeza de un cilindro experimenten pares de torsión adecuados?
 - El objetivo es que en 95% de los motores todos sus pernos experimenten pares de torsión adecuados. ¿Cuál debe ser la probabilidad de que un perno experimente un par de torsión adecuado para que se alcance este objetivo?
7. Un mensaje electrónico consiste en una cadena de bits (0s y 1s). El mensaje debe pasar por dos relevadores antes de que sean recibidos. En cada relevador la probabilidad es de 0.1 de que el bit se regrese antes de que sea relevado (es decir, que un 1 cambie a 0, o un 0 a 1). Encuentre la probabilidad de que el valor de un bit recibido en el destino final es igual al valor del que fue enviado.
8. La lectura de un termómetro calibrado en agua helada (temperatura real de 0°C) representa una variable aleatoria con función de densidad de probabilidad.
- $$f(x) = \begin{cases} k(1 - x^2) & -1 < x < 1 \\ 0 & \text{de otro modo} \end{cases}$$
- donde k es una constante.
- Determine el valor de k .
 - ¿Cuál es la probabilidad de que el termómetro indique una temperatura mayor a 0°C ?
 - ¿Cuál es la probabilidad de que la lectura esté dentro los 0.25°C de la temperatura real?
 - ¿Cuál es la media de la lectura?
 - ¿Cuál es la mediana de la lectura?
 - ¿Cuál es la desviación estándar?
9. Se lanzan dos dados. En virtud de que salen dos números diferentes, ¿cuál es la probabilidad de que en uno de los dados salga el número 6?

10. En un lote de diez componentes se extrae aleatoriamente dos para inspección. Suponga que, de hecho, exactamente dos de los diez componentes del lote están defectuosos. Sea X el número de componentes extraídos que están defectuosos.

- Determine $P(X = 0)$.
- Determine $P(X = 1)$.
- Determine $P(X = 2)$.
- Determine la función de masa de probabilidad de X .
- Determine la media de X .
- Determine la desviación estándar de X .

11. Un dispositivo eléctrico tiene dos fusibles. Sea que X represente el tiempo de vida del primer fusible y sea que Y denote el del segundo (ambos en años). Suponga que la función de densidad de probabilidad conjunta de X y Y es

$$f(x, y) = \begin{cases} \frac{1}{6} e^{-x/2 - y/3} & x > 0 \text{ y } y > 0 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine $P(X \leq 2 \text{ y } Y \leq 3)$.
- Determine la probabilidad de que ambos fusibles duren al menos tres años.
- Determine la función de densidad de probabilidad marginal de X .
- Determine la función de densidad de probabilidad marginal de Y .
- ¿ X y Y son independientes? Explique.

12. Sean A y B eventos con $P(A) = 0.3$ y $P(A \cup B) = 0.7$.

- ¿Para qué valor de $P(B)$, A y B serán mutuamente excluyentes?
- ¿Para qué valor de $P(B)$, A y B serán independientes?

13. Un fabricante de tablas deslizadoras en nieve tiene tres plantas en Estados Unidos, una en el este, otra en el oeste y otra en Canadá. Los registros de producción muestran que cada una de las plantas en Estados Unidos produjo diez mil tablas deslizadoras el mes pasado, mientras que la planta en Canadá produjo ocho mil. De las tablas producidas en Canadá el mes pasado, 4% tiene un defecto que provocó que se deslaminaran prematuramente. Los registros de las plantas de Estados Unidos muestran que 3% de las tablas producidas en el este y 6% de las del oeste también presentan este defecto.

- ¿Cuál es la proporción de las tablas fabricadas el mes pasado que estaban defectuosas?

- ¿Cuál es la probabilidad de que una tabla deslizadora esté defectuosa y se haya fabricado en Canadá?
- Dado que una tabla deslizadora está defectuosa, ¿cuál es la probabilidad de que se haya fabricado en Estados Unidos?

14. El artículo “Traps in Mineral Valuations—Proceed With Care” (W. Lonegan, en *Journal of the Australian Institute of Mining and Metallurgy*, 2001:18-22) modela el valor (en millones de dólares) de un depósito de minerales que aún no se ha abierto como una variable aleatoria X con una función de masa de probabilidad $p(x)$ dada por $p(10) = 0.40$, $p(60) = 0.50$, $p(80) = 0.10$ y $p(x) = 0$ para valores de x diferentes a 10, 60 u 80.

- ¿Este artículo considera el valor de un depósito mineral como una variable aleatoria discreta o continua?
- Calcule μ_X .
- Calcule σ_X .
- El proyecto será rentable si el valor es mayor a \$50 millones. ¿Cuál es la probabilidad de que el proyecto sea rentable?

15. Seis recién graduados fueron contratados por una empresa de ingeniería. A cada uno se le asignó aleatoriamente uno de los seis cubículos que están en fila al final de la oficina. Dos de los graduados son Bill y Cathy. ¿Cuál es la probabilidad de que a ellos se les asigne cubículos adyacentes?

16. Un clóset contiene cuatro pares de zapatos. Si se escogen aleatoriamente cuatro zapatos, ¿cuál es la probabilidad de que de los zapatos elegidos no se tenga un par?

17. Sean X y Y variables aleatorias independientes con $\mu_X = 2$, $\sigma_X = 1$, $\mu_Y = 2$ y $\sigma_Y = 3$. Encuentre las medias y varianzas de las siguientes cantidades.

- $3X$
- $X + Y$
- $X - Y$
- $2X + 6Y$

18. Sean X y Y variables aleatorias independientes con $\mu_X = 1$, $\sigma_X = 2$, $\mu_Y = 3$, $\sigma_Y = 1$ y $\rho_{X,Y} = 0.5$. Encuentre las medias y varianzas de las siguientes cantidades.

- $X + Y$
- $X - Y$
- $3X + 2Y$
- $5Y - 2X$

19. Un productor de acero está probando un nuevo aditivo para fabricar una aleación de acero. La función de masa de probabilidad conjunta de la intensidad de la tensión (en miles de libras/pulg²) y la concentración aditiva es

Concentración de aditivo	Intensidad de la tensión		
	100	150	200
0.02	0.05	0.06	0.11
0.04	0.01	0.08	0.10
0.06	0.04	0.08	0.17
0.08	0.04	0.14	0.12

- a) ¿Cuáles son las funciones de masa de probabilidad marginal de X (concentración aditiva) y Y (intensidad de la tensión)?
- b) ¿ X y Y son independientes? Explique.
- c) Dado que un espécimen tiene una concentración aditiva de 0.04, ¿cuál es la probabilidad de que la intensidad de la tensión sea de 150 o mayor?
- d) Dado que una especie tiene una concentración aditiva de 0.08, ¿cuál es la probabilidad de que la intensidad de la tensión sea mayor a 125?
- e) Cierta aplicación requiere que la intensidad de la tensión sea igual o mayor a 175. ¿Qué concentración aditiva debe utilizarse para que la probabilidad de cumplir con esta especificación sea máxima?

20. Con referencia al ejercicio 19.

- a) Determine μ_X .
- b) Determine μ_Y .
- c) Determine σ_X .
- d) Determine σ_Y .
- e) Determine $\text{Cov}(X, Y)$.
- f) Determine $\rho_{X, Y}$.

21. Con referencia al ejercicio 19.

- a) Calcule la función de masa condicional $p_{Y|X}(y | 0.06)$.
- b) Calcule la función de masa condicional $p_{X|Y}(x | 100)$.
- c) Calcule la esperanza condicional $E(Y | X = 0.06)$.
- d) Calcule la esperanza condicional $E(X | Y = 100)$.

22. Cierta planta opera en tres etapas al día. De todas las unidades fabricadas, 50% se produce durante la primera etapa, 30% en la segunda y 20% en la tercera. De todas las unidades

des fabricadas en la primera etapa, 1% está defectuosa, mientras que 2% de la segunda y 3% de la tercera están defectuosas.

- a) Se extrae una unidad aleatoriamente de la producción diaria y está defectuosa. ¿Cuál es la probabilidad de que se haya fabricado durante la primera etapa?
- b) Se extrae una unidad aleatoriamente de la producción diaria y no está defectuosa. ¿Cuál es la probabilidad de que haya sido fabricada durante la tercera etapa?

23. El artículo “Uncertainty and Climate Change” (G. Heal y B. Kristöm, en *Environmental and Resource Economics*, 2002: 3-39) considera tres escenarios, A, B y C, para el impacto del calentamiento global en el ingreso. Para cada escenario se especifica una función de masa de probabilidad para la pérdida de ingreso. Éstas se presentan en la siguiente tabla.

Pérdida (%)	Probabilidad		
	Escenario A	Escenario B	Escenario C
0	0.65	0.65	0.65
2	0	0	0.24
5	0.2	0.24	0.1
10	0	0	0.01
15	0.1	0.1	0
20	0	0.01	0
25	0.05	0	0

- a) Calcule la media y desviación estándar de la pérdida en el escenario A.
- b) Calcule la media y desviación estándar de la pérdida en el escenario B.
- c) Calcule la media y desviación estándar de la pérdida en el escenario C.
- d) En cada escenario, calcule la probabilidad de que la pérdida sea menor a 10 por ciento.

24. Con referencia al ejercicio 23, suponga que las probabilidades de que ocurra cada uno de los tres escenarios son $P(A) = 0.20$, $P(B) = 0.30$ y $P(C) = 0.50$.

- a) Determine la probabilidad de que ocurra el escenario A y que la pérdida sea de 5%.
- b) Encuentre la probabilidad de que la pérdida sea de 5%.
- c) Determine la probabilidad de que suceda el escenario A dado que la pérdida es de 5%.

25. Una caja contiene cuatro focos de 75W, tres de 60W y tres focos fundidos. Se selecciona aleatoriamente dos focos de

la caja. Sea X el número de focos seleccionados de 75 W, Y representa el número de focos seleccionados de 60 W.

- Determine la función de masa de probabilidad conjunta de X y Y .
- Determine μ_X .
- Determine μ_Y .
- Determine σ_X .
- Determine σ_Y .
- Determine $\text{Cov}(X, Y)$.
- Determine $\rho_{X,Y}$.

26. Una provisión de una solución de ácido clorhídrico (HCl) distribuida por cierto vendedor contiene pequeñas cantidades de diversas impurezas, incluyendo cobre y níquel. Sea X la cantidad de cobre y Y representa la cantidad de níquel, en partes por diez millones, en una botella de la solución seleccionada aleatoriamente. Suponga que la función de densidad de probabilidad conjunta de X y Y está dada por

$$f(x, y) = \begin{cases} c(x + y)^2 & 0 < x < 1 \text{ y } 0 < y < 1 \\ 0 & \text{de otro modo} \end{cases}$$

- Determine el valor de la constante c de tal forma que $f(x, y)$ sea una función de densidad conjunta.
- Calcule la función de densidad marginal $f_X(x)$.
- Calcule la función de densidad condicional $f_{X|Y}(y | x)$.
- Calcule la esperanza condicional $E(Y | X = 0.4)$.
- ¿ X y Y son independientes? Explique.

27. Con referencia al ejercicio 26.

- Determine μ_X .
- Determine σ_X^2 .
- Determine $\text{Cov}(X, Y)$.
- Determine $\rho_{X,Y}$.

28. Se lanza una moneda al aire cinco veces. ¿Qué secuencia es más probable, HTTHH o HHHHH? ¿O son igualmente probables? Explique.

29. Se lanza al aire una moneda de 25 centavos y otra de 5 centavos. La primera tiene una probabilidad de 0.4 de que caiga en “cara” y la segunda tiene una probabilidad de 0.6 de que caiga “cara”. Sea $X = 1$ si la primera moneda cae en “cara” y sea $X = 0$ si cae en “cruz”. Sea $Y = 1$ si la moneda de 5 centavos, la segunda, cae en “cara” y $Y = 0$ si cae en “cruz”.

- Determine la función de masa de probabilidad de X .
- Determine la función de masa de probabilidad de Y .
- ¿Es razonable suponer que X y Y son independientes? ¿Por qué?
- Determine la función de masa de probabilidad conjunta de X y Y .

30. Se lanzan dos dados. X representa el número que aparece en el primero y Y el que aparece en el segundo. Encuentre μ_{XY} .

31. Una caja contiene tres cartas, 1, 2 y 3. Se eligen aleatoriamente dos de ellas, se reemplaza la primera antes de que salga la segunda. X representa el número en la primera y Y representa en la segunda.

- Determine la función de masa de probabilidad conjunta de X y Y .
- Determine la función de masa de probabilidad marginal $p_X(x)$ y $p_Y(y)$.
- Determine μ_X y μ_Y .
- Determine μ_{XY} .
- Determine $\text{Cov}(X, Y)$.

32. Con referencia al ejercicio 31, suponga que no se reemplaza la primera carta antes de que salga la segunda.

- Determine la función de masa de probabilidad conjunta de X y Y .
- Determine la función de masa de probabilidad marginal $p_X(x)$ y $p_Y(y)$.
- Determine μ_X y μ_Y .
- Determine μ_{XY} .
- Determine $\text{Cov}(X, Y)$.

Capítulo 3

Propagación de errores

Introducción

La medición es fundamental en el trabajo de investigación. Con frecuencia científicos e ingenieros realizan cálculos con cantidades medidas; por ejemplo, para calcular la densidad de un objeto se divide una medición de su masa entre la de su volumen, o al calcular el área de un rectángulo se multiplican su longitud y anchura.

Cualquier procedimiento de medición tiene errores. Por consiguiente, en general los valores medidos son algo diferentes de los valores reales. Cuando se realiza un cálculo con mediciones, los errores en éstas producen un error en el valor calculado. Decimos que el error se **propaga** de las mediciones al valor calculado. Si se tiene cierto conocimiento con respecto al tamaño de los errores en las mediciones, como en la longitud y el ancho de un rectángulo, existen métodos para conocer la magnitud del error en una cantidad calculada como el área. La propagación de errores trata estos métodos y es el tema de este capítulo.

3.1 Error de medición

Una geóloga pesa una roca en una balanza. Toma cinco mediciones y obtiene los siguientes datos (en gramos):

251.3 252.5 250.8 251.1 250.4

Todas las mediciones son diferentes y es probable que ninguna sea igual a la masa real de la roca. A la diferencia entre un valor medido y el valor real se le llama **error** en el valor medido. Cualquier procedimiento de medición tiene muchas fuentes de error. Por ejemplo, supon-

ga que las mediciones de la roca se leían en una marca en una escala. Si la balanza no estaba calibrada adecuadamente, cada medición estará lejos de su valor real en cierta cantidad fija. Por tanto, una calibración imperfecta aporta errores de la misma magnitud en cada medición. La interpolación entre las marcas de graduación de la escala es otra fuente de error. La magnitud del error debida a la interpolación quizá varíe entre mediciones y es probable que sea benéfico para algunas mediciones y negativo para otras. Es razonable suponer que a largo plazo el promedio de los errores por interpolación será igual a cero.

En general, se puede pensar que el error de una medición lo integran el **error sistemático**, o **sesgo**, y el **error aleatorio**. El primero representa la parte del error que es igual para cada medición; el segundo varía entre mediciones y, en promedio, será igual a cero en el largo plazo. Algunas fuentes de error contribuyen con ambos tipos de error, el sesgo y el error aleatorio. Por ejemplo, considere el error de paralaje. Este último constituye la diferencia en la posición evidente de la marca cuando se observa desde ángulos diferentes. La magnitud de este tipo de error en cualquier medición especial depende de la posición del observador con respecto a la escala. Como consecuencia de que la posición variará un poco entre lecturas, el paralaje contribuye al error aleatorio. Si el observador tiende a apoyarse en algo de un lado en vez de otro, el paralaje también contribuirá al sesgo.

Cualquier medición se puede considerar como la suma del valor real más las contribuciones de cada uno de los dos componentes del error:

$$\text{Valor medido} = \text{Valor real} + \text{sesgo} + \text{error aleatorio} \quad (3.1)$$

Puesto que parte del error es aleatorio, es adecuado utilizar un modelo estadístico para estudiar los errores de medición. Se modela cada valor medido como una variable aleatoria, tomada de una población de mediciones posibles. La media μ de la población representa esa parte de la medición que es igual para toda medición. Por tanto, μ es la suma del valor real más el sesgo. La desviación estándar σ de la población representa la desviación estándar del error aleatorio. Ésta representa la variación debida al hecho de que cada medición tiene un valor diferente por su error aleatorio. Intuitivamente, σ constituye el tamaño de un error aleatorio estándar.

Se tiene interés en dos aspectos del proceso de medición. El primero es su **exactitud**. Ésta la determina el sesgo, que es la diferencia entre la media μ de la medición y el valor real de esta última. Entre más pequeño sea el sesgo, más exacto será el proceso de medición. Si la media μ es igual al valor real, el sesgo será igual a 0; en esta tesitura, al proceso de medición se le llama **no sesgado**.

El otro aspecto del proceso de medición de interés es la **precisión**. Ésta constituye el grado con que tienden a coincidir las mediciones repetidas de la misma cantidad. Si las mediciones repetidas resultan cercanas entre sí todo el tiempo, la precisión es alta. Si son muy dispersas, la precisión es baja. Por tanto, la precisión se determina mediante la desviación estándar σ del proceso de medición. Entre más pequeño sea el valor de σ , más preciso será aquél. Con frecuencia ingenieros y científicos se refieren a σ como **incertidumbre aleatoria** o **incertidumbre estadística** del proceso de medición. A σ se le llamará en forma más simple **incertidumbre**.

Cuando se notifica un valor medido, es importante reportar una estimación aproximada del sesgo y de la incertidumbre de éste, con la finalidad de describir la exactitud y la precisión de la medición. Generalmente es más fácil estimar la incertidumbre que el sesgo. Las figuras 3.1 y 3.2 ilustran el porqué de lo anterior. La figura 3.1 muestra un experimento

Resumen

- Un valor medido representa una variable aleatoria con media μ y desviación estándar σ .
- El sesgo en el proceso de medición constituye la diferencia entre la media de las mediciones y el valor real:

$$\text{Sesgo} = \mu - \text{valor real}$$

- La incertidumbre en el proceso de medición es la desviación estándar σ .
- Entre más pequeño sea el sesgo, más exacto será el proceso de medición.
- Entre más pequeña sea la incertidumbre, más preciso será el proceso de medición.

hipotético que implica mediciones repetidas, en condiciones diferentes, donde se considera al sesgo y la incertidumbre. Los conjuntos de mediciones en las figuras 3.1a y b están bastante cercanos, indicando que la incertidumbre es pequeña. Por su parte, los que corresponden a las figuras 3.1a y c se encuentran centrados cerca del valor real, indicando que el sesgo es pequeño.

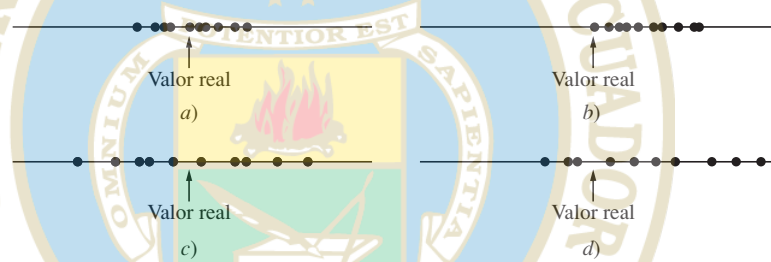


FIGURA 3.1 a) Tanto el sesgo como la incertidumbre son pequeños. b) El sesgo es grande; la incertidumbre es pequeña. c) El sesgo es pequeño; la incertidumbre es grande. d) Tanto el sesgo como la incertidumbre son grandes.

Por supuesto, en la vida real no se conoce el valor real que se está midiendo. Por tanto, los dibujos de las mediciones que se muestran en la figura 3.1 se parecerían a la figura 3.2 (p. 160). Se puede determinar que los conjuntos de mediciones en las figuras 3.2a y b tienen incertidumbre más pequeña. Pero sin información adicional acerca del valor real, no se puede calcular el sesgo.

Se concluye de las figuras 3.1 y 3.2 que *la incertidumbre se puede calcular de las mediciones repetidas, pero para estimar el sesgo, se debe tener información adicional acerca del valor real*. Se puede obtener esta información adicional, por ejemplo, midiendo repetidamente una cantidad usual cuyo valor real se conoce y estimar al sesgo como la diferencia entre el promedio de las mediciones y el valor real conocido. Otra manera para calcular el sesgo sería comparar el promedio de gran número de mediciones con otra hecha con un proceso más elaborado para el que se sabe que el sesgo es despreciable. La estimación del sesgo es esencialmente el proceso de calibración, para el cual se necesita información externa al dispositivo de medición.

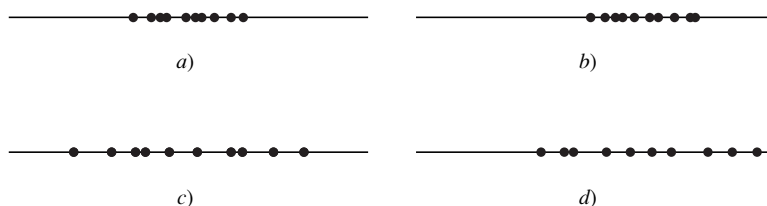


FIGURA 3.2 Se puede estimar la incertidumbre del conjunto de mediciones repetidas, pero si no se conoce el valor real no es posible calcular el sesgo.

Ejemplo

3.1

Se sabe que una muestra de laboratorio de gas tiene una concentración de monóxido de carbono (CO) de 50 partes por millón (ppm). Se utiliza un espectrómetro para tomar cinco mediciones independientes de esta concentración. Las cinco mediciones, en ppm, son 51, 47, 53, 53 y 48. Estime el sesgo y la incertidumbre en una medición del espectrómetro.

Solución

Se consideran las cinco mediciones como una muestra aleatoria de la población de mediciones posibles. El sesgo es igual a la media de esta población menos el valor real de 50. La incertidumbre representa la desviación estándar de la población. No se conoce la media ni la desviación estándar de la población, pero éstas se pueden aproximar con la media y la desviación estándar de la muestra. La media de las cinco mediciones es 50.4. Por tanto, se estima que el sesgo es de $50.4 - 50 = 0.4$ ppm. La desviación estándar de las cinco mediciones es 2.8 ppm. Por consecuencia, se estima que la incertidumbre en cada medición es de 2.8 ppm.

Ejemplo

3.2

Ahora se utiliza un espectrómetro diferente para medir la concentración de CO en otra muestra de gas. La concentración real de esta muestra es desconocida. Se hacen cinco mediciones (en ppm). Éstas son 62, 63, 61, 62 y 59. Estime la incertidumbre en una medición de este espectrómetro. ¿Se puede estimar el sesgo?

Solución

La incertidumbre en una sola medición se estima con la desviación estándar de la muestra, que es 1.5 ppm. La media de la muestra es 61.4 ppm, pero para estimar el sesgo se tendría que restar la concentración real de la media. Debido a que no se conoce la concentración real, no se puede estimar el sesgo.

En la práctica, las estimaciones aproximadas de la incertidumbre son a veces muy aproximadas. En los ejemplos 3.1 y 3.2 se sugiere estimar la incertidumbre σ con la desviación estándar de la muestra de cinco mediciones. Las estimaciones que se basan en muestras pequeñas parecidas a esta en ocasiones están muy equivocadas. Cuando es posible, es mejor estimar la incertidumbre con base en muestras grandes. Sin embargo, una estimación de una muestra pequeña es mejor que ninguna.

Resumen

Sean $X_1, \dots, X_n$ mediciones independientes, todas se hacen con el mismo proceso en la misma cantidad.

- La desviación estándar s de la muestra se puede utilizar para estimar la incertidumbre.
- Las estimaciones de la incertidumbre con frecuencia son muy aproximadas, especialmente cuando se basan en muestras pequeñas.
- Si se conoce el valor real, la media de la muestra $\bar{X}$ se puede utilizar para estimar el sesgo: $\text{Sesgo} \approx \bar{X}$, valor real.
- Si el valor real no se conoce, el sesgo no se puede estimar a partir de las mediciones repetidas.

Un ejemplo importante de la estimación de sesgo es la calibración de balanzas en supermercados y otros establecimientos comerciales para asegurar que no pesan sistemáticamente de más o de menos los artículos. A este procedimiento le sigue una serie de comparaciones con patrones externos, comenzando en el ámbito jurisdiccional y terminando cerca de París, Francia, donde se localiza el patrón mundial final para el peso (técnicamente la masa). Éste es el prototipo internacional del kilogramo, un cilindro de platino-iridio cuya masa es por definición exactamente 1 kg. Una réplica del kilogramo se localiza en el Laboratorio Nacional de Normas y Tecnología, en Washington, el cual sirve de patrón para todas las mediciones de peso en Estados Unidos. El uso de esta réplica, en vez del kilogramo, introduce un sesgo en cada medición de peso en ese país. Comparando la réplica de Estados Unidos con el kilogramo, el sesgo se ha estimado en -1.9×10^{-8} kg. Es decir, la réplica que se encuentra en Estados Unidos parece más ligera que el kilogramo en aproximadamente 19 partes en mil millones. Por esta razón, todas las mediciones de peso que se hacen en el Laboratorio Nacional de Normas y Tecnología se ajustan hacia arriba en 19 partes en mil millones para compensar. Observe que este factor de ajuste no se podía haber calculado pesando repetidamente la réplica; se requirió comparar con un patrón externo.

De ahora en adelante, a menos que se indique lo contrario, se supondrá que este sesgo se ha reducido a un nivel despreciable. Se describirán las mediciones en la forma

$$\text{Valor medido} \pm \sigma \quad (3.2)$$

donde σ representa la incertidumbre en el proceso que produjo el valor medido.

La expresión (3.2) tiene la forma $a \pm b$, donde a y b son números. Es importante darse cuenta que las expresiones que contienen el símbolo $\pm$ pueden tener muchos significados. El significado aquí es que a es un valor medido y b constituye la incertidumbre en a . Algunas personas usan $a \pm b$ para indicar que b es el valor máximo para el error, o que b es un múltiplo de la incertidumbre, generalmente dos o tres veces la incertidumbre. Se presentará incluso otro significado en el capítulo 5, donde se usará la notación $a \pm b$ para denotar un *intervalo de confianza*, que es un intervalo que se calcula de tal forma para que probablemente contenga al valor real. Siempre que se encuentre con el símbolo $\pm$ se debe asegurar que se comprende el contexto en el cual se utiliza.

Ejemplo

3.3

El espectrómetro del ejemplo 3.1 se ha recalibrado, por lo que se puede suponer que el sesgo es despreciable. Ahora aquél se utiliza para medir la concentración de CO en otra muestra de gas. La medición es 55.1 ppm. ¿Cómo se expresaría esta medición?

Solución

A partir de las mediciones repetidas del ejemplo 3.1, la incertidumbre en una medición de este instrumento se estimó de 2.8 ppm. Por tanto, se reporta la concentración de CO en esta muestra de gas como 55.1 ± 2.8 ppm.

Ejercicios de la sección 3.1

- Se mide cuatro veces el punto de ebullición del agua. Los resultados son 110.01°C , 110.02°C , 109.99°C y 110.01°C . ¿Cuál de los siguientes enunciados describe mejor este proceso de medición?
 - Exacto, pero no preciso.
 - Preciso, pero no exacto.
 - Ni exacto ni preciso.
 - Tanto exacto como preciso.
- Se usan dos aparatos para medir el punto de fusión del *p*-amino-benceno. Se toman números iguales de mediciones en cada uno de los aparatos. El resultado en el primer aparato es $90 \pm 1^\circ\text{C}$ y el resultado en el segundo es $90 \pm 2^\circ\text{C}$.
 - ¿Se puede decir qué aparato es más exacto? Si es así, indique cuál. Si no, explique por qué.
 - ¿Se puede decir qué aparato es más preciso? Si es así, diga cuál. Si no, explique por qué.
- La longitud de un objeto está dada como 3.21 ± 0.02 cm. Verdadero o falso:
 - La longitud se midió de 3.21 cm.
 - La longitud real del objeto es de 3.21 cm.
 - El sesgo en la medición es de 0.02 cm.
 - La incertidumbre en la medición es de 0.02 cm.
- Para algunos procesos de medición, la incertidumbre es aproximadamente proporcional al valor de la medición. Por ejemplo, se dice que cierta balanza tiene una incertidumbre de $\pm 2\%$. Se pesa un objeto en esta balanza.
 - Dado que la lectura es de 100 g, exprese la incertidumbre en esta medición, en gramos.
 - Dado que la lectura es de 50 g, exprese la incertidumbre en esta medición, en gramos.
- Una persona se para sobre una balanza de baño. La lectura es de 150 lb. Después de que la persona se baja, dicha lectura es de 2 libras.
 - ¿Es posible estimar la incertidumbre en esta medición? Si es así, estímelas. Si no, explique por qué.
 - ¿Es posible estimar el sesgo en esta medición? Si es así, estímelas. Si no, explique por qué.
- Una persona sube y baja de una balanza de baño cuatro veces. Las lecturas (en libras) son 148, 151, 150 y 152. Cada vez que la persona se baja de la balanza, la lectura es de 2 libras.
 - ¿Es posible estimar la incertidumbre en estas mediciones? Si es así, estímelas. Si no, explique por qué.
 - ¿Es posible estimar el sesgo en estas mediciones? Si es así, estímelas. Si no, explique por qué.
- En un escenario hipotético, el Laboratorio Nacional de Normas y Tecnología ha recibido una nueva réplica del kilogramo. Se pesa cinco veces. Las mediciones son las siguientes, en unidades de microgramos arriba de 1 kg: 114.3, 82.6, 136.4, 126.8, 100.7.
 - ¿Es posible estimar la incertidumbre en estas mediciones? Si es así, estímelas. Si no, explique por qué.

- b) ¿Es posible estimar el sesgo en estas mediciones? Si es así, estímelo. Si no, explique por qué.
8. Ahora se pesa el kilogramo cinco veces en una balanza diferente. Las mediciones son las siguientes, en unidades de microgramos arriba de 1 kg: 25.6, 26.8, 26.2, 26.8, 25.4.
- a) ¿Es posible estimar la incertidumbre en estas mediciones? Si es así, estímelas. Si no, explique por qué.
- b) ¿Es posible estimar el sesgo en estas mediciones? Si es así, estímelo. Si no, explique por qué.
9. Un nuevo y desconocido peso se pesa en la misma balanza que se utilizó en el ejercicio número 8 y la medición es 127 μg arriba de 1 kg. Utilizando la información del ejercicio 8, ¿es posible tener una medición más exacta? Si es así, ¿cuál es ésta? Si no, explique por qué.
10. El artículo “Calibration of an FTIR Spectrometer” (P. Pan-kratz, en *Statistical Case Studies for Industrial and Process Improvement*, SIAM-ASA, 1997:19-38) describe el uso de un espectrómetro para hacer cinco mediciones del contenido de carbono (en ppm) de cierta placa de silicio cuyo contenido real de carbono se sabía que era de 1.1447 ppm. Las mediciones fueron 1.0730, 1.0825, 1.0711, 1.0870 y 1.0979.
- a) ¿Es posible estimar la incertidumbre en estas mediciones? Si es así, estímelas. Si no, explique por qué.
- b) ¿Es posible estimar el sesgo en estas mediciones? Si es así, estímelo. Si no, explique por qué.
11. Se midió ocho veces la longitud de una varilla. Las mediciones en centímetros, en el orden en que fueron tomados, fueron 21.20, 21.22, 21.25, 21.26, 21.28, 21.30, 21.32 y 21.35.
- a) ¿Estas mediciones parecen ser una muestra aleatoria de una población de mediciones posibles? ¿Por qué sí? o ¿Por qué no?
- b) ¿Es posible estimar la incertidumbre en estas mediciones? Explique.

3.2 Combinaciones lineales de las mediciones

A menudo se suman constantes a las mediciones, se multiplican mediciones por constantes, o se suman dos o más mediciones. En esta sección se describe cómo se afectan las incertidumbres debido a estas operaciones aritméticas. Puesto que las mediciones son variables aleatorias y las incertidumbres son desviaciones estándar de estas variables aleatorias, los resultados que se usan para calcular las desviaciones estándar de combinaciones lineales de variables aleatorias se pueden aplicar para calcular las incertidumbres en combinaciones lineales de las mediciones. Los resultados para variables aleatorias independientes se presentaron en la sección 2.5; los resultados más generales se mostraron en la sección 2.6. Aquí se aplican estos resultados a las mediciones independientes. Al final se analizan las mediciones dependientes.

Se empieza por establecer los resultados básicos para calcular las incertidumbres en combinaciones lineales de mediciones independientes y después se continúa con algunos ejemplos.

Si X es una medición y c es una constante, entonces

$$\sigma_{cX} = |c|\sigma_X \quad (3.3)$$

Si $X_1, \dots, X_n$ son mediciones independientes y $c_1, \dots, c_n$ son constantes, entonces

$$\sigma_{c_1X_1 + \dots + c_nX_n} = \sqrt{c_1^2\sigma_{X_1}^2 + \dots + c_n^2\sigma_{X_n}^2} \quad (3.4)$$

Ejemplo**3.4**

El radio de un círculo mide 3.0 ± 0.1 cm. Estime la circunferencia y determine la incertidumbre en la estimación.

Solución

Sea R el radio del círculo. El valor medido de R es de 3.0 cm y la incertidumbre es la desviación estándar de esta medición, que es $\sigma_R = 0.1$ cm. La circunferencia está dada por $C = 2\pi R$. La incertidumbre en C es σ_C , la desviación estándar de C . Debido a que 2π es una constante, se tiene

$$\begin{aligned}\sigma_C &= |2\pi|\sigma_R && \text{(utilizando la ecuación 3.3)} \\ &= (6.28)(0.1 \text{ cm}) \\ &= 0.63 \text{ cm}\end{aligned}$$

La circunferencia es 18.84 ± 0.63 cm.

Ejemplo**3.5**

Un artículo se forma al colocar dos componentes, uno detrás de otro. Las longitudes de éstos se miden de manera independiente, por medio de un proceso que da una medición aleatoria con incertidumbre 0.1 cm. La longitud del artículo se estima sumando las dos longitudes medidas. Suponga que las mediciones son 4.10 y 3.70 cm. Estime la longitud del artículo y determine la incertidumbre en el cálculo.

Solución

Sea X la longitud medida del primer componente y Y la longitud medida del segundo. La estimación de la longitud es 7.80 cm. La incertidumbre es

$$\begin{aligned}\sigma_{X+Y} &= \sqrt{\sigma_X^2 + \sigma_Y^2} && \text{(utilizando la ecuación 3.4 con } c_1 = c_2 = 1) \\ &= \sqrt{(0.1)^2 + (0.1)^2} \\ &= 0.14 \text{ cm}\end{aligned}$$

La estimación de la longitud es 7.80 ± 0.14 cm.

Ejemplo**3.6**

Un topógrafo mide el perímetro de un terreno rectangular. Toma medidas de dos lados adyacentes, 50.11 ± 0.05 m y 75.21 ± 0.08 m. Estas mediciones son independientes. Estime el perímetro del terreno y determine la incertidumbre en la estimación.

Solución

Sean $X = 50.11$ y $Y = 75.21$ las dos mediciones. El perímetro se estima como $P = 2X + 2Y = 250.64$ m y la incertidumbre en P es

$$\begin{aligned}\sigma_P &= \sigma_{2X+2Y} \\ &= \sqrt{4\sigma_X^2 + 4\sigma_Y^2} && \text{(utilizando la ecuación 3.4)}\end{aligned}$$

$$\begin{aligned}
 &= \sqrt{4(0.05)^2 + 4(0.08)^2} \\
 &= 0.19 \text{ m}
 \end{aligned}$$

El perímetro es 250.64 ± 0.19 m.

Ejemplo

3.7

En el ejemplo 3.6 el asistente del topógrafo sugiere que se calcule la incertidumbre en P mediante un método diferente. Su razonamiento es que $P = X + X + Y + Y$, entonces

$$\begin{aligned}
 \sigma_P &= \sigma_{X+X+Y+Y} \\
 &= \sqrt{\sigma_X^2 + \sigma_X^2 + \sigma_Y^2 + \sigma_Y^2} \\
 &= \sqrt{(0.05)^2 + (0.05)^2 + (0.08)^2 + (0.08)^2} \\
 &= 0.13 \text{ m}
 \end{aligned}$$

Éste no concuerda con el valor de 0.19 m calculado en el ejemplo 3.6. ¿Cuál fue la equivocación?

Solución

La equivocación fue que de los cuatro términos de la suma no todos son independientes. Específicamente, $X + X$ no es la suma de cantidades independientes y tampoco $Y + Y$. Con el fin de utilizar la ecuación (3.4) para calcular la incertidumbre en P , se debe expresar a P como la suma de cantidades independientes; es decir, $P = 2X + 2Y$, como en el ejemplo 3.6.

Mediciones repetidas

Una de las mejores maneras de reducir la incertidumbre es tomar varias mediciones independientes y determinar el promedio de ellas. Las mediciones en este caso son una muestra aleatoria simple de una población y su promedio es la media de la muestra. Los métodos para el cálculo de la media y de la desviación estándar de una muestra se presentaron en las secciones 2.5 y 2.6. Estos métodos se pueden aplicar para calcular la media y la incertidumbre en el promedio de las mediciones repetidas independientes.

Si $X_1, \dots, X_n$ son mediciones n independientes, cada una con media μ e incertidumbre σ , entonces la media de la muestra $\bar{X}$ es una medición con media

$$\mu_{\bar{X}} = \mu \quad (3.5)$$

y con incertidumbre

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad (3.6)$$

Con un poco de paciencia se puede ver por qué son importantes estos resultados para las aplicaciones. De hecho muestran que si se realizan muchas mediciones independientes de la misma cantidad, entonces el promedio de éstas tiene la misma media de cada medición in-

dividual, pero la desviación estándar se reduce en un factor igual a la raíz cuadrada del tamaño de la muestra. En otras palabras, el promedio de varias mediciones repetidas tiene la misma exactitud y es más preciso que cualquier única medición.

Ejemplo 3.8

La longitud de un componente se mide con un proceso cuya incertidumbre es de 0.05 cm. Si se hicieron 25 mediciones independientes y se utiliza el promedio de éstas para estimar la longitud, ¿cuál será la incertidumbre? ¿Cuánto más preciso es el promedio de 25 mediciones que el de una sola medición?

Solución

La incertidumbre en el promedio de 25 mediciones es $0.05/\sqrt{25} = 0.01$ cm. La incertidumbre en una sola medición es 0.05 cm. La incertidumbre en el promedio de las 25 mediciones independientes es, por tanto, menor que el de una sola medición por un factor de 5, que es la raíz cuadrada del número de mediciones que se promedian. Por tanto, el promedio de 25 mediciones independientes es cinco veces más preciso que el de una sola medición.

Ejemplo 3.9

La masa de una roca se midió cinco veces en una balanza cuya incertidumbre no se conoce. Las cinco mediciones (en gramos) son 21.10, 21.05, 20.98, 21.12 y 21.05. Estime la masa de la roca y determine la incertidumbre en la estimación.

Solución

Sea $\bar{X}$ el promedio de las cinco mediciones y sea s la desviación estándar de la muestra. Se calcula $\bar{X} = 21.06$ g y $s = 0.0543$ g. Utilizando la ecuación (3.6) se estimaría que la longitud del componente es de $\bar{X} \pm \sigma/\sqrt{5}$. No se conoce σ , que es la incertidumbre, o la desviación estándar, del proceso de medición. Sin embargo, se puede aproximar a σ con s , la desviación estándar de la muestra de las cinco mediciones. Por tanto, se estima que la masa de la roca es de $21.06 \pm 0.0543/\sqrt{5}$, o 21.06 ± 0.02 g.

Ejemplo 3.10

En el ejemplo 3.6 dos lados adyacentes de un terreno rectangular se calcularon de $X = 50.11 \pm 0.05$ m y $Y = 75.21 \pm 0.08$ m. Se supone que el presupuesto para este proyecto es suficiente para hacer 14 mediciones más. Cada lado ya se ha medido una vez. Un ingeniero sugiere asignar las nuevas mediciones a cada lado equitativamente, por lo que éste será medido ocho veces. Un segundo ingeniero sugiere hacer las 14 mediciones en el lado más largo, ya que ese lado se mide con incertidumbre más grande. Estime la incertidumbre en el perímetro bajo cada plan. ¿Con cuál plan se obtiene la incertidumbre más pequeña?

Solución

Con el primer plan, sea $\bar{X}$ el promedio de las ocho mediciones del lado más corto y sea $\bar{Y}$ el promedio de las ocho mediciones del lado más largo. El perímetro se estimará con $2\bar{X} + 2\bar{Y}$. La incertidumbre en el perímetro con el primer plan es, por tanto,

$$\begin{aligned}
 \sigma_{2\bar{X}+2\bar{Y}} &= \sqrt{4\sigma_X^2 + 4\sigma_Y^2} && \text{(utilizando la ecuación 3.4)} \\
 &= \sqrt{4\left(\frac{\sigma_X}{\sqrt{8}}\right)^2 + 4\left(\frac{\sigma_Y}{\sqrt{8}}\right)^2} && \text{(utilizando la ecuación 3.6)} \\
 &= \sqrt{\frac{4(0.05)^2}{8} + \frac{4(0.08)^2}{8}} \\
 &= 0.067 \text{ m}
 \end{aligned}$$

Con el segundo plan, el perímetro se estimará con $2X + 2\bar{Y}$, donde X es una sola medición del lado más corto y $\bar{Y}$ es el promedio de las 15 mediciones del lado más largo. La incertidumbre en el perímetro con el segundo plan es, por tanto,

$$\begin{aligned}
 \sigma_{2X+2\bar{Y}} &= \sqrt{4\sigma_X^2 + 4\sigma_Y^2} && \text{(utilizando la ecuación 3.4)} \\
 &= \sqrt{4\sigma_X^2 + 4\left(\frac{\sigma_Y}{\sqrt{15}}\right)^2} && \text{(utilizando la ecuación 3.6)} \\
 &= \sqrt{4(0.05)^2 + \frac{4(0.08)^2}{15}} \\
 &= 0.11 \text{ m}
 \end{aligned}$$

El primer plan es mejor.

Mediciones repetidas con incertidumbres diferentes

A veces al repetir las mediciones se puede tener incertidumbres diferentes. Esto puede ocurrir, por ejemplo, cuando las mediciones se hacen con instrumentos diferentes. La mejor manera de combinar las mediciones en este caso es con un promedio ponderado, más que con la media de la muestra. En los ejemplos 3.11 y 3.12 se analiza esta idea.

Ejemplo

3.11

Un ingeniero mide el periodo de un péndulo (en segundos) de 2.0 ± 0.2 s. Se hizo otra medición independiente con un reloj más preciso y el resultado es de 2.2 ± 0.1 s. El promedio de estas dos mediciones es 2.1 s. Determine la incertidumbre en esta cantidad.

Solución

Sea X la medición con el reloj menos preciso, por lo que $X = 2.0$ s, con incertidumbre $\sigma_X = 0.2$ s. Sea Y la medición con el reloj más preciso, por lo que $Y = 2.2$ s, con incertidumbre $\sigma_Y = 0.1$ s. El promedio es $(1/2)X + (1/2)Y = 2.10$ y la incertidumbre en este promedio es

$$\begin{aligned}
 \sigma_{\text{prom}} &= \sqrt{\frac{1}{4}\sigma_X^2 + \frac{1}{4}\sigma_Y^2} \\
 &= \sqrt{\frac{1}{4}(0.2)^2 + \frac{1}{4}(0.1)^2} \\
 &= 0.11 \text{ s}
 \end{aligned}$$

Ejemplo

3.12

En el ejemplo 3.11, otro ingeniero sugiere que debido a que Y es una medición más precisa que X , podría ser más preciso un promedio ponderado en el cual Y fuera más pesado que X que el promedio no ponderado. Específicamente, el ingeniero sugiere que al elegir una constante adecuada c entre 0 y 1, el promedio ponderado $cX + (1 - c)Y$ podría tener una incertidumbre más pequeña que el promedio no ponderado $(1/2)X + (1/2)Y$ que se consideró en el ejemplo 3.11. Expresando la incertidumbre en el promedio ponderado $cX + (1 - c)Y$ en función de c se encuentra el valor de c que minimiza la incertidumbre.

Solución

La incertidumbre en el promedio ponderado es

$$\begin{aligned}\sigma &= \sqrt{c^2\sigma_X^2 + (1 - c)^2\sigma_Y^2} \\ &= \sqrt{0.04c^2 + 0.01(1 - c)^2} \\ &= \sqrt{0.05c^2 - 0.02c + 0.01}\end{aligned}$$

Ahora se debe encontrar el valor de c que minimiza a σ . Esto es equivalente a encontrar el valor de c que minimice a σ^2 . Al tomar la derivada de $\sigma^2 = 0.05c^2 - 0.02c + 0.01$ con respecto a c y se iguala a 0:

$$\frac{d\sigma^2}{dc} = 0.10c - 0.02 = 0$$

Despejando c , se obtiene

$$c = 0.2$$

Por tanto, el promedio ponderado más preciso es $0.2X + 0.8Y = 2.16$. La incertidumbre en esta estimación es

$$\sigma_{\text{mejor}} = \sqrt{(0.2)^2\sigma_X^2 + (0.8)^2\sigma_Y^2} = \sqrt{(0.2)^2(0.2)^2 + (0.8)^2(0.1)^2} = 0.09 \text{ s}$$

Observe que ésta es menor que la incertidumbre de 0.11 s que se encontró para el promedio no ponderado que se usó en el ejemplo 3.11.

La razón de los coeficientes X y Y en el mejor promedio ponderado es igual a la razón de las varianzas de Y y X : $\sigma_Y^2/\sigma_X^2 = 0.1^2/0.2^2 = 0.25 = 0.2/0.8 = c/(1 - c)$. Por tanto, se puede expresar los coeficientes en función de las varianzas: $c = \sigma_Y^2/(\sigma_X^2 + \sigma_Y^2) = 0.2$ y $1 - c = \sigma_X^2/(\sigma_X^2 + \sigma_Y^2) = 0.8$. Esta relación, en general, es válida.

Resumen

Si X y Y son mediciones *independientes* de la misma cantidad, con incertidumbres σ_X y σ_Y , respectivamente, entonces el promedio ponderado de X y Y con la incertidumbre más pequeña está dado $c_{\text{mejor}}X + (1 - c_{\text{mejor}})Y$ donde

$$c_{\text{mejor}} = \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} \quad 1 - c_{\text{mejor}} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2} \quad (3.7)$$

Combinaciones lineales de mediciones dependientes

Imagine que X y Y son mediciones con incertidumbres σ_X y σ_Y y se desea calcular la incertidumbre en la suma $X + Y$. Si X y Y son *dependientes*, la incertidumbre en la suma podría ser más grande o menor que en el caso independiente y no se puede determinar con sólo σ_X y σ_Y . Por ejemplo, si errores aleatorios positivos en X tienden a ocurrir al lado de errores aleatorios negativos en Y , y viceversa, aquéllos tenderán a contrarrestarse cuando se calcula la suma $X + Y$, así la incertidumbre en $X + Y$ será más pequeña que en el caso independiente. Por otro lado, si los errores aleatorios en X y Y tienden a tener el mismo signo, la incertidumbre en $X + Y$ será mayor que en el caso independiente.

La cantidad que mide la relación entre los errores aleatorios en X y Y es la covarianza, que se analizó en la sección 2.6. En general, si $X_1, \dots, X_n$ son mediciones y si se conoce la covarianza de cada par de mediciones, se puede utilizar la ecuación (2.72) (en la sección 2.6) para calcular la incertidumbre en una combinación lineal de las mediciones.

En la práctica, cuando las mediciones son dependientes, sucede con frecuencia el caso de que no se conoce lo suficiente acerca de la dependencia para cuantificarla. En estos casos se puede colocar un límite superior a la incertidumbre de una combinación lineal de las mediciones. Aquí se presenta el resultado; al final de esta sección se encuentra su demostración.

Si $X_1, \dots, X_n$ son mediciones y $c_1, \dots, c_n$ son constantes, entonces

$$\sigma_{c_1 X_1 + \dots + c_n X_n} \leq |c_1| \sigma_{X_1} + \dots + |c_n| \sigma_{X_n} \quad (3.8)$$

La expresión del lado derecho de la desigualdad (3.8) es una estimación conservadora de la incertidumbre en $c_1 X_1 + \dots + c_n X_n$.

Ejemplo 3.13

Un topógrafo está midiendo el **perímetro** de un terreno rectangular. Mide dos lados adyacentes de 50.11 ± 0.05 m y 75.21 ± 0.08 m. Estas mediciones no son necesariamente independientes. Determine con una estimación conservadora la incertidumbre del perímetro del terreno.

Solución

Sean X_1 y X_2 las dos mediciones. Las incertidumbres son entonces $\sigma_{X_1} = 0.05$ y $\sigma_{X_2} = 0.08$ y el perímetro está dado por $P = 2X_1 + 2X_2$. Utilizando la desigualdad (3.8), se obtiene

$$\begin{aligned} \sigma_P &= \sigma_{2X_1 + 2X_2} \\ &\leq 2\sigma_{X_1} + 2\sigma_{X_2} \\ &= 2(0.05) + 2(0.08) \\ &= 0.26 \text{ m} \end{aligned}$$

La incertidumbre en el perímetro no es mayor que 0.26 m. En el ejemplo 3.6 se calculó que la incertidumbre era de 0.19 m cuando X y Y son independientes.

Deducción de la desigualdad $\sigma_{c_1X_1+\dots+c_nX_n} \leq |c_1|\sigma_{X_1} + \dots + |c_n|\sigma_{X_n}$

Esta deducción requiere material de la sección 2.6. Sean $X_1, \dots, X_n$ variables aleatorias y sean $c_1, \dots, c_n$ constantes. Utilizando la ecuación (2.72) (de la sección 2.6),

$$\sigma_{c_1X_1+\dots+c_nX_n}^2 = c_1^2\sigma_{X_1}^2 + \dots + c_n^2\sigma_{X_n}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_i c_j \text{Cov}(X_i, X_j)$$

Ahora $\rho_{X_i, X_j} = \frac{\text{Cov}(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}$. Puesto que $|\rho_{X_i, X_j}| \leq 1$, por lo que

$$|\text{Cov}(X_i, X_j)| \leq \sigma_{X_i}\sigma_{X_j}$$

Ya que $c_i c_j \text{Cov}(X_i, X_j) \leq |c_i||c_j||\text{Cov}(X_i, X_j)|$, por lo que

$$c_i c_j \text{Cov}(X_i, X_j) \leq |c_i||c_j|\sigma_{X_i}\sigma_{X_j}$$

Sustituyendo, se obtiene

$$\sigma_{c_1X_1+\dots+c_nX_n}^2 \leq c_1^2\sigma_{X_1}^2 + \dots + c_n^2\sigma_{X_n}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n |c_i||c_j|\sigma_{X_i}\sigma_{X_j} \quad (3.9)$$

Puesto que $c_i^2 = |c_i|^2$, el lado derecho de la desigualdad (3.9) se puede factorizar:

$$c_1^2\sigma_{X_1}^2 + \dots + c_n^2\sigma_{X_n}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n |c_i||c_j|\sigma_{X_i}\sigma_{X_j} = (|c_1|\sigma_{X_1} + \dots + |c_n|\sigma_{X_n})^2$$

Sustituyendo en la desigualdad (3.9) y tomando raíces cuadradas, se obtiene

$$\sigma_{c_1X_1+\dots+c_nX_n} \leq |c_1|\sigma_{X_1} + \dots + |c_n|\sigma_{X_n}$$

Ejercicios de la sección 3.2

- Suponga que X y Y son mediciones independientes con incertidumbres $\sigma_X = 0.2$ y $\sigma_Y = 0.4$. Determine las incertidumbres en las cantidades siguientes:
 - $3X$
 - $X - Y$
 - $2X + 3Y$
- Una medición del diámetro de un cilindro tiene incertidumbre de 2 mm. ¿Cuántas mediciones se deben hacer para que se pueda calcular el diámetro con una incertidumbre de sólo 0.5 mm?
- En el artículo "The World's Longest Continued Series of Sea Level Observations" (M. Ekman en *Paleogeography*, 1988:73-77), la media del nivel anual de elevación de la tierra en Estocolmo, Suecia, se estimó de 4.93 ± 0.23 mm durante 1774-1884 y de 3.92 ± 0.19 mm durante 1885-1984. Estime la diferencia en la media anual de elevación entre estos dos espacios de tiempo y determine la incertidumbre en el cálculo.
- Un hueco cilíndrico es perforado con un bloque de acero y se fabrica un pistón cilíndrico que quepa en el hueco. El diámetro del hueco es 20.00 ± 0.01 cm y el diámetro del pistón es 19.90 ± 0.02 cm. La holgura es la mitad de la diferencia entre los diámetros. Estime la holgura y determine la incertidumbre en el cálculo.
- El ancho y la altura de una pieza de madera de 2×4 son en realidad de 1.5×3.5 pulg. Suponga que la incertidumbre en estas cantidades es despreciable. La longitud de una pieza de

2×4 se mide de 72 ± 0.1 pulg. Estime el volumen de la pieza de madera y determine la incertidumbre en la estimación.

6. El periodo T de un péndulo simple está dado por $T = 2\pi\sqrt{L/g}$ donde L es la longitud del péndulo y g la aceleración debido a la gravedad. Por tanto, si se miden L y T , se puede calcular g con $g = 4\pi^2 L/T^2$. Suponga que se sabe que el periodo es $T = 1.5$ s con incertidumbre despreciable y que se mide L de 0.559 ± 0.005 m. Estime g y determine la incertidumbre en la estimación.
7. La ley de Beer-Lambert relaciona la absorbancia A de una solución con concentración C de una especie en solución mediante $A = MLC$, donde L es la longitud de la trayectoria y M es el coeficiente de absortividad molar. Suponga que $C = 1.25$ mol/cm³ y $L = 1$ cm, ambos con incertidumbre despreciable y que $A = 1.30 \pm 0.05$. Estime M y determine la incertidumbre en la estimación.
8. En el flujo de Couette, dos placas planas grandes yacen una encima de otra, separándolas una fina capa de fluido. Si se aplica una fuerza cortante en la placa superior, la viscosidad del fluido también produce el movimiento de la placa inferior. La velocidad V en la placa superior en relación con la placa inferior está dada por $V = \tau h/\mu$, donde τ es la fuerza cortante aplicada a la placa superior, h es el espesor de la capa de fluido y μ es la viscosidad del fluido. Suponga que $\mu = 1.49$ Pa·s y $h = 10$ mm, ambos con incertidumbre despreciable.
 - a) Suponga que $\tau = 30.0 \pm 0.1$ Pa. Estime V y determine la incertidumbre en la estimación aproximada.
 - b) Si se desea estimar V con una incertidumbre de 0.2 mm/s, ¿cuál debe ser la incertidumbre en τ ?
9. De acuerdo con la ley del enfriamiento de Newton, la temperatura T de un cuerpo al tiempo t está dado por $T = T_a + (T_0 - T_a)e^{-kt}$, donde T_a es la temperatura ambiente, T_0 es la temperatura inicial y k es la constante de razón de enfriamiento. Para cierto tipo de recipiente de bebida, el valor de k se sabe que es 0.025 min^{-1} .
 - a) Suponga que $T_a = 36^\circ\text{F}$ exactamente que $T_0 = 72.0 \pm 0.5^\circ\text{F}$. Estime la temperatura T al tiempo $t = 10$ min y determine la incertidumbre en la estimación.
 - b) Suponga que $T_0 = 72^\circ\text{F}$ exactamente y que $T_a = 36.0 \pm 0.5^\circ\text{F}$. Estime la temperatura T al tiempo $t = 10$ min y determine la incertidumbre en la estimación.
10. En el artículo "Influence of Crack Width on Shear Behavior of SIFCON" (C. Fritz y H. Reinhardt, en *High Performance Fiber Reinforced Cement Composites: Proceedings of the International RILEM/ACI Workshop*, 1992), la máxima tensión de corte τ de un miembro de concreto agrietado está dada por $\tau = \tau_0(1 - kw)$, donde τ_0 es la tensión de corte máxima para una grieta de ancho cero, w es el ancho de la grieta en mm y k es una constante que se estima de los datos experimentales. Suponga que $k = 0.29 \pm 0.05 \text{ mm}^{-1}$. Dado que $\tau_0 = 50 \text{ MPa}$ y $w = 1.0$ mm, ambos con incertidumbre despreciable, estime τ y determine la incertidumbre en la estimación.
11. Se hacen nueve mediciones independientes de la longitud de una varilla. El promedio de las nueve mediciones es $\bar{X} = 5.238$ cm y la desviación estándar es $s = 0.081$ cm.
 - a) ¿Está la incertidumbre en el valor 5.238 cm más cerca de 0.009, 0.027 o 0.081 cm? Explique.
 - b) Otra varilla se mide una vez con el mismo proceso. La medición es 5.423 cm. ¿Está la incertidumbre de este valor más cerca de 0.009, 0.027 o 0.081 cm? Explique.
12. Cierta balanza tiene una incertidumbre de 3 g y un sesgo de 2 g.
 - a) Se hace una sola medición en esta balanza. ¿Cuáles son el sesgo y la incertidumbre en esta medición?
 - b) Se hacen cuatro mediciones independientes en esta balanza. ¿Cuáles son el sesgo y la incertidumbre en el promedio de estas mediciones?
 - c) Se hacen 400 mediciones independientes en esta balanza. ¿Cuáles son el sesgo y la incertidumbre en el promedio de estas mediciones?
 - d) Conforme se hacen más mediciones, ¿la incertidumbre se hace más pequeña, más grande o permanece constante?
 - e) Conforme se hacen más mediciones, ¿el sesgo se hace más pequeño, más grande o permanece constante?
13. El volumen de una roca se mide colocándola sobre un cilindro graduado parcialmente lleno de agua y se mide el aumento en el volumen. Se hacen ocho mediciones independientes. El promedio de las mediciones es 87.0 mL y la desviación estándar es 2.0 mL.
 - a) Estime el volumen de la roca y determine la incertidumbre en la estimación.
 - b) Se hacen ocho mediciones adicionales, hasta hacer un total de 16. ¿Cuál es la incertidumbre, aproximadamente, en el promedio de las 16 mediciones?
 - c) Aproximadamente ¿cuántas mediciones se necesitarían para reducir la incertidumbre a 0.4 mL?
14. Un estudiante mide la constante de un resorte k , se carga un resorte y se mide la elongación. (De acuerdo con la ley de

Hooke, si l es la carga y e la elongación, entonces $k = l/e$. Suponga que se hacen cinco mediciones independientes y los valores medidos de k (en N/m) son 36.4, 35.4, 38.6, 36.6 y 38.0.

- Estime la constante del resorte y determine la incertidumbre en la estimación.
- Determine un valor aproximado para la incertidumbre en el promedio de las diez mediciones.
- ¿Aproximadamente cuántas mediciones se deben hacer para reducir la incertidumbre a 0.3 N/m?
- Un segundo resorte, similar al primero, mide la constante del resorte una vez. El valor medido de k es 39.3. ¿Aproximadamente cuál es el valor de la incertidumbre?

15. Cierta proceso químico se opera diez veces a temperatura de 65°C y diez a 80°C. La producción en cada operación se midió como un porcentaje de un máximo teórico. Los datos se presentan en la tabla siguiente.

65°C	71.3	69.1	70.3	69.9	71.1	70.7	69.8	68.5	70.9	69.8
80°C	90.3	90.8	91.2	90.7	89.0	89.7	91.3	91.2	89.7	91.1

- Para cada temperatura, estime la media de la producción y determine la incertidumbre en la estimación.
- Estime la diferencia entre las medias de las producciones en las dos temperaturas y determine la incertidumbre en la estimación.

16. Se pesa un objeto cuatro veces y los resultados, en miligramos, son 234, 236, 233 y 229. Entonces el objeto se pesa cuatro veces en una balanza diferente y los resultados, en miligramos, son 236, 225, 245 y 240. Se usará el promedio de las ocho mediciones para estimar el peso. Alguien sugiere estimar la incertidumbre en este cálculo de la siguiente manera: calcule la desviación estándar de las ocho mediciones. Llame a esta cantidad s . La incertidumbre es entonces $s/\sqrt{8}$. ¿Es esto correcto? Explique.

17. La longitud de un componente se estima por medio de mediciones repetidas.

- Se hacen diez mediciones independientes con un instrumento cuya incertidumbre es 0.05 mm. Sea $\bar{X}$ el promedio de estas mediciones. Determine la incertidumbre en $\bar{X}$.
- Un nuevo dispositivo de medición, cuya incertidumbre es de 0.02 mm, está disponible. Se hacen cinco mediciones independientes con este dispositivo. Sea $\bar{Y}$ el promedio de estas mediciones. Determine la incertidumbre en $\bar{Y}$.
- Con el fin de disminuir la incertidumbre aún más, se decide combinar las estimaciones de $\bar{X}$ y $\bar{Y}$. Un ingeniero sugiere estimar la longitud con $(1/2)\bar{X} + (1/2)\bar{Y}$. Un segundo ingeniero argumenta que puesto que $\bar{X}$ está basado en diez mediciones mientras que $\bar{Y}$ está basado sólo en cinco, una mejor estimación es $(10/15)\bar{X} + (5/15)\bar{Y}$. Determine la incertidumbre en cada una de estas estimaciones. ¿Cuál es más pequeña?
- Determine el valor c de forma que el promedio ponderado $c\bar{X} + (1 - c)\bar{Y}$ tenga una incertidumbre mínima. Determine la incertidumbre de este promedio ponderado.

18. Las longitudes de dos componentes se medirán varias veces. La incertidumbre en cada medición de la longitud del primer componente es $\sigma_1 = 0.02$ cm y la incertidumbre en cada medición de la longitud del segundo componente es $\sigma_2 = 0.08$ cm. Sea $\bar{X}$ el promedio de las mediciones del primer componente y $\bar{Y}$ el promedio de las mediciones del segundo componente. La longitud total de los dos componentes se estimará con la cantidad $\bar{X} + \bar{Y}$.

- Determine la incertidumbre en la longitud total si el primer componente se mide cuatro veces y el segundo 12 veces.
- Determine la incertidumbre en la longitud total en función de n si el primer componente se mide n veces y el segundo componente se mide $16 - n$ veces.
- Determine la mejor manera de asignar 16 mediciones entre los componentes determinando el valor de n que minimiza la incertidumbre.

3.3 Incertidumbres para funciones de una medición

Los ejemplos que se han visto hasta ahora implican calcular incertidumbres en funciones lineales de mediciones. En muchos casos se desea estimar la incertidumbre de una función no lineal de una medición. Por ejemplo, si el radio R de un círculo mide 5.00 ± 0.01 cm, ¿cuál es la incertidumbre en el área A ? En términos estadísticos, se sabe que la desviación estándar σ_R es de 0.01 cm y se debe calcular la desviación estándar de A , donde A es la función de R dada por $A = \pi R^2$.

El tipo del problema que se desea resolver es: dada una variable aleatoria X , con desviación estándar conocida σ_X y dada una función $U = U(X)$, ¿cómo se calcula la desviación estándar σ_U ? Si U es una función lineal, los métodos de la sección 3.2 son aplicables. Si U no es lineal, aún se puede aproximar σ_U , multiplicando σ_X por el valor absoluto de la derivada dU/dX . La aproximación será buena en tanto σ_X sea pequeña.

Si X es una medida cuya incertidumbre σ_X es pequeña y si U es una función de X , entonces

$$\sigma_U \approx \left| \frac{dU}{dX} \right| \sigma_X \quad (3.10)$$

En la práctica, se evalúa la derivada dU/dX en la medición observada X .

La ecuación (3.10) se conoce como la fórmula de la **propagación de error**. Al final de esta sección se proporciona su demostración.

Las incertidumbres de la propagación de errores están sólo aproximadas

Las incertidumbres calculadas utilizando la ecuación (3.10) con frecuencia son sólo simples aproximaciones. Por esta razón, estas incertidumbres se deben expresar con no más de dos dígitos significativos. Efectivamente, algunos autores sugieren utilizar un solo dígito significativo.

Las funciones no lineales están sesgadas

Si X es una medición no sesgada de un valor real μ_X y si la función $U = U(X)$ es una función no lineal de X , entonces en la mayoría de los casos U será una estimación *sesgada* del valor real $U(\mu_X)$. En la práctica este sesgamiento generalmente se desprecia. Se puede demostrar con métodos avanzados que el tamaño del sesgamiento depende de manera fundamental de las magnitudes de σ_X y de la segunda derivada d^2U/dX^2 . Por tanto, conforme la incertidumbre σ_X sea pequeña, el sesgamiento en U en general también lo será, excepto en algunas circunstancias bastante inusuales cuando la segunda derivada es muy grande. Por supuesto, si X es una medición con sesgamiento no despreciable, entonces el sesgamiento en U podría ser grande. Estas ideas se analizan aún más en el ejercicio 22 de los ejercicios adicionales al final de este capítulo.

Ejemplo 3.14

El radio R de un círculo mide 5.00 ± 0.01 cm. Estime el área del círculo y determine la incertidumbre.

Solución

El área A está dada por $A = \pi R^2$. La estimación aproximada de A es $\pi(5.00 \text{ cm})^2 = 78.5 \text{ cm}^2$. Ahora, $\sigma_R = 0.01$ cm y $dA/dR = 2\pi R = 10\pi$ cm. La incertidumbre en A se puede encontrar como:

$$\begin{aligned}\sigma_A &= \left| \frac{dA}{dR} \right| \sigma_R \\ &= (10\pi \text{ cm})(0.01 \text{ cm}) \\ &= 0.31 \text{ cm}^2\end{aligned}$$

Se estima el área del círculo de $78.5 \pm 0.3 \text{ cm}^2$.

Ejemplo 3.15

Una roca que se identifica como cuarcita tamaño adoquín tiene una masa m de 674.0 g. Suponga que esta medición tiene una incertidumbre despreciable. El volumen V de la roca se medirá al colocarla en un cilindro graduado parcialmente lleno de agua y medir el volumen de agua desplazada. La densidad D de la roca se calculará como $D = m/V$. Suponga que el volumen del agua desplazada es 261.0 ± 0.1 mL. Estime la densidad de la roca y determine la incertidumbre.

Solución

Sustituyendo $V = 261.0$ mL, la estimación de la densidad D es $674.0/261.0 = 2.582$ g/mL. Tratando a $m = 674.0$ como una constante conocida, $dD/dV = -674.0/V^2 = -674.0/(261.0)^2 = -0.010$ g/mL². Se sabe que $\sigma_V = 0.1$ mL. La incertidumbre en D es, por tanto,

$$\begin{aligned}\sigma_D &= \left| \frac{dD}{dV} \right| \sigma_V \\ &= |-0.010|(0.1 \text{ g/mL}) \\ &= 0.001 \text{ g/mL}\end{aligned}$$

Se estima la densidad de 2.582 ± 0.001 g/mL.

Incertidumbres relativas para funciones de una medición

Se ha estado llamando a la desviación estándar σ_U de una medición U como la incertidumbre en U . Un nombre más completo para σ_U es la **incertidumbre absoluta**, ya que se expresa con las mismas unidades que la medición U . A veces se desea expresar la incertidumbre como una fracción del valor real, la cual (suponiendo que no hay sesgamiento) representa la media de la medición μ_U . De hecho se le llama **incertidumbre relativa** en U . Esta última también se denomina **coeficiente de variación**. En la práctica, puesto que μ_U no se conoce, el valor medido de U se usa en su lugar cuando se calcula la incertidumbre relativa.

Resumen

Si U es una medición cuyo valor real es μ_U y cuya incertidumbre es σ_U , entonces la incertidumbre relativa en U es la cantidad σ_U/μ_U .

La respectiva incertidumbre es un número puro, sin unidades. Con frecuencia se expresa como porcentaje. En la práctica μ_U no se conoce, por lo que si el sesgamiento es despreciable, se estima la respectiva incertidumbre con σ_U/U .

Hay dos maneras de calcular la incertidumbre relativa en una cantidad U . Simplemente con sólo utilizar la ecuación (3.10) para calcular la incertidumbre absoluta σ_U y dividir entre U . Para desarrollar el segundo método, se calculará la incertidumbre absoluta en $\ln U$:

$$\sigma_{\ln U} = \frac{d(\ln U)}{dU} \sigma_U = \frac{\sigma_U}{U}$$

Esta ecuación muestra que la incertidumbre absoluta en $\ln U$ es igual a la incertidumbre relativa en U . Por tanto, la segunda manera de calcular la incertidumbre relativa en U es calcular $\ln U$ y después utilizar la ecuación (3.10) para calcular la incertidumbre absoluta en $\ln U$.

Resumen

Hay dos métodos para aproximar la incertidumbre relativa σ_U/U de una función $U = U(X)$:

1. Calcule σ_U utilizando la ecuación (3.10) y después divida entre U .
2. Calcule $\ln U$ y use la ecuación (3.10) para encontrar $\sigma_{\ln U}$, que es igual a σ_U/U .

Ambos métodos funcionan en cada ejemplo, por lo que se puede utilizar el que sea más fácil para un problema específico. Esta elección generalmente se determina por lo que sea más fácil de calcular, la derivada de U o de $\ln U$.

Ejemplo

3.16

El radio de un círculo mide 5.00 ± 0.01 cm. Estime el área y determine la incertidumbre relativa.

Solución

En el ejemplo 3.14, el área $A = \pi R^2$ se estimó de 78.5 ± 0.3 cm². La incertidumbre absoluta es, por tanto, $\sigma_A = 0.3$ cm² y la incertidumbre relativa es $\sigma_A/A = 0.3/78.5 = 0.004$. Por tanto, se puede expresar al área como $A = 78.5$ cm² ± 0.4 por ciento.

Si no se hubiese ya calculado σ_A , sería más fácil calcular la incertidumbre relativa al calcular la incertidumbre absoluta en $\ln A$. Puesto que $\ln A = \ln \pi + 2 \ln R$, $d \ln A/dR = 2/R = 0.4$. La incertidumbre relativa en A es

$$\begin{aligned}
 \frac{\sigma_A}{A} &= \sigma_{\ln A} \\
 &= \left| \frac{d \ln A}{dR} \right| \sigma_R \\
 &= 0.4 \sigma_R \\
 &= (0.4)(0.01) \\
 &= 0.4\%
 \end{aligned}$$

Ejemplo 3.17

La aceleración de una masa que desciende en un plano inclinado sin fricción está dada por $a = g \sin \theta$, donde g es la aceleración debida a la gravedad y θ es el ángulo de inclinación del plano. Suponga que la incertidumbre en g es despreciable. Si $\theta = 0.60 \pm 0.01$ rad, determine la incertidumbre relativa en a .

Solución

La incertidumbre relativa en a es la incertidumbre absoluta en $\ln a$. Ahora $\ln a = \ln g + \ln(\sin \theta)$, donde $\ln g$ es constante. Por tanto, $d \ln a / d\theta = d \ln(\sin \theta) / d\theta = \cos \theta / \sin \theta = \cot \theta = \cot(0.60) = 1.46$. La incertidumbre en θ es $\sigma_\theta = 0.01$. La incertidumbre relativa en a es, por tanto,

$$\begin{aligned}
 \frac{\sigma_a}{a} &= \sigma_{\ln a} \\
 &= \left| \frac{d \ln a}{d\theta} \right| \sigma_\theta \\
 &= (1.46)(0.01) \\
 &= 1.5\%
 \end{aligned}$$

Observe que la incertidumbre relativa en $a = g \sin \theta$ no depende de la constante g .

Deducción de la fórmula de propagación de errores

Se deduce la fórmula de la propagación de errores para una función no lineal U de una variable aleatoria X aproximándola con una función lineal y después utilizando los métodos de la sección 3.2. Para encontrar una aproximación lineal a U , se usa una serie de Taylor aproximada a primer orden. Esto se conoce como **linealizando el problema**; ésta es una técnica comúnmente usada en ciencia e ingeniería.

Sea $U(X)$ una función derivable. Sea μ_X cualquier punto. Entonces si X está cerca de μ_X , la serie de Taylor aproximada a primer orden para $U(X)$ es

$$U(X) - U(\mu_X) \approx \frac{dU}{dX}(X - \mu_X) \quad (3.11)$$

La derivada dU/dX se evalúa en μ_X .

Ahora sea X una medición y sea $U(X)$ (a la que se llamará U) una cantidad calculada a partir de X . Sea μ_X la media de X . Para cualquier medición razonablemente precisa, X estará bastante cerca de μ_X para que la aproximación en serie de Taylor sea válida.

Sumando $U(\mu_X)$ en ambos lados de la ecuación (3.11) se obtiene

$$U \approx U(\mu_X) + \frac{dU}{dX}(X - \mu_X)$$

Multiplicando por dU/dX y arreglando los términos

$$U \approx \left(U(\mu_X) - \frac{dU}{dX}\mu_X \right) + \frac{dU}{dX}X$$

Ahora la cantidad dU/dX es una constante, ya que ésta se encuentra evaluada en μ_X . Por tanto, la cantidad $U(\mu_X) - (dU/dX)\mu_X$ es también constante. Por tanto, de la ecuación (2.46) (en la sección 2.5) se tiene que

$$\sigma_U = \left| \frac{dU}{dX} \right| \sigma_X$$

Ésta es la fórmula de **propagación de errores**. Cuando se aplica, se evalúa la derivada dU/dX en la medición observada X , puesto que no se conoce el valor de μ_X .

Ejercicios de la sección 3.3

- Determine la incertidumbre en Y , dado que $X = 4.0 \pm 0.4$ y $Y = 5X$ exactamente de 5 m y que el tiempo medido es $t = 1.01 \pm 0.02$ s. Estime g y encuentre la incertidumbre en la estimación. (Observe que $g = 2s/t^2$.)
 - $Y = X^2$
 - $Y = \sqrt{X}$
 - $Y = 1/X$
 - $Y = \ln X$
 - $Y = e^X$
 - $Y = \sin X$ (X está en unidades de radianes)
- Dado que X y Y están relacionados por la ecuación dada y que $X = 2.0 \pm 0.2$, estime Y y determine la incertidumbre en la estimación.
 - $XY = 1$
 - $X/Y = 2$
 - $X\sqrt{Y} = 3$
 - $Y\sqrt{X} = 4$
- La aceleración g debida a la gravedad se estima dejando caer un objeto y se mide el tiempo que le toma recorrer cierta distancia. Suponga que se sabe que la distancia s es exacta-
 - Suponga que $g = 9.80 \text{ m/s}^2$ exactamente y que $L = 0.742 \pm 0.005$ m. Estime T y determine la incertidumbre en la estimación.
 - Suponga que $L = 0.742$ m exactamente y que $T = 1.73 \pm 0.01$ s. Estime g y determine la incertidumbre en la estimación.
- La altura h de aumento en el vaso capilar del agua en un tubo de vidrio limpio está dada por $h = k/r$, donde r es el radio del tubo y k es la constante que depende de la temperatura

del agua. Suponga que a una temperatura de 10°C , $k = 7.57 \text{ mm}^2$. Suponga que el radio del tubo es $2.0 \pm 0.1 \text{ mm}$. Estime h y determine la incertidumbre en la estimación.

7. La velocidad de fricción F del agua que fluye a través de un tubo está dada por $F = \sqrt{gdh/4l}$, donde g representa la aceleración debida a la gravedad, d el diámetro del tubo, l la longitud del tubo y h la pérdida de carga. Estime F y determine la incertidumbre en la estimación, suponga que $g = 9.80 \text{ m/s}^2$ exactamente y que
 - a) $d = 0.15 \text{ m}$ y $l = 30.0 \text{ m}$, ambos con incertidumbre despreciable y $h = 5.33 \pm 0.02 \text{ m}$.
 - b) $h = 5.33 \text{ m}$ y $l = 30.0 \text{ m}$, ambos con incertidumbre despreciable y $d = 0.15 \pm 0.03 \text{ m}$.
 - c) $d = 0.15 \text{ m}$ y $h = 5.33 \text{ m}$, ambos con incertidumbre despreciable y $l = 30.00 \pm 0.04 \text{ m}$.
8. El índice de refracción n de una pieza de vidrio está relacionado con el ángulo crítico θ por $n = 1/\sin\theta$. Suponga que el ángulo crítico se ha medido de $0.70 \pm 0.02 \text{ rad}$. Estime el índice de refracción y determine la incertidumbre en la estimación.
9. La densidad de una roca será medida colocándola en un cilindro graduado parcialmente lleno de agua y después se mide el volumen de agua desplazado. La densidad D está dada por $D = m/(V_1 - V_0)$, donde m es la masa de la roca, V_0 es el volumen inicial del agua y V_1 es el volumen del agua más la roca. Suponga que la masa de la roca es de 750 g , con incertidumbre despreciable y que $V_0 = 500.0 \pm 0.1 \text{ mL}$ y $V_1 = 813.2 \pm 0.1 \text{ mL}$. Estime la densidad de la roca y determine la incertidumbre en la estimación.
10. La conversión de cianuro de amonio en urea es una reacción de segundo orden. Esto significa que la concentración C de cianuro de amonio al tiempo t está dado por $1/C = kt + 1/C_0$, donde C_0 es la concentración inicial y que k es la razón constante. Suponga que se sabe que la concentración inicial es de 0.1 mol/L exactamente. Suponga que el tiempo se puede medir con incertidumbre despreciable.
 - a) Después de 45 minutos se mide la concentración de cianuro de amonio de $0.0811 \pm 0.0005 \text{ mol/L}$. Estime la razón constante k y determine la incertidumbre en la estimación.
 - b) Use el resultado del inciso a) para estimar el tiempo en el que la concentración de cianuro de amonio será de 0.0750 mol/L y determine la incertidumbre en la estimación.
11. Convierta las siguientes incertidumbres absolutas en incertidumbres relativas.
 - a) 37.2 ± 0.1
 - b) 8.040 ± 0.003
 - c) 936 ± 37
 - d) 54.8 ± 0.3
12. Convierta las siguientes incertidumbres relativas en incertidumbres absolutas.
 - a) $48.41 \pm 0.3\%$
 - b) $991.7 \pm 0.6\%$
 - c) $0.011 \pm 0.9\%$
 - d) $7.86 \pm 1\%$
13. La aceleración g debida a la gravedad se estima dejando caer un objeto y se mide el tiempo que le toma recorrer cierta distancia. Suponga que se conoce que la distancia s es exactamente 2.2 m . El tiempo que se mide es $t = 0.67 \pm 0.02 \text{ s}$. Estime g y determine la incertidumbre relativa en la estimación. (Observe que $g = 2s/t^2$.)
14. Con referencia al ejercicio 4, suponga que $T = 298.4 \pm 0.2 \text{ K}$. Estime V y determine la incertidumbre relativa en la estimación.
15. Con referencia al ejercicio 5.
 - a) Suponga que $g = 9.80 \text{ m/s}^2$ exactamente y que $L = 0.855 \pm 0.005 \text{ m}$. Estime T y determine la incertidumbre relativa en la estimación.
 - b) Suponga que $L = 0.855 \text{ m}$ exactamente y que $T = 1.856 \pm 0.005 \text{ s}$. Estime g y determine la incertidumbre relativa en la estimación.
16. Con referencia al ejercicio 6, suponga que el radio del tubo es $r = 2.5 \pm 0.2 \text{ mm}$ y que $k = 7.57 \text{ mm}^2$ con incertidumbre despreciable. Estime h y determine la incertidumbre relativa en la estimación.
17. Con referencia al ejercicio 7, estime F y determine la incertidumbre relativa en la estimación, suponga que $g = 9.80 \text{ m/s}^2$ exactamente y que
 - a) $d = 0.20 \text{ m}$ y $l = 35.0 \text{ m}$, ambos con incertidumbre despreciable y $h = 4.51 \pm 0.03 \text{ m}$.
 - b) $h = 4.51 \text{ m}$ y $l = 35.0 \text{ m}$, ambos con incertidumbre despreciable y $d = 0.20 \pm 0.008 \text{ m}$.
 - c) $d = 0.20 \text{ m}$ y $h = 4.51 \text{ m}$, ambos con incertidumbre despreciable y $l = 35.00 \pm 0.4 \text{ m}$.
18. Con referencia al ejercicio 8, suponga que el ángulo crítico se midió de $0.90 \pm 0.01 \text{ rad}$. Estime el índice de refracción y determine la incertidumbre relativa en la estimación.

19. Con referencia al ejercicio 9, suponga que la masa de la roca es 288.2 g con incertidumbre despreciable, el volumen inicial del agua en el cilindro es $400 \pm 0.1 \text{ mL}$ y el volumen de agua más la roca es $516 \pm 0.2 \text{ mL}$. Estime la densidad de la roca y determine la incertidumbre relativa en la estimación.
20. En una operación de una reacción química a cierta temperatura, la concentración C de cierto reactivo al tiempo t está dada por $1/C = kt + 1/C_0$, donde C_0 es la concentración inicial y k es la razón constante. Suponga que se sabe que la concentración inicial es 0.04 mol/L exactamente. Suponga que el tiempo se mide con incertidumbre despreciable.
- a) Después de 30 s , la concentración C se mide de $0.0038 \pm 2.0 \times 10^{-4} \text{ mol/L}$. Estime la razón constante k y determine la incertidumbre relativa en la estimación.
- b) Después de 50 s , la concentración C se mide de $0.0024 \pm 2.0 \times 10^{-4} \text{ mol/L}$. Estime la razón constante k y determine la incertidumbre relativa en la estimación.
- c) Denote las estimaciones de la razón constante k en los incisos a) y b) por $\hat{k}_1$ y $\hat{k}_2$, respectivamente. La media geométrica $\sqrt{\hat{k}_1 \hat{k}_2}$ se usa como estimación de k . Determine la respectiva incertidumbre en la estimación.

3.4 Incertidumbres para funciones de varias mediciones

Con frecuencia se necesita estimar una cantidad como una función de varias mediciones. Por ejemplo, se podría medir la masa m y el volumen V de una roca y calcular la densidad como $D = m/V$. En el ejemplo 3.15 se mostró cómo estimar la incertidumbre en D cuando una de las cantidades, en este caso V , fue medida con incertidumbre mientras que m se trató como una constante conocida. Sin embargo, en la práctica se podría necesitar estimar la incertidumbre en D cuando tanto m como V se miden con incertidumbre.

En esta sección se aprenderá cómo estimar la incertidumbre en una cantidad que es una función de varias mediciones inciertas independientes. La fórmula básica está dada aquí.

Si $X_1, X_2, \dots, X_n$ son mediciones independientes cuyas incertidumbres $\sigma_{X_1}, \sigma_{X_2}, \dots, \sigma_{X_n}$ son pequeñas y si $U = U(X_1, X_2, \dots, X_n)$ es una función de $X_1, X_2, \dots, X_n$, entonces

$$\sigma_U \approx \sqrt{\left(\frac{\partial U}{\partial X_1}\right)^2 \sigma_{X_1}^2 + \left(\frac{\partial U}{\partial X_2}\right)^2 \sigma_{X_2}^2 + \dots + \left(\frac{\partial U}{\partial X_n}\right)^2 \sigma_{X_n}^2} \quad (3.12)$$

En la práctica, se evalúan las derivadas parciales en el punto $(X_1, X_2, \dots, X_n)$.

La ecuación (3.12) representa la **fórmula de propagación de errores multivariada**. Es importante observar que es válida sólo cuando las mediciones $X_1, X_2, \dots, X_n$ son independientes. Una deducción de la fórmula se proporciona al final de la sección. Como en el caso de una medición, las incertidumbres calculadas con la fórmula de la propagación de errores son frecuentemente aproximaciones.

Las funciones no lineales de las mediciones son, en general, sesgadas (véase el análisis con respecto a funciones de una medición en la sección 3.3). Sin embargo, en tanto las mediciones $X_1, \dots, X_n$ no estén sesgadas y las incertidumbres $\sigma_{X_1}, \dots, \sigma_{X_n}$ sean todas pequeñas, el sesgamiento en U por lo regular será lo suficientemente pequeño para despreciarlo. Excepciones a esta regla son bastante inusuales, pueden ocurrir cuando algunas de las derivadas par-

ciales de segundo o más alto orden de U con respecto a X_i son bastante grandes. Por supuesto, si una o más de las $X_1, \dots, X_n$ son considerablemente sesgadas, entonces U podría también estarlo. Estas ideas se exploran aún más en el ejercicio 23 de los ejercicios adicionales al final de este capítulo.

Ahora se presentan algunos ejemplos que ilustran el uso de la propagación de errores multivariada.

Ejemplo

3.18

Suponga que la masa de una roca se mide de $m = 674.0 \pm 1.0$ g y el volumen de la roca se mide de $V = 261.0 \pm 0.1$ mL. Estime la densidad de la roca y determine la incertidumbre en el cálculo en la estimación.

Solución

Sustituyendo $m = 674.0$ g y $V = 261.0$ mL, la estimación de la densidad D es $674.0/261.0 = 2.582$ g/mL. Puesto que $D = m/V$, las derivadas parciales de D son

$$\begin{aligned}\frac{\partial D}{\partial m} &= \frac{1}{V} = 0.0038 \text{ mL}^{-1} \\ \frac{\partial D}{\partial V} &= \frac{-m}{V^2} = -0.0099 \text{ g/mL}^2\end{aligned}$$

La incertidumbre en D es, por tanto,

$$\begin{aligned}\sigma_D &= \sqrt{\left(\frac{\partial D}{\partial m}\right)^2 \sigma_m^2 + \left(\frac{\partial D}{\partial V}\right)^2 \sigma_V^2} \\ &= \sqrt{(0.0038)^2 (1.0)^2 + (-0.0099)^2 (0.1)^2} \\ &= 0.0040 \text{ g/mL}\end{aligned}$$

La densidad de la roca es 2.582 ± 0.004 g/mL.

Uno de los beneficios de la fórmula de propagación de errores multivariada es que permite determinar las mediciones que contribuyen más a la incertidumbre en el resultado final. El ejemplo 3.19 ilustra esto último.

Ejemplo

3.19

La densidad de la roca en el ejemplo 3.18 se estimó otra vez con equipo diferente, con el fin de mejorar la precisión. ¿Qué mejoraría más la precisión del cálculo de la densidad: reducir la incertidumbre en el cálculo de la masa a 0.5 g o a 0.05 mL?

Solución

Del ejemplo 3.18, $\sigma_D = \sqrt{(0.0038)^2 \sigma_m^2 + (-0.0099)^2 \sigma_V^2}$. Se ha hecho una elección entre tener $\sigma_m = 0.5$ y $\sigma_V = 0.1$, o tener $\sigma_m = 1.0$ y $\sigma_V = 0.05$. Los resultados de la primera elección en $\sigma_D = 0.002$ g/mL, mientras que en la segunda elección en $\sigma_D = 0.004$ g/mL. Es mejor reducir σ_m a 0.5 g.

Ejemplo

3.20

Dos resistores con resistencias R_1 y R_2 están conectados en paralelo. La resistencia combinada R está dada por $R = (R_1 R_2)/(R_1 + R_2)$. Si R_1 mide $100 \pm 10 \, \Omega$ y R_2 $20 \pm 1 \, \Omega$, estime R y determine la incertidumbre en la estimación.

Solución

La estimación de R es $(100)(20)/(100 + 20) = 16.67 \, \Omega$. Para calcular σ_R , primero se calculan las derivadas parciales de R :

$$\frac{\partial R}{\partial R_1} = \left(\frac{R_2}{R_1 + R_2} \right)^2 = 0.0278$$

$$\frac{\partial R}{\partial R_2} = \left(\frac{R_1}{R_1 + R_2} \right)^2 = 0.694$$

Ahora $\sigma_{R_1} = 10 \, \Omega$ y $\sigma_{R_2} = 1 \, \Omega$. Por tanto,

$$\begin{aligned} \sigma_R &= \sqrt{\left(\frac{\partial R}{\partial R_1} \right)^2 \sigma_{R_1}^2 + \left(\frac{\partial R}{\partial R_2} \right)^2 \sigma_{R_2}^2} \\ &= \sqrt{(0.0278)^2 (10)^2 + (0.694)^2 (1)^2} \\ &= 0.75 \, \Omega \end{aligned}$$

La resistencia combinada es $16.67 \pm 0.75 \, \Omega$.

Ejemplo

3.21

En el ejemplo 3.20, el resistor de $100 \pm 10 \, \Omega$ se puede reemplazar con uno más costoso de $100 \pm 1 \, \Omega$. ¿En cuánto se reduciría la incertidumbre en la resistencia combinada por esto último? ¿Es útil hacer el reemplazo?

Solución

Utilizando el método del ejemplo 3.20, la incertidumbre en la resistencia combinada R con el nuevo resistor sería

$$\sqrt{(0.0278)^2 (1)^2 + (0.694)^2 (1)^2} = 0.69 \, \Omega$$

No hay mucha reducción de la incertidumbre de $0.75 \, \Omega$ utilizando el resistor viejo. Casi toda la incertidumbre en la resistencia combinada se debe a la incertidumbre en el resistor de $20 \, \Omega$. En este contexto, la que se encuentra en el resistor de $100 \, \Omega$ se puede despreciar para la mayoría de los propósitos prácticos. Hay muy poco beneficio en reemplazar este resistor.

Observe que en el ejemplo 3.20, un componente (el resistor de $100 \, \Omega$) tenía una incertidumbre más grande, tanto en términos absolutos como relativos del valor medido, que el otro. Aún así, el ejemplo 3.21 mostraba que la incertidumbre en la resistencia combinada estaba sólo ligeramente afectada por la incertidumbre en este componente. La lección es que no se

puede predecir el impacto de las incertidumbres en las mediciones individuales sobre la incertidumbre en el cálculo final de las magnitudes de las incertidumbres solas. Se debe utilizar la fórmula de propagación de errores.

Incertidumbres para funciones de mediciones dependientes

Si $X_1, X_2, \dots, X_n$ no son independientes, la incertidumbre en una función $U = U(X_1, X_2, \dots, X_n)$ se puede estimar si se conoce la covarianza de cada par (X_i, X_j) . (La covarianza se analiza en la sección 2.6.) En muchas situaciones, las covarianzas no se conocen. En estos casos, se puede hacer una estimación conservadora de la incertidumbre en U . Aquí se presenta este resultado.

Si $X_1, X_2, \dots, X_n$ son mediciones cuyas incertidumbres $\sigma_{X_1}, \sigma_{X_2}, \dots, \sigma_{X_n}$, son pequeñas y si $U = U(X_1, X_2, \dots, X_n)$ es una función de $(X_1, X_2, \dots, X_n)$, entonces una estimación conservadora de σ_U está dada por

$$\sigma_U \leq \left| \frac{\partial U}{\partial X_1} \right| \sigma_{X_1} + \left| \frac{\partial U}{\partial X_2} \right| \sigma_{X_2} + \dots + \left| \frac{\partial U}{\partial X_n} \right| \sigma_{X_n} \quad (3.13)$$

En la práctica se evalúan las derivadas parciales en el punto $(X_1, X_2, \dots, X_n)$.

La desigualdad (3.13) es válida en casi todas las situaciones prácticas; en principio puede fallar si algunas de las segundas derivadas parciales de U son muy grandes.

Ejemplo

3.22

Con referencia al ejemplo 3.20, determine una estimación conservadora para la incertidumbre en la resistencia R total si no se sabe si R_1 y R_2 son independientes.

Solución

Se tiene $\sigma_{R_1} = 10 \, \Omega$, $\sigma_{R_2} = 1 \, \Omega$, $\partial R / \partial R_1 = 0.0278$, y $\partial R / \partial R_2 = 0.694$. Por tanto,

$$\begin{aligned} \sigma_R &\leq \left| \frac{\partial R}{\partial R_1} \right| \sigma_{R_1} + \left| \frac{\partial R}{\partial R_2} \right| \sigma_{R_2} \\ &= (0.0278)(10) + (0.694)(1) \\ &= 0.97 \, \Omega \end{aligned}$$

La incertidumbre en la resistencia total se estimó conservadoramente de $0.97 \, \Omega$. En el ejemplo 3.20 se calculó la incertidumbre de $0.75 \, \Omega$ cuando R_1 y R_2 son independientes.

Incertidumbres relativas para funciones de varias mediciones

En la sección 3.3 se presentaron los métodos para calcular las incertidumbres relativas para funciones de una variable. Los métodos de cálculo de incertidumbres relativas para las funciones de varias variables son similares.

Hay dos métodos para aproximar la incertidumbre relativa σ_U/U en una función $U = U(X_1, X_2, \dots, X_n)$:

1. Calcule σ_U utilizando la ecuación (3.12) y después divida entre U .
2. Calcule $\ln U$ y después use la ecuación (3.12) para encontrar $\sigma_{\ln U}$ que es igual a σ_U/U .

Ambos métodos funcionan en cada ejemplo, por tanto se puede utilizar cualquiera de los dos, el que sea más fácil, para un problema específico. Esta elección se determina generalmente por lo que sea más fácil para calcular las derivadas parciales de U o de $\ln U$.

Ejemplo 3.23

Dos lados perpendiculares de un rectángulo se miden de $X = 2.0 \pm 0.1$ cm y $Y = 3.2 \pm 0.2$ cm. Determine la incertidumbre relativa en el área $A = XY$.

Solución

La incertidumbre relativa se calcula fácilmente encontrando la incertidumbre absoluta en $\ln A = \ln X + \ln Y$. Se inicia calculando las derivadas parciales de $\ln A$:

$$\frac{\partial \ln A}{\partial X} = \frac{1}{X} = 0.50 \quad \frac{\partial \ln A}{\partial Y} = \frac{1}{Y} = 0.31$$

Se ha dado que $\sigma_X = 0.1$ y $\sigma_Y = 0.2$. La incertidumbre relativa en A es

$$\begin{aligned} \frac{\sigma_A}{A} = \sigma_{\ln A} &= \sqrt{\left(\frac{\partial \ln A}{\partial X}\right)^2 \sigma_X^2 + \left(\frac{\partial \ln A}{\partial Y}\right)^2 \sigma_Y^2} \\ &= \sqrt{(0.50)^2 (0.1)^2 + (0.31)^2 (0.2)^2} \\ &= 0.080 \end{aligned}$$

La incertidumbre relativa en A es 0.080 u 8%. El área del rectángulo es $6.4 \text{ cm}^2 \pm 8\%$.

Ejemplo 3.24

Una máquina de Atwood consta de dos masas X y Y ($X > Y$) fijas en los extremos de una cuerda ligera que pasa por una polea ligera sin fricción. Cuando se sueltan las masas, la más grande X se acelera hacia abajo con una aceleración

$$a = g \frac{X - Y}{X + Y}$$

Suponga que X y Y se miden como $X = 100 \pm 1$ g y $Y = 50 \pm 1$ g. Suponga que g , la aceleración debida a la gravedad, es conocida con incertidumbre despreciable. Encuentre la incertidumbre relativa en la aceleración a .

Solución

La incertidumbre relativa en a es igual a la incertidumbre absoluta en $\ln a = \ln g + \ln(X - Y) - \ln(X + Y)$. Tratamos a g como una constante, ya que su incertidumbre es despreciable. Las derivadas parciales son

$$\frac{\partial \ln a}{\partial X} = \frac{1}{X - Y} - \frac{1}{X + Y} = 0.0133$$

$$\frac{\partial \ln a}{\partial Y} = -\frac{1}{X - Y} - \frac{1}{X + Y} = -0.0267$$

Las incertidumbres en X y Y son $\sigma_X = \sigma_Y = 1$. La incertidumbre relativa en a es

$$\begin{aligned} \frac{\sigma_a}{a} = \sigma_{\ln a} &= \sqrt{\left(\frac{\partial \ln a}{\partial X}\right)^2 \sigma_X^2 + \left(\frac{\partial \ln a}{\partial Y}\right)^2 \sigma_Y^2} \\ &= \sqrt{(0.0133)^2(1)^2 + (-0.0267)^2(1)^2} \\ &= 0.030 \end{aligned}$$

La incertidumbre relativa en a es 0.030, o 3%. Observe que este valor no depende de g .

Deducción de la fórmula de propagación de errores multivariada

Se deduce la fórmula de propagación de errores para una función no lineal U de una variable aleatoria X al aproximar a la función con una función lineal multivariada (es decir, al linealizar el problema) y después utilizando los métodos de la sección 3.2. Para encontrar una aproximación lineal para U se usa una aproximación con una serie de Taylor de primer orden. Sea $U = U(X_1, X_2, \dots, X_n)$ una función cuyas derivadas parciales existen. Sea $(\mu_1, \mu_2, \dots, \mu_n)$ un punto. Entonces si $X_1, X_2, \dots, X_n$ son cercanas a $\mu_1, \mu_2, \dots, \mu_n$ respectivamente, la linealización de U es

$$\begin{aligned} U(X_1, X_2, \dots, X_n) - U(\mu_1, \mu_2, \dots, \mu_n) &\approx \frac{\partial U}{\partial X_1}(X_1 - \mu_1) \\ &+ \frac{\partial U}{\partial X_2}(X_2 - \mu_2) + \dots + \frac{\partial U}{\partial X_n}(X_n - \mu_n) \end{aligned}$$

Cada derivada parcial se evalúa en el punto $(\mu_1, \mu_2, \dots, \mu_n)$.

Si $X_1, X_2, \dots, X_n$ son mediciones *independientes*, la aproximación lineal conduce a un método para aproximar la incertidumbre en U , dadas las incertidumbres en $X_1, X_2, \dots, X_n$. La deducción es similar al caso de una variable que se presentó al final de la sección 3.3. Sean $\mu_1, \mu_2, \dots, \mu_n$ las medias de $X_1, X_2, \dots, X_n$, respectivamente. Entonces para cualesquiera mediciones razonablemente precisas, $X_1, X_2, \dots, X_n$ estarán bastante cerca de $\mu_1, \mu_2, \dots, \mu_n$ para que la linealización sea válida.

Se puede reescribir la ecuación (3.14) como

$$U \approx \left(U(\mu_1, \mu_2, \dots, \mu_n) - \frac{\partial U}{\partial X_1} \mu_1 - \frac{\partial U}{\partial X_2} \mu_2 - \dots - \frac{\partial U}{\partial X_n} \mu_n \right) + \frac{\partial U}{\partial X_1} X_1 + \frac{\partial U}{\partial X_2} X_2 + \dots + \frac{\partial U}{\partial X_n} X_n \quad (3.15)$$

Las cantidades $\partial U/\partial X_1, \partial U/\partial X_2, \dots, \partial U/\partial X_n$ son todas constantes, puesto que se evalúan en el punto $(\mu_1, \mu_2, \dots, \mu_n)$. Por tanto, la cantidad

$$U(\mu_1, \mu_2, \dots, \mu_n) - \frac{\partial U}{\partial X_1} \mu_1 - \frac{\partial U}{\partial X_2} \mu_2 - \dots - \frac{\partial U}{\partial X_n} \mu_n$$

también es constante. En este contexto, de la ecuación (2.40) (de la sección 2.5) y de la ecuación (3.4) (de la sección 3.2) se tiene que

$$\sigma_U \approx \sqrt{\left(\frac{\partial U}{\partial X_1}\right)^2 \sigma_{X_1}^2 + \left(\frac{\partial U}{\partial X_2}\right)^2 \sigma_{X_2}^2 + \dots + \left(\frac{\partial U}{\partial X_n}\right)^2 \sigma_{X_n}^2}$$

Ejercicios de la sección 3.4

- Determine la incertidumbre en U , suponiendo que $X = 10.0 \pm 0.5$, $Y = 5 \pm 0.1$ y
 - $U = XY^2$
 - $U = X^2 + Y^2$
 - $U = (X + Y^2)/2$
- Con referencia al ejercicio 8 de la sección 3.2, suponga que $\tau = 30.0 \pm 0.1$ Pa, $h = 10.0 \pm 0.2$ mm y $\mu = 1.49$ Pa·s con incertidumbre despreciable.
 - Estime V y determine la incertidumbre en la estimación.
 - ¿Qué proveería una reducción más grande en la incertidumbre en V : reducir la incertidumbre en τ a 0.01 Pa o reducir la incertidumbre en h a 0.1 mm?
- Entra aire en un compresor con una presión P_1 y sale con una presión P_2 , la presión intermedia está dada por $P_3 = \sqrt{P_1 P_2}$. Suponga que $P_1 = 10.1 \pm 0.3$ MPa y $P_2 = 20.1 \pm 0.4$ MPa.
 - Estime P_3 y determine la incertidumbre en la estimación.
 - ¿Qué proveería una reducción más grande en la incertidumbre en P_3 : reducir la incertidumbre en P_1 a 0.2 MPa o reducir la incertidumbre en P_2 a 0.2 MPa?
- Una manera de medir el contenido de agua en una tierra es pesarla antes y después de secarla en un horno. El contenido de agua es $W = (M_1 - M_2)/M_1$, donde M_1 es la masa antes de secarla y M_2 después de secarla. Suponiendo que $M_1 = 1.32 \pm 0.01$ kg y $M_2 = 1.04 \pm 0.01$ kg.
 - Estime W y determine la incertidumbre en la estimación.
 - ¿Qué proveería una reducción más grande en la incertidumbre en W : reducir la incertidumbre en M_1 a 0.005 kg o reducirla en M_2 en 0.005 kg?
- La ecuación de lentes dice que si un objeto se coloca a una distancia p de una lente y se forma una imagen a una distancia q de aquélla, entonces la longitud focal f satisface la ecuación $1/f = 1/p + 1/q$. Suponga $p = 2.3 \pm 0.2$ cm y $q = 3 \pm 0.2$ cm.
 - Calcule f y determine la incertidumbre.
 - ¿Qué proveería una reducción más grande en la incertidumbre en f : reducir la incertidumbre en p a 0.1 cm o en q a 0.1 cm?
- La presión P , temperatura T y volumen V de una mol de gas ideal está relacionada con la ecuación $PV = 8.31 T$, cuando P se mide en kilopascales, T se mide en grados Kelvin y V se mide en litros.

- a) Suponga que $P = 242.52 \pm 0.03$ kPa y $V = 10.103 \pm 0.002$ L. Estime T y determine la incertidumbre en la estimación.
- b) Suponga que $P = 242.52 \pm 0.03$ kPa y $T = 290.11 \pm 0.02$ K. Estime V y determine la incertidumbre en la estimación.
- c) Suponga que $V = 10.103 \pm 0.002$ L y $T = 290.11 \pm 0.02$ K. Estime P y determine la incertidumbre en la estimación.
7. Con referencia al ejercicio 7 en la sección 3.3, suponga que $g = 9.80$ m/s² exactamente, $d = 0.18 \pm 0.02$ m, $h = 4.86 \pm 0.06$ m y $l = 32.04 \pm 0.01$ m. Estime F y determine la incertidumbre en la estimación.
8. En el artículo “Temperature-Dependent Optical Constants of Water Ice in the Near Infrared: New Results and Critical Review of the Available Measurements” (B. Rajaram, D. Glandorf y colaboradores, en *Applied Optics*, 2001:4449-4462), el índice imaginario de refracción del agua congelada se presenta para diferentes frecuencias y temperaturas. A frecuencia de 372.1 cm⁻¹ y temperatura de 166 K, el índice se estima de 0.00116 . A la misma frecuencia y a temperatura de 196 K, el índice se estima de 0.00129 . La incertidumbre se reporta de 10^{-4} para cada uno de estos índices. La razón de los índices se estima de $0.00116/0.00129 = 0.899$. Determine la incertidumbre en esta razón.
9. Con referencia al ejercicio 10 de la sección 3.2, suponga que $\tau_0 = 50 \pm 1$ MPa, $w = 1.2 \pm 0.1$ mm y $k = 0.29 \pm 0.05$ mm⁻¹.
- a) Estime τ y determine la incertidumbre en la estimación.
- b) ¿Qué proveería una reducción más grande en la incertidumbre en τ : reducir la incertidumbre en τ_0 a 0.1 MPa, en w a 0.01 mm, o en k a 0.025 mm⁻¹?
- c) Un nuevo proceso algo más costoso permitiría que tanto τ_0 como w sean medidos con incertidumbre despreciable. ¿Vale la pena implementar el proceso? Explique.
10. De acuerdo con la ley de Snell, el ángulo de refracción θ_2 de un rayo de luz que viaja en un medio cuyo índice de refracción n está relacionado con el ángulo de incidencia θ_1 de un rayo que está viajando en el vacío por medio de la ecuación $\sin\theta_1 = n \sin\theta_2$. Suponga que $\theta_1 = 0.3672 \pm 0.005$ rad y $\theta_2 = 0.2943 \pm 0.004$ rad. Estime n y determine la incertidumbre en la estimación.
11. Arqueólogos que estudian los métodos de almacenamiento de carne empleados en el territorio Nunamiut, en Alaska, han desarrollado un índice de secado de carne. Lo que sigue

es una versión ligeramente simplificada del índice en el artículo “A Zooarchaeological Signature for Meat Storage: Rethinking the Drying Utility Index” (T. Friesen, en *American Antiquity*, 2001:315-331). Sea m el peso de carne, b el peso del hueso y g el peso neto de alguna parte del caribú. El índice de secado de carne y está dado por $y = mb/g$. Suponga que para una costilla especial del caribú, se hicieron las siguientes mediciones (en gramos): $g = 3\,867.4 \pm 0.3$, $b = 1\,037.0 \pm 0.2$, $m = 2\,650.4 \pm 0.1$.

- a) Estime y y determine la incertidumbre en la estimación.
- b) ¿Qué proveería una reducción más grande en la incertidumbre en y : reducir la incertidumbre en g a 0.1 g, en b a 0.1 g, o en m a 0 ?

12. La resistencia R (en ohms) de un conductor cilíndrico está dado por $R = kl/d^2$, donde l es la longitud, d es el diámetro y k es una constante de proporcionalidad. Suponga que $l = 14 \pm 0.1$ cm y $d = 4.4 \pm 0.1$ cm.

- a) Estime R y determine la incertidumbre en la estimación. Su respuesta estará en relación con la constante de proporcionalidad k .
- b) ¿Qué proveería una reducción más grande en la incertidumbre en R : reducir la incertidumbre en l a 0.05 cm o en d a 0.05 cm?

13. Un cable cilíndrico de radio R se alarga cuando se somete a una fuerza de tensión F . Sea L_0 la longitud inicial del cable y L_1 la longitud final. El módulo de Young para el material está dado por

$$Y = \frac{FL_0}{\pi R^2(L_1 - L_0)}$$

Suponga que $F = 800 \pm 1$ N, $R = 0.75 \pm 0.1$ mm, $L_0 = 25 \pm 0.1$ mm y $L_1 = 30 \pm 0.1$ mm.

- a) Estime Y y determine la incertidumbre en la estimación.
- b) De las incertidumbres en F , R , L_0 y L_1 , sólo una tiene un efecto no despreciable en la incertidumbre en Y . ¿Cuál es?

14. De acuerdo con la ley del enfriamiento de Newton, el tiempo t necesario para que un objeto con una temperatura inicial T_0 se enfríe a la temperatura T con una temperatura del ambiente T_a está dada por

$$t = \frac{\ln(T_0 - T_a)}{k} - \frac{\ln(T - T_a)}{k}$$

donde k es una constante. Suponga que para cierto tipo de recipiente, $k = 0.025$ min⁻¹. Sea t el número de minutos ne-

cesario para enfriar el recipiente a 50°F. Suponga que $T_0 = 70.1 \pm 0.2^\circ\text{F}$ y $T_a = 35.7 \pm 0.1^\circ\text{F}$. Estime t y determine la incertidumbre en la estimación.

15. Con referencia al ejercicio 14, en un experimento para determinar el valor de k , la temperatura T al tiempo $t = 10$ min se mide de $T = 54.1 \pm 0.2^\circ\text{F}$. Suponga que $T_0 = 70.1 \pm 0.2^\circ\text{F}$ y $T_a = 35.7 \pm 0.1^\circ\text{F}$. Estime k y determine la incertidumbre en la estimación.

16. El desplazamiento vertical v de una grieta infiltrada con fibra de concreto en un miembro a una tensión de corte máxima está dada por $v = a + bw$, donde w es el ancho de la grieta y se estima a y b a partir de los datos como $a = 2.5 \pm 0.1$ mm y $b = 0.05 \pm 0.01$. Suponga que $w = 1.2 \pm 0.1$ mm.

- Estime v y determine la incertidumbre en la estimación.
- De las incertidumbres en w , a y b , sólo una tiene un efecto en la incertidumbre no despreciable en v . ¿Cuál es?

17. La forma de una bacteria se puede aproximar con un cilindro de radio r y altura h terminado en cada extremo con un hemisferio. El volumen y el área superficial de la bacteria están dados por

$$V = \pi r^2(h + 4r/3)$$

$$S = 2\pi r(h + 2r)$$

Se sabe que la razón R con la que se absorbe una sustancia química en la bacteria es $R = c(s/V)$, donde c es una constante de proporcionalidad. Suponga que para cierta bacteria, $r = 0.9 \pm 0.1$ μm y $h = 1.7 \pm 0.1$ μm .

- ¿Los valores calculados de S y V son independientes? Explique.
- Suponiendo que las mediciones de r y h sean independientes, estime R y determine la incertidumbre en la estimación. Su respuesta estará en relación con c .

18. Estime U y determine la incertidumbre respectiva en la estimación, suponiendo que $X = 5 \pm 0.2$, $Y = 10 \pm 0.5$ y

- $U = X\sqrt{Y}$
- $U = 2Y/\sqrt{X}$
- $U = X^2 + Y^2$

19. Con referencia al ejercicio 8 en la sección 3.2, suponga que $\tau = 35.2 \pm 0.1$ Pa, $h = 12.0 \pm 0.3$ mm y $\mu = 1.49$ Pa·s con incertidumbre despreciable. Estime V y determine la incertidumbre relativa en la estimación.

20. Con referencia al ejercicio 3, suponga que $P_1 = 15.3 \pm 0.2$ MPa y $P_2 = 25.8 \pm 0.1$ MPa. Estime P_3 y determine la incertidumbre relativa en la estimación.

21. Con referencia al ejercicio 5, suponga que $p = 4.3 \pm 0.1$ cm y $q = 2.1 \pm 0.2$ cm. Estime f y determine la incertidumbre relativa en la estimación.

22. Con referencia al ejercicio 6.

- Suponga que $P = 224.51 \pm 0.04$ kPa y $V = 11.237 \pm 0.002$ L. Estime T y determine la incertidumbre relativa en la estimación.
- Suponga que $P = 224.51 \pm 0.04$ kPa y $T = 289.33 \pm 0.02$ K. Estime V y determine la incertidumbre relativa en la estimación.
- Suponga que $V = 11.203 \pm 0.002$ L y $T = 289.33 \pm 0.02$ K. Estime P y determine la incertidumbre relativa en la estimación.

23. Con referencia al ejercicio 10, estime n y determine la incertidumbre relativa en la estimación, de las siguientes mediciones: $\theta_1 = 0.216 \pm 0.003$ rad y $\theta_2 = 0.456 \pm 0.005$ rad.

24. Con referencia al ejercicio 12, suponga que $l = 10$ cm $\pm$ 0.5% y $d = 10.4$ cm $\pm$ 0.5%.

- Haga una estimación de R y determine la incertidumbre relativa en la estimación. ¿La incertidumbre relativa depende de k ?
- Suponga que ya sea l o d se puede volver a medir con una incertidumbre relativa de 0.2%. ¿Cuál se debe volver a medir para dar una mejoría más grande en la incertidumbre relativa de la resistencia?

25. Con referencia al ejercicio 13, suponga que $F = 750 \pm 1$ N, $R = 0.65 \pm 0.09$ mm, $L_0 = 23.7 \pm 0.2$ mm y $L_1 = 27.7 \pm 0.2$ mm. Estime Y y determine la incertidumbre relativa en la estimación.

26. Con referencia al ejercicio 14, suponga que $T_0 = 73.1 \pm 0.1^\circ\text{F}$, $T_a = 37.5 \pm 0.2^\circ\text{F}$, $k = 0.032$ min⁻¹ con incertidumbre despreciable y $T = 50^\circ\text{F}$ exactamente. Estime t y determine la incertidumbre relativa en la estimación.

27. Con referencia al ejercicio 17, suponga que para cierta bacteria $r = 0.8 \pm 0.1$ μm y $h = 1.9 \pm 0.1$ μm .

- Estime S y determine la incertidumbre relativa en la estimación.

- b) Estime V y determine la incertidumbre relativa en la estimación.
- c) Estime R y determine la incertidumbre relativa en la estimación.
- d) ¿La incertidumbre relativa en R depende de c ?

28. Sean X y Y mediciones independientes y sean c , n y m constantes. Demuestre que la incertidumbre relativa en $U = X^n Y^m$ es

$$\frac{\sigma_U}{U} = \sqrt{\left(n \frac{\sigma_X}{X}\right)^2 + \left(m \frac{\sigma_Y}{Y}\right)^2}$$

Ejercicios adicionales del capítulo 3

- Suponga que X , Y y Z son mediciones independientes con $X = 25 \pm 1$, $Y = 5 \pm 0.3$ y $Z = 3.5 \pm 0.2$. Determine las incertidumbres en cada una de las cantidades siguientes:
 - $X + YZ$
 - $X/(Y - Z)$
 - $X\sqrt{Y + e^Z}$
 - $X \ln(Y^2 + Z)$
- Suponga que X , Y y Z son mediciones independientes y que la incertidumbre relativa en X es 5%, en Y es 10% y en Z es 15%. Determine la incertidumbre relativa en cada una de las cantidades siguientes:
 - XY/Z
 - $X^{3/2}(YZ)^{1/3}$
 - $X^3\sqrt{Z/Y}$
- Un artículo se forma colocando dos componentes uno tras otro. La longitud de cada componente será medido.
 - Si la incertidumbre en la medición de la longitud de cada componente es 0.1 mm, ¿cuál es la incertidumbre en la longitud combinada de los dos componentes?
 - Si se desea estimar la longitud del artículo con una incertidumbre de 0.05 mm, ¿cuál debe ser la incertidumbre en la medición de cada componente? Suponga que las incertidumbres en las dos mediciones son iguales.
- Para algunas mutaciones genéticas, se piensa que la frecuencia del gen mutante en los hombres aumenta linealmente con la edad. Si m_1 es la frecuencia a la edad t_1 y m_2 es la frecuencia a la edad t_2 , entonces la razón anual de aumento se estima con $r = (m_2 - m_1)/(t_2 - t_1)$. En un análisis de reacciones en la cadena de la polimerasa, la frecuencia en hombres de 20 años se calculó de 17.7 ± 1.7 por μg de ADN y la frecuencia en hombres de 40 años se calculó de 35.9 ± 5.8 por μg de ADN. Suponga que la edad se mide con incertidumbre despreciable.
 - Estime la razón anual del aumento y determine la incertidumbre en la estimación.
 - Determine la incertidumbre relativa en la estimación de la razón del aumento.
- La ecuación de Darcy-Weisbach establece que la potencia de la capacidad de generación en un sistema hidroeléctrico que se pierde debido a la pérdida de carga está dada por $P = \eta\gamma QH$, donde η es la eficiencia de la turbina, γ la gravedad específica del agua, Q la razón de flujo y H la pérdida de carga. Suponga que $\eta = 0.85 \pm 0.02$, $H = 3.71 \pm 0.10$ m, $Q = 60 \pm 1$ m³/s y $\gamma = 9\,800$ N/m³ con incertidumbre despreciable.
 - Estime la pérdida de potencia (las unidades serán watts) y determine la incertidumbre en la estimación.
 - Determine la incertidumbre relativa en la estimación en la pérdida de potencia.
 - ¿Qué proveería una reducción más grande en la incertidumbre en P : reducir la incertidumbre en η a 0.01, en H a 0.05 o en Q a 0.5?
- Sean A y B dos variantes (alelos) del ADN en cierta ubicación en el genoma. Sea p la proporción de alelos en una población que son del tipo A y q del tipo B . El principio de equilibrio Hardy-Weinberg establece que la proporción P_{AB} de organismos que son del tipo AB es igual a pq . En una visión general de población de una especie especial, la proporción de alelos de tipo A se estima de 0.360 ± 0.048 y la del tipo B se estima independientemente de 0.250 ± 0.043 .
 - Estime la proporción de organismos que son de tipo AB y determine la incertidumbre en la estimación.
 - Determine la incertidumbre relativa en la proporción.
 - ¿Qué proveería una reducción más grande en la incertidumbre en la proporción: reducir la incertidumbre en la proporción del tipo A a 0.02 o del tipo B a 0.02?

7. La capacidad calorífica de un calorímetro se sabe que es de $4 \text{ kJ/}^\circ\text{C}$, con incertidumbre despreciable. El número de calorías alimentarias (kilocalorías) por gramo de una sustancia está dado por $C = cH(\Delta T)/m$, donde C es el número de calorías alimentarias, H la capacidad calorífica del calorímetro, ΔT el aumento en la temperatura en $^\circ\text{C}$ causado por quemar la sustancia en el calorímetro, m la masa de la sustancia en gramos y $c = 0.2390 \text{ cal/kJ}$ es el factor de conversión de kilo joules a calorías alimentarias. Una cantidad de mayonesa con masa $0.40 \pm 0.01 \text{ g}$ es quemada en un calorímetro. El aumento de la temperatura es $2.75 \pm 0.02^\circ\text{C}$.
- Estime el número de calorías alimentarias por gramo de mayonesa y determine la incertidumbre en la estimación.
 - Determine la incertidumbre relativa en el número estimado de calorías alimentarias.
 - ¿Qué proveería una reducción más grande en la incertidumbre en C : reducir la incertidumbre en la masa a 0.005 g o en ΔT a 0.01°C ?
8. Se hicieron 22 mediciones independientes de la dureza de una soldadura, utilizando la balanza A de Rockwell. El promedio fue de 65.52 y la desviación estándar de 0.63 .
- Estime la dureza de esta soldadura y determine la incertidumbre en la estimación.
 - En una sola medición hecha de la dureza de otra soldadura del mismo metal base, la medición es de 61.3 . ¿Cuál es la incertidumbre en esta medición?
9. El artículo “Insights into Present-Day Crustal Motion in the Central Mediterranean Area from GPS Surveys” (M. Anzi-dei, P. Baldi y colaboradores, en *Geophysical Journal International*, 2001:98-100) informa que los componentes de la velocidad de la corteza terrestre en Zimmerwald, Suiza, son $22.10 \pm 0.34 \text{ mm/año}$ en dirección norte y $14.3 \pm 0.32 \text{ mm/año}$ en dirección este.
- Estime la velocidad de la corteza terrestre y determine la incertidumbre en la estimación.
 - Utilizando la respuesta del inciso (a), estime el número de años que le tomará a la corteza desplazarse 100 mm y determine la incertidumbre en la estimación.
10. Si dos gases tienen masas molares M_1 y M_2 , la ley de Graham establece que la razón R de sus tasas de efusión a través de una pequeña abertura está dada por $R = \sqrt{M_1/M_2}$. La tasa de efusión de un gas desconocido que pasa a través de una pequeña abertura se midió de 1.66 ± 0.03 veces más grande que la tasa de efusión del dióxido de carbono. La masa molar de éste se puede tomar igual a 44 g/mol con incertidumbre despreciable.
- Estime la masa molar del gas desconocido y determine la incertidumbre en la estimación.
 - Determine la incertidumbre relativa en la estimación de la masa molar.
11. Un artículo laminado está hecho de seis capas. Las dos exteriores tienen un espesor de $1.25 \pm 0.10 \text{ mm}$ y las cuatro interiores tienen un espesor de $0.80 \pm 0.05 \text{ mm}$. Suponga que los espesores de las capas son independientes. Estime el espesor del artículo y determine la incertidumbre en la estimación.
12. El artículo “Effect of Varying Solids Concentration and Organic Loading on the Performance of Temperature Phased Anaerobic Digestion Process” (S. Vanderburgh y T. Ellis, en *Water Environment Research*, 2002:142-148) analiza experimentos para determinar el efecto de la concentración de sólidos en el desarrollo de los métodos de tratamiento para lodos de agua residuales. En el primer experimento, la concentración de sólidos (en g/L) fue 43.94 ± 1.18 . En el segundo, que era independiente del primero, la concentración era 48.66 ± 1.76 . Estime la diferencia en la concentración entre dos experimentos y determine la incertidumbre en la estimación.
13. En el artículo “Measurements of the Thermal Conductivity and Thermal Diffusivity of Polymer Melts with the Short-Hot-Wire Method” (X. Zhang, W. Hendro y colaboradores, en *International Journal of Thermophysics*, 2002:1077-1090), la difusión térmica de un líquido medido por el método transitorio del alambre corto caliente está dado por
- $$\lambda = \frac{VIA}{\pi la}$$
- donde λ es la difusión térmica; V y I el voltaje y la corriente aplicada al alambre caliente, respectivamente; l la longitud del alambre; y A y a las cantidades que implican la temperatura, cuyos valores se estiman por separado. En este artículo, las incertidumbres relativas de estas cantidades están dadas de la siguiente manera: V , 0.01% ; I , 0.01% ; l , 1% ; A , 0.1% ; a , 1% .
- Encuentre la incertidumbre relativa en λ .
 - ¿Qué reduciría más la incertidumbre relativa: disminuir la incertidumbre relativa en l a 0.5% o reducir las incertidumbres relativas en V , I , y A para cada una a 0 ?
14. Un cable está formado por varios alambres paralelos. La fuerza del cable se puede estimar de las intensidades de los alambres por separado por cualesquiera de los dos métodos. En el *método del cable dúctil*, su fuerza se estima como la suma de las intensidades de los alambres. En el *método de*

cable quebradizo, su fuerza se estima como la fuerza del alambre más débil multiplicada por el número de alambres. Un cable especial está compuesto de 12 alambres. Cuatro de ellos tienen una fuerza de $6\,000 \pm 20$ lb, cuatro más de $5\,700 \pm 30$ lb y otros cuatro de $6\,200 \pm 40$ lb.

- Estime la fuerza del cable y determine la incertidumbre en la estimación, utilizando el método de cable dúctil.
- Estime la fuerza del cable y determine la incertidumbre en la estimación, utilizando el método de cable quebradizo.

15. Con referencia al ejercicio 14, un cable está compuesto de 16 alambres. La fuerza de cada uno es $5\,000 \pm 20$ lb.

- ¿La fuerza estimada del cable será la misma con el método de cable dúctil como con el método de cable quebradizo?
- ¿La incertidumbre en la fuerza estimada del cable será la misma con el método de cable dúctil como con el método de cable quebradizo? Explique por qué sí o por qué no.

16. La media de la producción del proceso A se estima de 80 ± 5 , donde las unidades son el porcentaje de un máximo teórico. La media de la producción del proceso B se estima de 90 ± 3 . El aumento relativo obtenido del proceso B se estima de $(90 - 80)/80 = 0.125$. Determine la incertidumbre en la estimación.

17. La razón de flujo del agua por un tubo cilíndrico está dado por $Q = \pi r^2 v$, donde r es el radio del tubo y v es la velocidad de flujo.

- Suponga que $r = 3.00 \pm 0.03$ y $v = 4 \pm 0.2$ m/s. Estime Q y determine la incertidumbre en la estimación.
- Suponga que $r = 4.00 \pm 0.04$ m y $v = 2 \pm 0.1$ m/s. Estime Q y determine la incertidumbre en la estimación.
- Si r y v no se han medido, pero se sabe que la incertidumbre relativa en r es 1% y que la incertidumbre relativa en v es 5%, ¿es posible calcular la incertidumbre relativa en Q ? Si es así, calcule la incertidumbre relativa. Si no, explique qué información adicional se necesita.

18. La conversión del ciclobutano (C_4H_8) a etileno (C_2H_4) es una reacción de primer orden. Esto significa que la concentración de ciclobutano al tiempo t está dada por $\ln C = \ln C_0 - kt$, donde C es la concentración al tiempo t , C_0 es la concentración inicial, t es el tiempo desde que la reacción empezó y k es la razón constante. Suponga que $C_0 = 0.2$ mol/L con incertidumbre despreciable. Después de 300 segundos a temperatura constante, la concentración se mide de $C =$

0.174 ± 0.005 mol/L. Suponga que el tiempo se puede medir con incertidumbre despreciable.

- Estime la razón constante k y determine la incertidumbre en la estimación. Las unidades de k serán s^{-1} .
- Determine la incertidumbre relativa en k .
- La vida media $t_{1/2}$ de la reacción es el tiempo que se necesita para que la concentración se reduzca a la mitad de su valor inicial. La vida media está relacionada con la razón constante por $t_{1/2} = (\ln 2)/k$. Utilizando el resultado que encontró en el inciso (a), determine la incertidumbre en la vida media.
- Determine la incertidumbre en la vida media.

19. La descomposición del dióxido de nitrógeno (NO_2) en monóxido de nitrógeno (NO) y oxígeno es una reacción de segundo orden. Esto significa que la concentración C de NO_2 al tiempo t está dada por $1/C = kt + 1/C_0$, donde C_0 es la concentración inicial y k es la tasa constante. Suponga que se sabe que la concentración inicial es exactamente de 0.03 mol/L. Suponga que el tiempo se puede medir con incertidumbre despreciable.

- Después de 40 s, la concentración C se mide de $0.0023 \pm 2.0 \times 10^{-4}$ mol/L. Estime la tasa constante k y determine la incertidumbre en la estimación.
- Después de 50 s, la concentración C se mide de $0.0018 \pm 2.0 \times 10^{-4}$ mol/L. Estime la tasa constante k y determine la incertidumbre en la estimación.
- Denote los estimadores de la razón constante k en los incisos a) y b) por $\hat{k}_1$ y $\hat{k}_2$, respectivamente. El promedio $(\hat{k}_1 + \hat{k}_2)/2$ se usa como un estimador aproximado de k . Determine la incertidumbre en esta estimación.
- Determine el valor de c para que el promedio ponderado $c\hat{k}_1 + (1 - c)\hat{k}_2$ tenga la incertidumbre más pequeña.

20. Dos estudiantes quieren medir la aceleración a de un carro que baja rodando por un plano inclinado. El carro arranca desde el reposo y viaja una distancia s hacia abajo del plano. El primer estudiante estima la aceleración al medir la velocidad instantánea v en cuanto el carro ha viajado s metros y usa la fórmula $a = v^2/2s$. El segundo estima la aceleración midiendo el tiempo, en segundos, que le toma al carro viajar los s metros y usa la fórmula $a = 2s/t^2$. Suponga que $s = 1$ m y que tiene una incertidumbre despreciable en s . Suponga que $v = 3.2 \pm 0.1$ m/s y que $t = 0.63 \pm 0.01$ s. Suponga que las mediciones de v y t son independientes.

- Calcule la aceleración utilizando el método del primer estudiante. Llame a esta estimación a_1 . Determine la incertidumbre a_1 .

- b) Determine la aceleración utilizando el método del segundo estudiante. Llame a esta estimación a_2 . Determine la incertidumbre a_2 .
- c) Encuentre el promedio ponderado de a_1 y a_2 que tiene la incertidumbre más pequeña. Encuentre la incertidumbre de este promedio ponderado.

21. Una pista tiene la forma de un cuadrado limitado con dos lados opuestos por semicírculos. La longitud de un lado del cuadrado se mide de 181.2 ± 0.1 m.

- a) Determine el área del cuadrado y su incertidumbre.
- b) Determine el área de uno de los semicírculos y su incertidumbre.
- c) Sea S el área del cuadrado como se calculó en el inciso a) y sea C el área de uno de los semicírculos como se calcularon en el inciso b). El área encerrada por la pista es $A = S + 2C$. Alguien calcula la incertidumbre en A como $\sigma_A = \sqrt{\sigma_S^2 + 4\sigma_C^2}$. ¿es esto correcto? Si es así, explique por qué. Si no, calcule la incertidumbre en A correctamente.

22. Si X es una medición no sesgada de un valor real μ_X y $U(X)$ es una función no lineal de X , entonces en la mayoría de los casos U es un estimador sesgado del valor real $U(\mu_X)$. En la mayoría de los casos se ignora el sesgo. Sin embargo, si es importante reducir el sesgo, un estimador con *corrección de sesgo* es $U(X) - (1/2)(d^2U/dX^2)\sigma_X^2$. En general el estimador con corrección de sesgo es sesgado, pero tiene sesgo menor que $U(X)$.

Suponga que el radio de un círculo se mide de $r = 3.0 \pm 0.1$ cm.

- a) Estime el área A y determine la incertidumbre de la estimación, sin corrección de sesgo.
- b) Calcule el estimador con corrección de sesgo de A .
- c) Compare la diferencia entre el estimador con corrección de sesgo y el estimador sin corrección de sesgo de la incertidumbre en el estimador sin corrección de sesgo. ¿Es importante la corrección por sesgo en este caso? Explique.

23. Si $X_1, X_2, \dots, X_n$ son mediciones independientes no sesgadas de los valores reales $\mu_1, \mu_2, \dots, \mu_n$ y $U(X_1, X_2, \dots, X_n)$ es una función no lineal de $X_1, X_2, \dots, X_n$, entonces, en general, $U(X_1, X_2, \dots, X_n)$ es un estimador sesgado del valor real $U(\mu_1, \mu_2, \dots, \mu_n)$. Un estimador con corrección de sesgo es $U(X_1, X_2, \dots, X_n) - (1/2) \sum_{i=1}^n (\partial^2 U / \partial X_i^2) \sigma_{X_i}^2$.

Cuando entra aire en un compresor a presión P_1 y sale a presión P_2 , la presión intermedia está dada por $P_3 = \sqrt{P_1 P_2}$. Suponga que $P_1 = 8.1 \pm 0.1$ MPa y $P_2 = 15.4 \pm 0.2$ MPa.

- a) Estime P_3 y determine la incertidumbre en la estimación, sin corrección de sesgo.
- b) Estime P_3 con corrección de sesgo.
- c) Compare la diferencia entre los estimadores con y sin corrección de sesgo con la incertidumbre en el estimador sin corrección de sesgo. ¿Es importante, en este caso, la corrección por sesgo? Explique.

Capítulo 4

Distribuciones comúnmente usadas

Introducción

La inferencia estadística consiste en extraer una muestra de una población y analizar sus datos con el propósito de aprender acerca de ello. Muchas veces se tiene un conocimiento superficial de la función de masa de probabilidad o de la función de densidad de probabilidad de la población. En estos casos la función de masa o de densidad de probabilidad se aproxima mediante una de muchas familias comunes de curvas o funciones. En este capítulo se describen algunas de estas funciones comunes y las condiciones en que es apropiado utilizar cada una.

4.1 Distribución de Bernoulli

Imagine un experimento que tenga dos resultados. Al primero se le llama “éxito” y al otro “fracaso”. La probabilidad de éxito se denota por p . Por consecuencia, la probabilidad de fracaso es $1 - p$. Lo anterior representa un **ensayo de Bernoulli** con probabilidad de éxito p . El más sencillo de este tipo es el lanzamiento al aire de una moneda. Los posibles resultados son “cara” o “cruz”. Si “cara” se define como éxito, entonces p constituye esa probabilidad. En una moneda, $p = 1/2$. Otro ejemplo de ese ensayo es la selección de un componente a partir de una población de componentes, pero algunos están defectuosos. Si se define como “éxito” a uno de éstos, entonces p significa la proporción de componentes defectuosos en la población.

Para cualquier ensayo de Bernoulli se define a la variable aleatoria X así: Si el experimento propicia “éxito”, entonces $X = 1$. De lo contrario, $X = 0$. De ahí que X sea una variable aleatoria discreta, con función de masa de probabilidad $p(x)$ definida por

$$p(0) = P(X = 0) = 1 - p$$

$$p(1) = P(X = 1) = p$$

$$p(x) = 0 \text{ para cualquier valor de } x \text{ diferente a } 0 \text{ o } 1$$

Se dice que la variable aleatoria X sigue una **distribución de Bernoulli** con parámetro p . La notación es $X \sim \text{Bernoulli}(p)$. La figura 4.1 muestra histogramas de probabilidad para las funciones de masa de probabilidad de Bernoulli(0.5) y de Bernoulli(0.8).

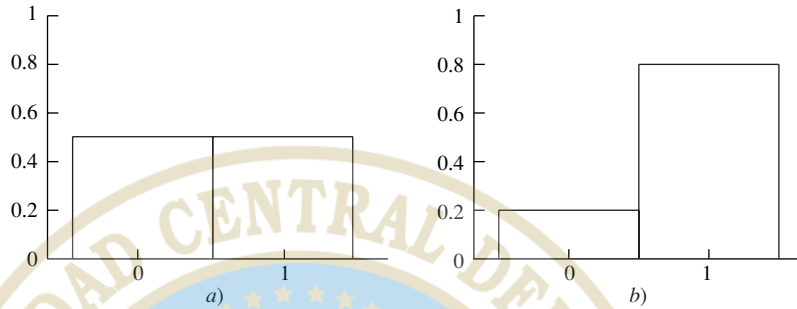


FIGURA 4.1 a) Histograma de probabilidad de Bernoulli(0.5). b) Histograma de probabilidad de Bernoulli(0.8).

Ejemplo 4.1

Cuando se lanza al aire una moneda hay una probabilidad de 0.5 de que caiga en “cara”. Sea $X = 1$ si la moneda cae en “cara” y $X = 0$ si cae en “cruz”. ¿Cuál es la distribución de X ?

Solución

Puesto que $X = 1$ cuando cae “cara”, ésta es resultado de éxito. La probabilidad de éxito, $P(X = 1)$, es igual a 0.5. Por tanto, $X \sim \text{Bernoulli}(0.5)$.

Ejemplo 4.2

Cuando se lanza un dado hay una probabilidad de $1/6$ de que salga 6. Sea $X = 1$ si el dado cae seis y $X = 0$ en cualquier otro caso. ¿Cuál es la distribución de X ?

Solución

La probabilidad de éxito es $p = P(X = 1) = 1/6$. Por lo que $X \sim \text{Bernoulli}(1/6)$.

Ejemplo 4.3

Diez por ciento de los componentes fabricados mediante determinado proceso está defectuoso. Se selecciona un componente aleatoriamente. Sea $X = 1$ si el componente está defectuoso y $X = 0$ en cualquier otro caso. ¿Cuál es la distribución de X ?

Solución

La probabilidad de éxito es $p = P(X = 1) = 0.1$. Por lo que $X \sim \text{Bernoulli}(0.1)$.

Media y varianza de una variable aleatoria de Bernoulli

Es fácil calcular la media y la varianza de una variable aleatoria Bernoulli. Si $X \sim \text{Bernoulli}(p)$, entonces, al usar las ecuaciones (2.29) y (2.30) (en la sección 2.4), se calcula

$$\begin{aligned}\mu_X &= (0)(1 - p) + (1)(p) \\ &= p \\ \sigma_X^2 &= (0 - p)^2(1 - p) + (1 - p)^2(p) \\ &= p(1 - p)\end{aligned}$$

Resumen	
Si $X \sim \text{Bernoulli}(p)$, entonces	
$\mu_X = p$	(4.1)
$\sigma_X^2 = p(1 - p)$	(4.2)

Ejemplo
4.4

Con referencia al ejemplo 4.3, determine μ_X y σ_X^2 .

Solución

Puesto que $X \sim \text{Bernoulli}(0.1)$, la probabilidad de éxito p es igual a 0.1. Al usar las ecuaciones (4.1) y (4.2), $\mu_X = 0.1$ y $\sigma_X^2 = 0.1(1 - 0.1) = 0.09$.

Ejercicios para la sección 4.1

1. Un jugador de basquetbol está a punto de tirar hacia la parte superior del tablero. La probabilidad de que anote el tiro es de 0.55.

a) Sea $X = 1$, si anota el tiro, si no lo hace, $X = 0$. Determine la media y la varianza de X .

b) Si anota el tiro, su equipo obtiene dos puntos; si lo falla, su equipo no recibe puntos. Sea Y el número de puntos anotados. ¿Tiene una distribución de Bernoulli? Si es así, encuentre la probabilidad de éxito. Si no, explique por qué.

c) Determine la media y varianza de Y .
2. En un restaurante de comida rápida, 25% de las órdenes para beber es una bebida pequeña, 35% una mediana y 40% una grande. Sea $X = 1$ si se escoge aleatoriamente una orden de una bebida pequeña y $X = 0$ en cualquier otro caso. Sea $Y = 1$ si la orden es una bebida mediana y $Y = 0$ en cualquier otro caso. Sea $Z = 1$ si la orden es una bebida pequeña o mediana y $Z = 0$ para cualquier otro caso.

a) Sea p_X la probabilidad de éxito de X . Determine p_X .

b) Sea p_Y la probabilidad de éxito de Y . Determine p_Y .

c) Sea p_Z la probabilidad de éxito de Z . Determine p_Z .

d) ¿Es posible que X y Y sean iguales a 1?

e) ¿Es $p_Z = p_X + p_Y$?

f) ¿Es $Z = X + Y$? Explique.

3. Cuando se aplica cierto barniz a una superficie de cerámica, 5% es la probabilidad de que se decolore, 20% de que se agriete, y 23% de que se decolore o no se agriete, o ambas. Sea $X = 1$ si se produce una decoloración y $X = 0$ en cualquier otro caso; $Y = 1$ si hay alguna grieta y $Y = 0$ en cualquier otro caso; $Z = 1$ si hay decoloración o grieta, o ambas, y $Z = 0$ en cualquier otro caso.
- Sea p_X la probabilidad de éxito de X . Determine p_X .
 - Sea p_Y la probabilidad de éxito de Y . Determine p_Y .
 - Sea p_Z la probabilidad de éxito de Z . Determine p_Z .
 - ¿Es posible que X y Y sean igual a 1?
 - ¿Es $p_Z = p_X + p_Y$?
 - ¿Es $Z = X + Y$? Explique.
4. Sean X y Y variables aleatorias de Bernoulli. Sea $Z = X + Y$.
- Demuestre que si X y Y no pueden ser iguales a 1, entonces Z es variable aleatoria de Bernoulli.
 - Demuestre que si X y Y no pueden ser iguales a 1, entonces $p_Z = p_X + p_Y$.
 - Demuestre que si X y Y pueden ser iguales a 1, entonces Z no es una variable aleatoria de Bernoulli.
5. Se lanza al aire una moneda de 1 y de 5 centavos. Sea $X = 1$ si sale “cara” en la moneda de 1 centavo y $X = 0$ en cualquier otro caso. Sea $Y = 1$ si sale “cara” en la moneda de 5 centavos y $Y = 0$ en cualquier otro caso. Sea $Z = 1$ si sale “cara” en ambas monedas y $Z = 0$ en cualquier otro caso.
- Sea p_X la probabilidad de éxito de X . Determine p_X .
 - Sea p_Y la probabilidad de éxito de Y . Determine p_Y .
 - Sea p_Z la probabilidad de éxito de Z . Determine p_Z .
 - ¿Son X y Y independientes?
 - ¿Es $p_Z = p_X p_Y$?
 - ¿Es $Z = XY$? Explique.
6. Se lanzan dos dados. Sea $X = 1$ si sale el mismo número en ambos y $X = 0$ en cualquier otro caso. Sea $Y = 1$ si la suma es 6 y $Y = 0$ en cualquier otro caso. Sea $Z = 1$ si sale el mismo número en los dados y ambos suman 6 (es decir, que salga 3 en los dos dados) y $Z = 0$ en cualquier otro caso.
- Sea p_X la probabilidad de éxito de X . Determine p_X .
 - Sea p_Y la probabilidad de éxito de Y . Determine p_Y .
 - Sea p_Z la probabilidad de éxito de Z . Determine p_Z .
 - ¿Son X y Y independientes?
 - ¿Es $p_Z = p_X p_Y$?
 - ¿Es $Z = XY$? Explique.
7. Sean X y Y variables aleatorias de Bernoulli. Sea $Z = XY$.
- Demuestre que Z es una variable aleatoria de Bernoulli.
 - Demuestre que si X y Y son independientes, entonces $p_Z = p_X p_Y$.

4.2 La distribución binomial

Extraer un solo componente de una población y determinar si está o no defectuoso es ejemplo de un ensayo de Bernoulli. En la práctica, es posible extraer varios componentes de una gran población y contar el número de elementos defectuosos. Esto implica realizar diversos ensayos de Bernoulli independientes y contar el número de éxitos. El número de éxitos es una variable aleatoria, que tiene una **distribución binomial**.

Ahora se presenta una descripción formal de la distribución binomial. Suponga que se lleva a cabo una serie de n ensayos de Bernoulli, cada uno con la misma probabilidad de éxito p . Además, suponga que los ensayos son *independientes*; esto es, que el resultado de un ensayo no influye en los resultados de alguno de los otros ensayos. Sea la variable aleatoria X igual al número de éxitos en n ensayos, entonces X tiene la **distribución binomial** con parámetros n y p . La notación es $X \sim \text{Bin}(n, p)$. X es una variable aleatoria discreta y sus posibles valores son $0, 1, \dots, n$.

Resumen

Si se realiza un total de n ensayos de Bernoulli y si

- Los ensayos son independientes
- Cada ensayo tiene la misma probabilidad de éxito p
- X es el número de éxitos en los n ensayos

entonces X tiene la distribución binomial con parámetros n y p , que se denota como $X \sim \text{Bin}(n, p)$.

Ejemplo

4.5

Se lanza al aire diez veces una moneda. Sea X el número de caras que aparecen. ¿Cuál es la distribución de X ?

Solución

Hay diez ensayos de Bernoulli independientes, cada uno con probabilidad de éxito de $p = 0.5$. La variable aleatoria X es igual al número de éxitos en los diez ensayos. Por consiguiente, $X \sim \text{Bin}(10, 0.5)$.

Recuerde, de la discusión de independencia en la sección 1.1, que cuando se toma una muestra de una población finita tangible, es posible tratar a los elementos de la muestra como independientes si es que la población es muy grande en comparación con el tamaño muestral. De lo contrario, los elementos de la muestra no serían independientes. En algunos casos el objetivo al extraer una muestra suele ser clasificar a cada elemento de la muestra en una de dos categorías. Por ejemplo, puede extraerse cierto número de elementos de una población y clasificar a cada uno como defectuoso o no. En estos casos cada elemento de la muestra representa un ensayo de Bernoulli, con una categoría contada como éxito y la otra como fracaso. Cuando la población de elementos es grande comparada con el número de elementos de la muestra, esos ensayos son independientes y, por razones prácticas, su número de éxitos tiene distribución binomial. Sin embargo, cuando el tamaño de la población no es tan grande, en comparación con la muestra, los ensayos no son independientes y su número de éxitos no tiene distribución binomial. La regla general es que si el tamaño muestral es de 5% o menos de la población, se puede emplear la distribución binomial.

Resumen

Suponga que una población finita contiene elementos de dos tipos, éxitos y fracasos, y que se extrae una muestra aleatoria simple de una población. Entonces, si el tamaño muestral no es mayor a 5% de aquella, se puede utilizar la distribución binomial para modelar el número de éxitos.

Ejemplo

4.6

Un lote contiene varios miles de componentes, de éstos 10% están defectuosos. Se extraen siete componentes de la población. Sea X el número de componentes defectuosos en la muestra. ¿Cuál es la distribución de X ?

Solución

Puesto que el tamaño muestral es pequeño en comparación con la población (es decir, menor a 5%), su número de éxitos representa una distribución binomial. Por tanto, se modela X con la distribución binomial $\text{Bin}(7, 0.1)$.

Función de masa de probabilidad de una variable aleatoria binomial

Ahora se deducirá la función de masa de probabilidad de una variable aleatoria binomial con un ejemplo. En una moneda específica existe una probabilidad de 0.6 de que salga “cara”. Se lanza al aire la moneda tres veces. Sea X el número de caras. Entonces $X \sim \text{Bin}(3, 0.6)$. Se calculará $P(X = 2)$.

Hay tres arreglos con dos “caras” en los tres lanzamientos de una moneda, HHT, HTH y THH. Primero se calcula la probabilidad de HHT. Este evento constituye una secuencia de eventos independientes: H en el primero, H en el segundo, y T en el tercer lanzamientos, respectivamente. Por separado se conoce las probabilidades de cada uno de ellos:

$$P(\text{H en el primer lanzamiento}) = 0.6, P(\text{H en el segundo}) = 0.6, P(\text{T en el tercero}) = 0.4$$

Como consecuencia de que los eventos son independientes, la probabilidad de que todos se presenten es igual al producto de sus probabilidades (ecuación 2.20 de la sección 2.3). Por tanto,

$$P(\text{HHT}) = (0.6)(0.6)(0.4) = (0.6)^2(0.4)^1$$

De forma similar, $P(\text{HTH}) = (0.6)(0.4)(0.6) = (0.6)^2(0.4)^1$ y $P(\text{THH}) = (0.4)(0.6)(0.6) = (0.6)^2(0.4)^1$. Es fácil ver que todos los diferentes arreglos de dos “caras” y una “cruz” tienen la misma probabilidad. Ahora

$$\begin{aligned} P(X = 2) &= P(\text{HHT o HTH o THH}) \\ &= P(\text{HHT}) + P(\text{HTH}) + P(\text{THH}) \\ &= (0.6)^2(0.4)^1 + (0.6)^2(0.4)^1 + (0.6)^2(0.4)^1 \\ &= 3(0.6)^2(0.4)^1 \end{aligned}$$

Al examinar este resultado se observa que el número 3 representa el número de arreglos de dos éxitos (“cara”) y un fracaso (“cruz”), 0.6 es la probabilidad de éxito p , el exponente 2 es el número de éxitos, 0.4 es la probabilidad de fracaso $1 - p$ y el exponente 1 es el número de fracasos.

Ahora se puede generalizar este resultado para generar una fórmula de la probabilidad de x éxitos en n ensayos de Bernoulli independientes con probabilidad de éxito p , en términos de x , n y p . En otras palabras, es posible calcular $P(X = x)$ donde $X \sim \text{Bin}(n, p)$. Se puede ver que

$$P(X = x) = (\text{número de arreglos de } x \text{ éxitos en } n \text{ ensayos}) \cdot p^x(1 - p)^{n-x} \quad (4.3)$$

Ahora todo lo que se necesita hacer es una expresión del número de arreglos de x éxitos en n ensayos. Para describir este número, se necesita la notación factorial. Para cualquier entero positivo n , la cantidad $n!$ (que se lee como “ n factorial”) es el número

$$(n)(n-1)(n-2) \dots (3)(2)(1)$$

Asimismo, se define $0! = 1$. El número de arreglos de x éxitos en n ensayos es $n!/x!(n-x)!$ (En la sección 2.2 se presentó una deducción de este resultado.) Ahora se puede definir la función de masa de probabilidad para una variable aleatoria binomial.

Si $X \sim \text{Bin}(n, p)$, la función de masa de probabilidad de X es

$$p(x) = P(X = x) = \begin{cases} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} & x = 0, 1, \dots, n \\ 0 & \text{de otro modo} \end{cases} \quad (4.4)$$

La figura 4.2 muestra los histogramas de probabilidad para las funciones de masa de probabilidad $\text{Bin}(10, 0.4)$ y $\text{Bin}(20, 0.1)$.

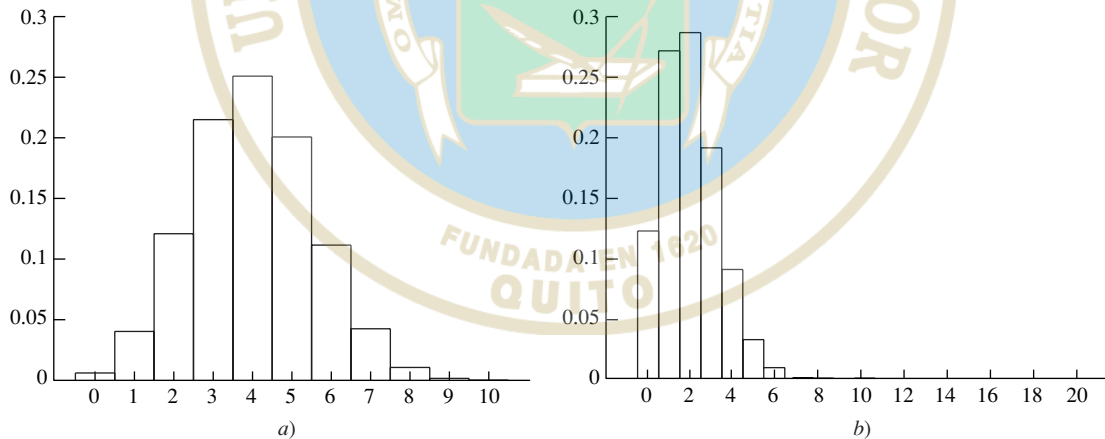


FIGURA 4.2 a) Histograma de probabilidad $\text{Bin}(10, 0.4)$. b) Histograma de probabilidad $\text{Bin}(20, 0.1)$.

Ejemplo 4.7

Determine la función de masa de probabilidad de la variable aleatoria X si $X \sim \text{Bin}(10, 0.4)$. Determine $P(X = 5)$.

Solución

Se emplea la ecuación (4.4) con $n = 10$ y $p = 0.4$. La función de masa de probabilidad es

$$p(x) = \begin{cases} \frac{10!}{x!(10-x)!} (0.4)^x (0.6)^{10-x} & x = 0, 1, \dots, 10 \\ 0 & \text{de otro modo} \end{cases}$$

$$\begin{aligned} P(X = 5) = p(5) &= \frac{10!}{5!(10-5)!} (0.4)^5 (0.6)^{10-5} \\ &= 0.2007 \end{aligned}$$

Ejemplo**4.8**

Se lanza al aire ocho veces un dado. Determine la probabilidad de que no salgan más de dos números seis.

Solución

Cada lanzamiento del dado es un experimento Bernoulli con una probabilidad de éxito de $1/6$. Sea X el número de seises en los ocho lanzamientos. Entonces $X \sim \text{Bin}(8, 1/6)$. Se necesita determinar a $P(X \leq 2)$. Con el uso de la función de masa de probabilidad,

$$\begin{aligned} P(X \leq 2) &= P(X = 0 \text{ o } X = 1 \text{ o } X = 2) \\ &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{8!}{0!(8-0)!} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{8-0} + \frac{8!}{1!(8-1)!} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{8-1} \\ &\quad + \frac{8!}{2!(8-2)!} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{8-2} \\ &= 0.2326 + 0.3721 + 0.2605 \\ &= 0.8652 \end{aligned}$$

La tabla A.1 (en el Apéndice A) presenta probabilidades binomiales de la forma $P(X \leq x)$ para $n \leq 20$ y valores seleccionados de p . Los ejemplos 4.9 y 4.10 muestran el uso de esta tabla.

Ejemplo**4.9**

Una gran compañía industrial hace un descuento en cualquier factura que se pague en un lapso de 30 días. De todas las facturas, 10% recibió el descuento. En una auditoría de la compañía se seleccionó aleatoriamente 12 facturas. ¿Cuál es la probabilidad de que menos de cuatro de las 12 facturas de la muestra tengan descuento?

Solución

Sea X el número de facturas en la muestra que recibe descuento. Entonces $X \sim \text{Bin}(12, 0.1)$. La probabilidad de que menos de cuatro facturas tengan descuento es $P(X \leq 3)$. Se consulta la tabla A.1 con $n = 12$, $p = 0.1$ y $x = 3$. Se encuentra que $P(X \leq 3) = 0.974$.

Algunas veces, la mejor manera para calcular la probabilidad de un evento consiste en estimar la probabilidad de que no ocurra el evento, y después restársela a 1. El ejemplo 4.10 proporciona ejemplo de esto último.

Ejemplo

4.10

Con referencia al ejemplo 4.9, ¿cuál es la probabilidad de que más de una de las 12 facturas de la muestra tenga descuento?

Solución

Sea X el número de facturas en la muestra que recibe descuento. Se desea calcular la probabilidad $P(X > 1)$. La tabla A.1 presenta las probabilidades en la forma $P(X \leq x)$. Por consecuencia, se observa que $P(X > 1) = 1 - P(X \leq 1)$. Consultando la tabla con $n = 12$, $p = 0.1$, $x = 1$, se encontró que $P(X \leq 1) = 0.659$. Por tanto, $P(X > 1) = 1 - 0.659 = 0.341$.

Una variable aleatoria binomial constituye la suma de variables aleatorias de Bernoulli

Suponga que se realiza n ensayos de Bernoulli independientes, cada uno con probabilidad de éxito p . Sean $Y_1, \dots, Y_n$ definidas de la siguiente manera: $Y_i = 1$ si el i -ésimo experimento da como resultado un éxito, y $Y_i = 0$, de otro modo. Entonces cada una de las variables aleatorias Y_i sigue una distribución Bernoulli(p). Ahora, sea que X represente el número de éxitos en los n ensayos. Entonces, $X \sim \text{Bin}(n, p)$. Puesto que cada Y_i es 0 o 1, la suma $Y_1 + \dots + Y_n$ es igual al número de los Y_i que tienen el valor 1, que es el número de éxitos en los n ensayos. Por tanto, $X = Y_1 + \dots + Y_n$. Esto último demuestra que una variable aleatoria binomial se puede expresar como la suma de variables aleatorias de Bernoulli. Dicho en otro modo, extraer un solo valor de una población $\text{Bin}(n, p)$ equivale a extraer una muestra de tamaño n de una población Bernoulli(p), y luego sumar los valores de la muestra.

La media y varianza de una variable aleatoria binomial

Es fácil calcular la media de una variable aleatoria binomial. Por ejemplo, si se lanza al aire una moneda durante diez veces, se espera ver, en promedio, cinco veces “cara”. El número cinco proviene de la multiplicación de la probabilidad de éxito (0.5) por el número de experimentos (10). Este método generalmente funciona. Si se realizara n ensayos de Bernoulli, cada uno con una probabilidad de éxito p , el número promedio de éxitos es np . Por consiguiente, si $X \sim \text{Bin}(n, p)$, entonces $\mu_X = np$. Se puede comprobar esta intuición cuando se observa que X es la suma de n variables de Bernoulli, cada una con media p . Por tanto, la media de X es la suma de las medias de variables aleatorias de Bernoulli que la componen, que es igual a np .

Se puede calcular a σ_X^2 al observar que X es la suma de variables aleatorias independientes de Bernoulli y recordando que la varianza de una variable aleatoria Bernoulli $p(1 - p)$. Por tanto, la varianza de X es la suma de las varianzas de las variables aleatorias de Bernoulli que la integran, que es igual a $np(1 - p)$.

Resumen

Si $X \sim \text{Bin}(n, p)$, entonces la media y la varianza de X están dadas por

$$\mu_X = np \quad (4.5)$$

$$\sigma_X^2 = np(1 - p) \quad (4.6)$$

Con el uso de la función de masa de probabilidad binomial (ecuación 4.4), se puede, en principio, calcular la media y la varianza de una variable aleatoria binomial mediante las definiciones de la media y la varianza de una variable aleatoria discreta (ecuaciones 2.29 y 2.30 en la sección 2.4). Estas expresiones implican sumatorias que son tediosas de evaluar. Es más sencillo considerar una variable aleatoria binomial como una suma de variables de Bernoulli aleatorias independientes.

Uso de una proporción muestral para estimar la probabilidad de éxito

En muchos casos no se conoce la probabilidad de éxito p asociada con cierto ensayo de Bernoulli, y se desea estimar su valor. Una forma natural de esto último consiste en realizar experimentos independientes n y contar el número X de éxitos. Para estimar la probabilidad de éxito p se calcula la proporción muestral $\hat{p}$.

$$\hat{p} = \frac{\text{número de éxitos}}{\text{número de ensayos}} = \frac{X}{n}$$

Esta notación sigue un patrón que es importante conocer. La probabilidad de éxito, que se desconoce, está representada por p . La proporción muestral, la cual se conoce, se representa por $\hat{p}$. El “sombbrero” (^) indica que $\hat{p}$ se utiliza para estimar un valor desconocido p .

Ejemplo 4.11

Un ingeniero que supervisa el control de calidad está probando la calibración de una máquina que empaca helado en contenedores. En una muestra de 20 de éstos, tres no están del todo llenos. Estime la probabilidad p de que la máquina no llene bien un contenedor.

Solución

La proporción muestral de contenedores no llenos es $\hat{p} = 3/20 = 0.15$. Se estima que la probabilidad p de que la máquina no llene bien un contenedor es también igual a 0.15.

Incertidumbre en una proporción muestral

Es importante considerar que la proporción muestral $\hat{p}$ es sólo una *estimación* de la probabilidad de éxito p , y que, en general, *no es igual a p* . Si se tomara otra muestra, probablemente el valor de $\hat{p}$ sería diferente. Es decir, hay incertidumbre en $\hat{p}$. Para que $\hat{p}$ sea una estimación útil, se debe calcular su sesgo y su incertidumbre. Ahora se hace esto. Sea n el tamaño muestral y X el número de éxitos, donde $X \sim \text{Bin}(n, p)$.

El sesgo es la diferencia $\mu_{\hat{p}} - p$. En virtud de que $\hat{p} = X/n$, se tiene, a partir de la ecuación (2.41), (de la sección 2.5) que

$$\begin{aligned}\mu_{\hat{p}} &= \mu_{X/n} = \frac{\mu_X}{n} \\ &= \frac{np}{n} = p\end{aligned}$$

Debido a que $\mu_{\hat{p}} = p$, $\hat{p}$ no es sesgado; en otras palabras, su sesgo es 0.

La incertidumbre es la desviación estándar $\sigma_{\hat{p}}$. A partir de la ecuación (4.6), la desviación estándar de X es $\sigma_X = \sqrt{np(1-p)}$. Puesto que $\hat{p} = X/n$, se tiene a partir de la ecuación (2.43) (de la sección 2.5) que

$$\begin{aligned}\sigma_{\hat{p}} &= \sigma_{X/n} = \frac{\sigma_X}{n} \\ &= \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}\end{aligned}$$

En la práctica, cuando se calcula la incertidumbre de $\hat{p}$, no se conoce la probabilidad de éxito p , por lo que se le aproxima con $\hat{p}$.

Resumen

Si $X \sim \text{Bin}(n, p)$, entonces la proporción muestral $\hat{p} = X/n$ se emplea para estimar la probabilidad de éxito p .

- $\hat{p}$ no sesgado.
- La incertidumbre en $\hat{p}$ es

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \quad (4.7)$$

En la práctica, cuando se calcula $\sigma_{\hat{p}}$ se sustituye $\hat{p}$ por p , dado que no se conoce p .

Ejemplo

4.12

Un comisionado de seguridad en una gran ciudad quiere estimar la proporción de edificios en la ciudad que viola los códigos de incendios. Se elige una muestra aleatoria de 40 edificios para inspeccionarlos, y se descubre que cuatro no cumplen el código de incendios. Estime la proporción de edificios en la ciudad que violan éste y encuentre la incertidumbre en la estimación.

Solución

Sea p la proporción de edificios en la ciudad que no cumple el código de incendios. El tamaño muestral (número de ensayos) es $n = 40$. El número de edificios con violaciones (éxitos) es $X = 4$. Se estima p con la proporción muestral $\hat{p}$.

$$\hat{p} = \frac{X}{n} = \frac{4}{40} = 0.10$$

Al usar la ecuación (4.7), la incertidumbre en $\hat{p}$ es

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Al sustituir $\hat{p} = 0.1$ por p y 40 por n , se obtiene

$$\begin{aligned}\sigma_{\hat{p}} &= \sqrt{\frac{(0.10)(0.90)}{40}} \\ &= 0.047\end{aligned}$$

En el ejemplo 4.12 se obtuvo que la incertidumbre en la proporción muestral era muy grande. Se puede reducir aquélla al aumentar el tamaño de esta última. El ejemplo 4.13 muestra cómo calcular el tamaño necesario para reducir la incertidumbre a una cantidad específica.

Ejemplo 4.13

En el ejemplo 4.12, ¿cuántos edificios adicionales deben inspeccionarse para que la incertidumbre de la proporción muestral de los edificios que no cumple el código sea sólo de 0.02?

Solución

Se necesita determinar el valor de n de tal forma que $\sigma_{\hat{p}} = \sqrt{p(1-p)/n} = 0.02$. Al aproximar p con $\hat{p} = 0.1$, se obtiene

$$\sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{n}} = 0.02$$

Al despejar n se tiene que $n = 225$. Ya se ha señalado 40 edificios, por lo que ahora se necesita indicar 185 más.

Algunas veces se desea estimar el valor de una función $f(p)$ de una probabilidad de éxito p . En estos casos se estima $f(p)$ con $f(\hat{p})$, donde $\hat{p}$ es la proporción muestral. Posteriormente se puede emplear el método de propagación de errores (ecuación 3.10 de la sección 3.3) para encontrar la incertidumbre en $f(\hat{p})$.

Ejemplo 4.14

En una muestra de 100 llantas nuevas de automóvil, se descubrió que siete de éstas tienen imperfecciones en el dibujo. Si se seleccionan aleatoriamente cuatro llantas nuevas y se instalan en un automóvil, estime la probabilidad de que ninguna de ellas tenga alguna imperfección, y determine la incertidumbre en la estimación.

Solución

Sea p la probabilidad de que una llanta no tenga imperfecciones. Se inicia al calcular la proporción muestral $\hat{p}$ y determinando su incertidumbre. La proporción muestral es $\hat{p} = 93/100 = 0.93$. La incertidumbre en $\hat{p}$ está dada por $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$. Se sustituye $n = 100$ y $\hat{p} = 0.93$ en p para obtener

$$\sigma_{\hat{p}} = \sqrt{\frac{(0.93)(0.07)}{100}} = 0.0255$$

Ahora la probabilidad de que ninguna de las cuatro llantas tenga alguna imperfección es p^4 . Lo anterior se estima con $\hat{p}^4 = 0.93^4 = 0.7481$. Se emplea la ecuación (3.10) para calcular la incertidumbre en $\hat{p}^4$:

$$\begin{aligned}\sigma_{\hat{p}^4} &\approx \left| \frac{d}{d\hat{p}} \hat{p}^4 \right| \sigma_{\hat{p}} \\ &= 4\hat{p}^3 \sigma_{\hat{p}} \\ &= 4(0.93)^3 (0.0255) \\ &= 0.082\end{aligned}$$

Ejercicios para la sección 4.2

- Sea $X \sim \text{Bin}(8, 0.4)$. Determine
 - $P(X = 2)$
 - $P(X = 4)$
 - $P(X < 2)$
 - $P(X > 6)$
 - μ_X
 - σ_X^2
- Se toma una muestra de cinco elementos de una población grande en la cual 10% de los elementos está defectuoso.
 - Determine la probabilidad de que ninguno de los elementos de la muestra esté defectuoso.
 - Determine la probabilidad de que sólo uno de ellos tenga defectos.
 - Determine la probabilidad de que uno o más de los elementos de la muestra estén defectuosos.
 - Determine la probabilidad de que menos de dos elementos de la muestra tenga defectos.
- Se lanza al aire una moneda diez veces.
 - ¿Cuál es la probabilidad de obtener exactamente tres veces “cara”?
 - Determine la media del número de caras obtenidas.
 - Determine la varianza del número de caras obtenidas.
 - Determine la desviación estándar del número de caras obtenidas.
- En un cargamento grande de llantas de automóvil, 5% tiene cierta imperfección. Se eligen aleatoriamente cuatro llantas para instalarlas en el automóvil.
 - ¿Cuál es la probabilidad de que ninguna de las llantas tenga imperfección?
 - ¿Cuál es la probabilidad de que sólo una de las llantas tenga imperfección?
 - ¿Cuál es la probabilidad de que una o más de las llantas tenga imperfección?
- En un patrón aleatorio de ocho bits utilizado para probar un microcircuito, cada bit tiene la misma probabilidad de ser 0 o 1. Suponga que los valores de los bits son independientes.
 - ¿Cuál es la probabilidad de que todos los bits sean 1?
 - ¿Cuál es la probabilidad de que exactamente tres de los bits sean 1?
 - ¿Cuál es la probabilidad de que al menos seis de los bits sean 1?
 - ¿Cuál es la probabilidad de que al menos dos de los bits sean 1?
- Un ingeniero que supervisa el control de calidad, selecciona una muestra aleatoria de 100 varillas de acero de la producción del día, descubre que 92 de ellas satisfacen las especificaciones.
 - Estime la proporción, de la producción de ese día, que satisface las especificaciones y determine la incertidumbre en la estimación.

- b) Estime el número de varillas que se debe seleccionar para reducir la incertidumbre a 1 por ciento.
7. En una muestra aleatoria de 100 partes ordenadas al vendedor A, 12 estaban defectuosas. En una muestra aleatoria de 200 partes ordenadas al vendedor B, diez estaban defectuosas.
- a) Estime la proporción de partes del vendedor A que están defectuosas y determine la incertidumbre en la estimación.
- b) Estime la proporción de partes del vendedor B que están defectuosas y encuentre la incertidumbre en la estimación.
- c) Estime la diferencia en las proporciones y determine la incertidumbre en la estimación.
8. De los elementos producidos en determinado proceso, 20% estaba defectuoso; de ellos, se puede reparar 60 por ciento.
- a) Determine la probabilidad de que un elemento elegido de forma aleatoria esté defectuoso y no se pueda reparar.
- b) Determine la probabilidad de que sólo dos de los 20 elementos seleccionados aleatoriamente esté defectuoso y no se pueda reparar.
9. De los pernos manufacturados por cierta aplicación, 90% satisface la longitud especificada y se puede utilizar inmediatamente, 6% está demasiado largo y sólo se puede usar después de que sea cortado, y 4% está demasiado corto y debe desecharse.
- a) Determine la probabilidad de que un perno seleccionado aleatoriamente se pueda utilizar (inmediatamente o después de ser cortados).
- b) Determine la probabilidad de que menos de nueve de una muestra de diez pernos se puedan utilizar (inmediatamente o después de ser cortados).
10. Una distribuidora recibe importante cargamento de componentes. A la empresa le gustaría aceptar el cargamento si 10% o menos de los componentes está defectuoso y rechazarlo si más de 10% presenta defecto. Se opta por seleccionar diez de éstos, y regresar el envío si más de uno tiene defectos.
- a) Si la proporción de pernos defectuosos en la muestra es de hecho 10%, ¿cuál es la probabilidad de que la distribuidora regrese el cargamento?
- b) Si la proporción de pernos defectuosos en la muestra es 20%, ¿cuál es la probabilidad de que la empresa regrese el cargamento?
- c) Si la proporción de pernos con defecto en la muestra es 2%, ¿cuál es la probabilidad de que la compañía regrese el cargamento?
- d) La distribuidora decide que aceptará el cargamento sólo si ninguno de los elementos de la muestra está defectuoso. ¿Cuál es el número mínimo de elementos que debe seleccionar si quiere tener una probabilidad no mayor a 0.01 de aceptar el cargamento si es que 20% de los componentes del cargamento tiene defectos?
11. Cierta cargamento viene con la garantía de que contiene no más de 15% de unidades defectuosas. Si la proporción de unidades defectuosas es mayor a 15%, aquél será regresado. Se extrae una muestra aleatoria de diez unidades. Sea X el número de unidades defectuosas en la muestra.
- a) Si, de hecho, 15% de las unidades en el cargamento está defectuoso (por lo que apenas el cargamento es aceptable), ¿a qué es igual $P(X \geq 7)$?
- b) Con base en la respuesta del inciso (a), si 15% de las unidades del cargamento está defectuoso, ¿siete piezas defectuosas en una muestra de diez es un número inusualmente grande?
- c) Si se descubre que siete de las diez unidades de la muestra está defectuoso, ¿esto sería una evidencia de que se debe regresar el cargamento? Explique.
- d) Si, de hecho, 15% de las unidades en el cargamento está defectuoso, ¿a qué es igual $P(X \geq 2)$?
- e) Con base en la respuesta al inciso (b), si 15% de las unidades del cargamento está defectuoso, ¿dos muestras defectuosas entre diez sería un número inusualmente grande?
- f) Si se descubre que dos de las diez unidades de la muestra están defectuosas, ¿ello sería una evidencia de que se debe regresar el cargamento? Explique.
12. Una compañía aseguradora ofrece descuentos a los propietarios de casas que deseen instalar detectores de humo en sus hogares. Un representante de la empresa afirma que 80% o más de los asegurados tiene detectores de humo. Usted toma una muestra aleatoria de ocho asegurados. Sea X el número de éstos en la muestra que tiene detectores de humo.
- a) Si exactamente 80% de los asegurados tiene detectores de humo (por lo que la afirmación del representante es verdadera, pero apenas), ¿a qué es igual $P(X \leq 1)$?
- b) Con base en la respuesta al inciso (a), si 80% de los asegurados tiene detectores de humo, ¿uno de aquéllos en una muestra de ocho sería un número inusualmente pequeño?

- c) Si se descubre que uno de los ocho asegurados de la muestra tenía un detector de humo, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
- d) Si exactamente 80% de los asegurados tiene detectores de humo, ¿a qué es igual $P(X \leq 6)$?
- e) Con base en la respuesta al inciso (d), si 80% de los asegurados tiene detectores de humo, ¿seis de aquéllos en una muestra de ocho sería un número inusualmente pequeño?
- f) Si usted descubre que seis de los ocho asegurados de la muestra tenían detectores de humo, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
13. Unas figurillas de porcelana se venden a 10 dólares si no tienen imperfección, y a 3 dólares si la presentan. Entre las figurillas de cierta compañía, 90% no tiene imperfecciones y 10% sí tiene. En una muestra de 100 figurillas ya vendidas, sea Y el ingreso ganado por su venta y X el número de éstas que no presenta imperfecciones.
- Expresa Y como una función de X .
 - Determine μ_Y .
 - Determine σ_Y .
14. El diseño de un sistema requiere la instalación de dos componentes idénticos. El sistema funcionará si al menos uno de los componentes lo hace. Un diseño alternativo necesita cuatro de estos componentes, y el sistema funcionará si al menos dos de los cuatro componentes funcionan. Si la probabilidad de que el componente funcione es de 0.9, y si los componentes funcionan de forma independiente, ¿qué diseño tiene la mayor probabilidad de funcionar?
16. Con referencia al ejemplo 4.14 estime la probabilidad de que sólo una de las cuatro llantas tenga alguna imperfección, y determine la incertidumbre en la estimación.
17. Si p es la probabilidad de éxito, a la cantidad $p/(1 - p)$ se le llama *oportunidad*. Ésta se estima en la investigación médica. El artículo "A Study of Twelve Southern California Communities with Differing Levels and Types of Air Pollution" (J. Peters, E. Avol y colaboradores, en *The American Journal of Respiratory and Critical Care Medicine*, 1999: 760-767) presenta una evaluación de la salud respiratoria de niños del sur de California. Suponga que a 88 infantes, en una muestra de 612 estudiados, se les diagnosticó bronquitis durante los últimos 12 meses.
- Estime la proporción p de niños a los que se les ha diagnosticado bronquitis y encuentre la incertidumbre en la estimación.
 - Estime la oportunidad y determine la incertidumbre en la estimación.

4.3 La distribución de Poisson

La distribución de Poisson se utiliza con frecuencia en el trabajo científico. Una manera de considerarla es como una aproximación de la distribución binomial cuando n es grande y p es pequeña. Esto último se muestra con un ejemplo.

Una masa contiene 10 000 átomos de una sustancia radiactiva. La probabilidad de que cierto átomo decaiga en un periodo de un minuto es de 0.0002. Sea X el número de átomos que decae en un minuto. Se puede considerar a cada átomo como un ensayo de Bernoulli, en los que el éxito ocurre si el átomo decae. Por tanto, X es el número de éxitos en 10 000 ensayos de Bernoulli independientes, cada uno con probabilidad de éxito de 0.0002, de tal forma que la distribución de X es $\text{Bin}(10\,000, 0.0002)$. La media de X es $\mu_X = (10\,000)(0.0002) = 2$.

Otra masa contiene 5 000 átomos y cada uno de éstos tiene probabilidad de 0.0004 de decaer en un intervalo de un minuto. Sea Y el número de átomos de esta masa que decae en un minuto. Siguiendo la lógica del párrafo anterior, $Y \sim \text{Bin}(5\,000, 0.0004)$ y $\mu_Y = (5\,000)(0.0004) = 2$.

En cada uno de estos casos, el número de ensayos n y la probabilidad de éxito p son diferentes, pero el número promedio de éxitos, que es igual al producto np , es el mismo. Ahora suponga que se quiere calcular la probabilidad de que sólo tres átomos decaigan en un

minuto para cada una de estas masas. Mediante la función de masa de probabilidad binomial, se calcula de la siguiente manera:

$$P(X = 3) = \frac{10\,000!}{3! \, 9\,997!} (0.0002)^3 (0.9998)^{9997} = 0.180465091$$

$$P(Y = 3) = \frac{5\,000!}{3! \, 4\,997!} (0.0004)^3 (0.9996)^{4997} = 0.180483143.$$

Estas probabilidades son casi iguales entre sí. Aunque a partir de la fórmula de la función de masa de probabilidad binomial esto no es obvio, cuando n es grande y p es pequeña la función de masa depende por completo de la media np , y muy pocos de los valores específicos de n y p . Por consiguiente, se puede aproximar la función de masa binomial con una cantidad que dependa sólo del producto np . Específicamente, si n es grande y p es pequeña, y $\lambda = np$, se puede demostrar mediante métodos avanzados que para todas las x ,

$$\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \approx e^{-\lambda} \frac{\lambda^x}{x!} \quad (4.8)$$

Esto conduce a la definición de una nueva función de probabilidad, denominada función de masa de probabilidad de Poisson, que se define mediante

$$p(x) = P(X = x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{si } x \text{ es un entero no negativo} \\ 0 & \text{de otro modo} \end{cases} \quad (4.9)$$

Si X es una variable aleatoria cuya función de masa de probabilidad está dada por la ecuación (4.9), entonces X sigue una **distribución de Poisson** con parámetro λ . La notación es $X \sim \text{Poisson}(\lambda)$.

Ejemplo

4.15

Si $X \sim \text{Poisson}(3)$, calcule $P(X = 2)$, $P(X = 10)$, $P(X = 0)$, $P(X = -1)$ y $P(X = 0.5)$.

Solución

Cuando se usa la función de masa de probabilidad (4.9), con $\lambda = 3$, se obtiene

$$P(X = 2) = e^{-3} \frac{3^2}{2!} = 0.2240$$

$$P(X = 10) = e^{-3} \frac{3^{10}}{10!} = 0.0008$$

$$P(X = 0) = e^{-3} \frac{3^0}{0!} = 0.0498$$

$$P(X = -1) = 0$$

debido a que -1 no es un entero no negativo,

$$P(X = 0.5) = 0$$

debido a que 0.5 no es un entero no negativo

Ejemplo

4.16

Si $X \sim \text{Poisson}(4)$, calcule $P(X \leq 2)$ y $P(X > 1)$.

Solución

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= e^{-4} \frac{4^0}{0!} + e^{-4} \frac{4^1}{1!} + e^{-4} \frac{4^2}{2!} \\ &= 0.0183 + 0.0733 + 0.1465 \\ &= 0.2381 \end{aligned}$$

Para encontrar $P(X > 1)$ se escribe

$$P(X > 1) = P(X = 2) + P(X = 3) + \dots$$

Esto último conduce a una suma infinita que es difícil de calcular. En vez de esto, se escribe

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \left(e^{-4} \frac{4^0}{0!} + e^{-4} \frac{4^1}{1!} \right) \\ &= 1 - (0.0183 + 0.0733) \\ &= 0.9084 \end{aligned}$$

Para las masas radiactivas descritas al inicio de esta sección, se utiliza la función de masa de Poisson para aproximar a $P(X = x)$ o $P(Y = x)$ sustituyendo a $\lambda = 2$ en la ecuación (4.9). La tabla 4.1 muestra que la aproximación es excelente.

TABLA 4.1 Ejemplo de una aproximación de Poisson a la función de masa de probabilidad binomial*

x	$P(X = x),$ $X \sim \text{Bin}(10\,000, 0.0002)$	$P(Y = x),$ $Y \sim \text{Bin}(5\,000, 0.0004)$	Aproximación de Poisson Poisson (2)
0	0.135308215	0.135281146	0.135335283
1	0.270670565	0.270670559	0.270670566
2	0.270697637	0.270724715	0.270670566
3	0.180465092	0.180483143	0.180447044
4	0.090223521	0.090223516	0.090223522
5	0.036082189	0.036074965	0.036089409
6	0.012023787	0.012017770	0.012029803
7	0.003433993	0.003430901	0.003437087
8	0.000858069	0.000856867	0.000859272
9	0.000190568	0.000190186	0.000190949

* Cuando n es grande y p pequeña, la función de masa de probabilidad Bin(n, p) también se aproxima mediante la función de masa de probabilidad de Poisson (λ) (ecuación 4.9), con $\lambda = np$. Aquí $X \sim \text{Bin}(10\,000, 0.0002)$ y $Y \sim \text{Bin}(5\,000, 0.0004)$, por lo que $\lambda = np = 2$, y la aproximación de Poisson es Poisson(2).

Resumen

Si $X \sim \text{Poisson}(\lambda)$, entonces

- X es una variable aleatoria discreta, cuyos posibles valores son enteros no negativos.
- El parámetro λ es una constante positiva.
- La función de masa de probabilidad de X es

$$p(x) = P(X = x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{si } x \text{ es un entero no negativo} \\ 0 & \text{de otro modo} \end{cases}$$

- La función de masa de probabilidad de Poisson se aproxima mucho a la función de masa de probabilidad binomial cuando n es grande, p es pequeña y $\lambda = np$.

Media y varianza de una variable aleatoria de Poisson

Para calcular la media y la varianza de una variable aleatoria de Poisson, se emplea la función de masa de probabilidad junto con las definiciones dadas por las ecuaciones (2.29) y (2.30) (de la sección 2.4). Al final de esta sección se muestran deducciones rigurosas de la media y de la varianza con este método. Aquí se presenta un enfoque intuitivo. Si $X \sim \text{Poisson}(\lambda)$ se puede considerar a X como una variable aleatoria binomial con n grande, p pequeña y $np = \lambda$. Dado que la media de una variable aleatoria binomial es np , se tiene que la media de una variable aleatoria de Poisson es λ . La varianza de una variable aleatoria binomial es $np(1 - p)$. Puesto que p es muy pequeña, se puede reemplazar $1 - p$ con 1, y concluir que la varianza de una variable aleatoria de Poisson es $np = \lambda$. Observe que esta última es igual a su media.

Resumen

- Si $X \sim \text{Poisson}(\lambda)$, entonces la media y la varianza de X están dadas por

$$\mu_X = \lambda \quad (4.10)$$

$$\sigma_X^2 = \lambda \quad (4.11)$$

La figura 4.3 muestra histogramas de probabilidad para las funciones de masa de probabilidad de Poisson(1) y de Poisson(10).

Uno de los primeros usos industriales de la distribución Poisson se aplicó en la fabricación de cervezas. Un paso fundamental en dicho proceso es la adición de la cultura de la levadura para preparar la malta para la fermentación. Se mantiene a las células vivas de levadura suspendidas en un medio líquido. Debido a que las células están vivas, su concentración en el medio cambia con el tiempo. Por tanto, antes de que se agregue la levadura, se necesita calcular la concentración de células de levadura por unidad en el volumen de la suspensión, para asegurarse de que se añadió la cantidad correcta.

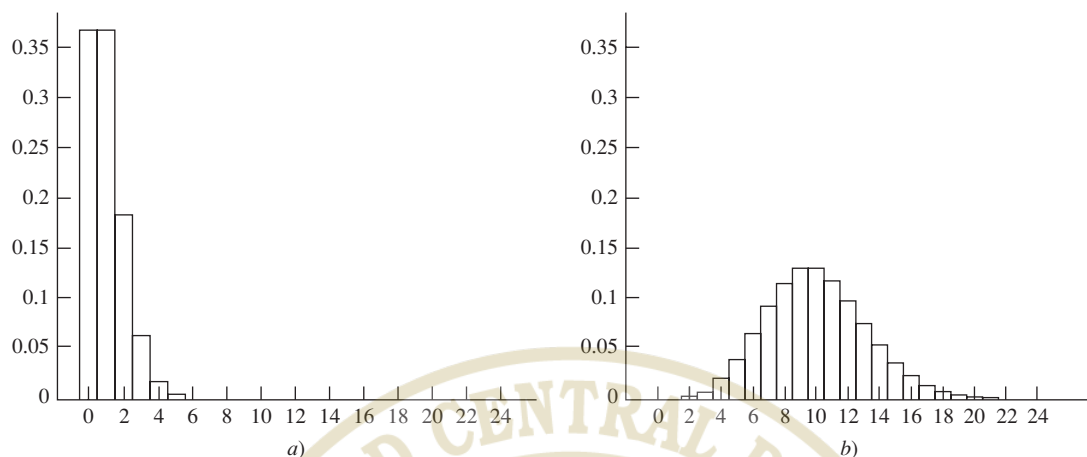


FIGURA 4.3 a) Histograma de probabilidad de Poisson(1), b) Histograma de probabilidad de Poisson(10).

Hasta principios del siglo xx, lo anterior consistía un problema para los fabricantes de cerveza. Ellos estimaban la concentración al extraer un pequeño volumen de la suspensión y contar las células de levadura en ésta utilizando un microscopio. Por supuesto que las estimaciones determinadas así estaban sujetas a la incertidumbre, pero nadie sabía cómo calcular ésta. Por tanto, nadie sabía en cuánto podía diferir la concentración de la muestra de la concentración verdadera.

William Sealy Gosset, un joven de 25 años, fue contratado por la compañía cervecera Guinness, de Dublín, Irlanda, y descubrió en 1904 que el número de células de levadura en el volumen de la suspensión de una muestra seguía una distribución de Poisson. Entonces desarrolló métodos para calcular la incertidumbre. El descubrimiento de Gosset no sólo le permitió a Guinness hacer un producto más consistente, sino que demostró que la distribución de Poisson puede tener aplicaciones importantes en muchas situaciones. Gosset quería publicar su resultado, pero sus jefes consideraron que su descubrimiento era información privada y se lo prohibieron. De todos modos Gosset lo publicó, pero para esconder este hecho a sus jefes, utilizó el seudónimo “Estudiante”.

En el ejemplo 4.17 se seguirá una lógica de razonamiento que conduce al resultado de “Estudiante”. Antes de esto, se mencionará que cuatro años después de publicar ese resultado, hizo otro descubrimiento que resolvió uno de los problemas pendientes más importantes de la estadística, y que ha tenido, desde entonces, una profunda influencia en los trabajos de casi todos los campos de la ciencia. Ese resultado se analiza en la sección 5.3.

Ejemplo

4.17

Unas partículas (por ejemplo, células de levadura) están suspendidas en un medio líquido con concentración de diez partículas por mL. Se agita por completo un volumen grande de la suspensión y después se extrae 1 mL. ¿Cuál es la probabilidad de que sólo se extraigan ocho partículas?

Solución

Siempre y cuando el volumen extraído sea una fracción pequeña del total, la solución a este problema no depende del volumen total de la suspensión, sino sólo de la concentración de partículas en éste. Sea V el volumen total de la suspensión, en mL. Entonces, el número total de partículas en la suspensión es $10V$. Considere a cada una de éstas como un ensayo de Bernoulli. Una partícula tiene “éxito” si es retirada. Ahora, se extrae 1 mL del total de V mL. Por consecuencia, la cantidad que se retirará representa $1/V$ del total, de ahí que cada partícula tenga una probabilidad de $1/V$ de ser retirada. Sea X el número de partículas extraídas. De este modo, X representa el número de éxitos en $10V$ ensayos de Bernoulli, cada uno con probabilidad de éxito de $1/V$. Por tanto, $X \sim \text{Bin}(10V, 1/V)$. Puesto que V es grande, $10V$ es grande y $1/V$ es pequeño. Por consiguiente, en una aproximación muy cercana, $X \sim \text{Poisson}(10)$. Se calcula $P(X = 8)$ con la función de masa de probabilidad Poisson: $P(X = 8) = e^{-10} 10^8 / 8! = 0.1126$.

En el ejemplo 4.17, λ tenía un valor de 10 ya que el número promedio de partículas en 1 mL de suspensión (el volumen retirado) era 10.

Ejemplo 4.18

Unas partículas están suspendidas en un medio líquido con concentración de seis partículas por mL. Se agita por completo un volumen grande de la suspensión, y después se extrae 3 mL. ¿Cuál es la probabilidad de que sólo se retiren 15 partículas?

Solución

Sea X el número de partículas extraídas. El número promedio de partículas en un volumen de 3 mL es 18. Entonces $X \sim \text{Poisson}(18)$. La probabilidad de que se extraigan sólo 15 partículas es

$$\begin{aligned} P(X = 15) &= e^{-18} \frac{18^{15}}{15!} \\ &= 0.0786 \end{aligned}$$

Observe que para que las soluciones de los ejemplos 4.17 y 4.18 sean correctas, es importante que la cantidad extraída de la suspensión no sea una fracción demasiado grande del total. Por ejemplo, si el volumen total en el ejemplo 4.18 fuera de 3 mL, de tal forma que se extrajera toda la cantidad, se tendría la certeza de que se retiraron las 18 partículas, por lo que la probabilidad de extraer 15 partículas sería igual a cero.

Ejemplo 4.19

La abuela hornea galletas de chispas de chocolates en grupos de 100. Ella agrega 300 chispas en la masa. Cuando las galletas están hechas, le ofrece una. ¿Cuál es la probabilidad de que su galleta no tenga chispas de chocolate?

Solución

Éste es otro caso de partículas en suspensión. Sea X el número de chispas en su galleta. La media del número de chispas es tres en cada galleta, de forma que $X \sim \text{Poisson}(3)$. De ahí que $P(X = 0) = e^{-3} 3^0 / 0! = 0.0498$.

Ejemplo

4.20

Los nietos de la abuela se han estado quejando de que ella es muy tacaña con las chispas de chocolate. La abuela acuerda agregar las suficientes chispas a la masa de tal forma que sólo 1% de las galletas no tendrá chispas de chocolate. ¿Cuántas chispas debe incluir en la masa de 100 galletas para lograr su propósito?

Solución

Sea n el número de chispas que se incluirá en una masa de 100 galletas, y X el número de chispas en su galleta. La media del número de chispas es $0.01n$ por cada galleta, por lo que $X \sim \text{Poisson}(0.01n)$. Se debe determinar el valor de n para el cual $P(X = 0) = 0.01$. Utilizando la función de masa de probabilidad Poisson($0.01n$)

$$\begin{aligned} P(X = 0) &= e^{-0.01n} \frac{(0.01n)^0}{0!} \\ &= e^{-0.01n} \end{aligned}$$

Al hacer $e^{-0.01n} = 0.01$, se obtiene que $n \approx 461$.

Los ejemplos 4.17 a 4.20 demuestran que para que las partículas se distribuyan uniformemente en un medio, el número de partículas que por casualidad está en una pequeña porción del medio sigue una distribución de Poisson. En estos ejemplos, las partículas eran reales y el medio era un espacio natural. Sin embargo, existen muchos casos en los que las “partículas” representan eventos y el medio es el tiempo. Anteriormente se vio tal ejemplo, en el que resulta ser que el número de eventos de decaimiento radiactivo en un intervalo fijo sigue una distribución de Poisson. Ahora se presenta otro.

Ejemplo

4.21

Suponga que el número de visitas a cierto sitio web durante un intervalo fijo sigue una distribución de Poisson. Suponga que la media de la razón de visitas es de cinco en cada minuto. Determine la probabilidad de que haya sólo 17 visitas en los siguientes tres minutos.

Solución

Sea X el número de visitas en tres minutos. La media del número de visitas en tres minutos es $(5)(3) = 15$, por lo que $X \sim \text{Poisson}(15)$. Utilizando la función de masa de probabilidad de Poisson(15),

$$\begin{aligned} P(X = 17) &= e^{-15} \frac{15^{17}}{17!} \\ &= 0.0847 \end{aligned}$$

Ejemplo

4.22

En el ejemplo 4.21, sea X el número de visitas en t minutos. Determine la función de masa de probabilidad de X , en función de t .

Solución

La media del número de visitas en t minutos es $5t$, por lo que $X \sim \text{Poisson}(5t)$. La función de masa de probabilidad de X es

$$p(x) = P(X = x) = \frac{e^{-5t}(5t)^x}{x!} \quad x = 0, 1, 2, \dots$$

Uso de la distribución de Poisson para estimar una razón

A menudo se realizan experimentos para estimar una razón λ que represente la media del número de eventos que ocurren en una unidad de tiempo o espacio. En estos experimentos se cuenta el número de eventos X que ocurre en t unidades, y se estima la razón λ con la cantidad $\hat{\lambda} = X/t$. (Observe que en virtud de que la cantidad X/t se utiliza para estimar λ , ésta se denota como $\hat{\lambda}$.) Si los números de eventos en intervalos disjuntos son independientes, y si no es posible que los eventos ocurran simultáneamente, entonces X sigue una distribución de Poisson. Al proceso que da como resultado dichos eventos se le denomina **proceso de Poisson**. Puesto que la media del número de eventos que ocurre en t unidades de tiempo o espacio es igual a λt , $X \sim \text{Poisson}(\lambda t)$.

Resumen

Sea λ la media del número de eventos que ocurre en una unidad de tiempo o espacio.
 Sea X el número de eventos que ocurre en t unidades de tiempo o espacio. Entonces si $X \sim \text{Poisson}(\lambda t)$, λ se estima con $\hat{\lambda} = X/t$.

Ejemplo

4.23

Una suspensión contiene partículas en una concentración desconocida de λ por mL. Se agita por completo la suspensión, y después se extraen 4 mL y se cuentan 17 partículas. Estime λ .

Solución

Sea $X = 17$ el número de partículas contadas y $t = 4$ mL el volumen extraído de la suspensión. Entonces $\hat{\lambda} = X/t = 17/4 = 4.25$ partículas por mL.

Incertidumbre en la razón estimada

Es importante darse cuenta que la razón estimada o concentración $\hat{\lambda}$ sólo representa una *estimación* de la verdadera razón o concentración λ . En general, $\hat{\lambda}$ no es igual a λ . Si se repitiera el experimento, probablemente el valor de $\hat{\lambda}$ sería diferente. En otras palabras, hay incertidumbre en $\hat{\lambda}$. Para que $\hat{\lambda}$ sea una estimación útil, se debe calcular su sesgo e incertidumbre. Los cálculos son similares a los de la proporción muestral presentados en la sección 4.2. Sea X el número de eventos contados en t unidades de tiempo o espacio, y suponga que $X \sim \text{Poisson}(\lambda t)$.

El sesgo es la diferencia $\mu_{\hat{\lambda}} - \lambda$. Dado que $\hat{\lambda} = X/t$, se tiene a partir de la ecuación (2.41) (de la sección 2.5) que

$$\begin{aligned}\mu_{\hat{\lambda}} &= \mu_{X/t} = \frac{\mu_X}{t} \\ &= \frac{\lambda t}{t} = \lambda\end{aligned}$$

Puesto que $\mu_{\hat{\lambda}} = \lambda$, $\hat{\lambda}$ no es sesgado.

La incertidumbre es la desviación estándar $\sigma_{\hat{\lambda}}$. Dado que $\hat{\lambda} = X/t$, se tiene a partir de la ecuación (2.43) (de la sección 2.5) que $\sigma_{\hat{\lambda}} = \sigma_X/t$. Debido a que $X \sim \text{Poisson}(\lambda t)$, se tiene a partir de la ecuación (4.11) que $\sigma_X = \sqrt{\lambda t}$. Por consecuencia,

$$\sigma_{\hat{\lambda}} = \frac{\sigma_X}{t} = \frac{\sqrt{\lambda t}}{t} = \sqrt{\frac{\lambda}{t}}$$

En la práctica, no se conoce el valor de λ , por lo que se puede aproximar con $\hat{\lambda}$.

Resumen

Si $X \sim \text{Poisson}(\lambda t)$ se estima la razón λ con $\hat{\lambda} = \frac{X}{t}$.

- $\hat{\lambda}$ no es sesgado.
- La incertidumbre en $\hat{\lambda}$ es

$$\sigma_{\hat{\lambda}} = \sqrt{\frac{\lambda}{t}} \quad (4.12)$$

En la práctica se sustituye $\hat{\lambda}$ por λ en la ecuación (4.12), ya que no se conoce λ .

Ejemplo

4.24

Se saca una muestra de 5 mL de una suspensión, y se cuenta 47 partículas. Estime la media del número de partículas por mL y encuentre la incertidumbre en la estimación.

Solución

El número de partículas contadas es $X = 47$. El volumen extraído es $t = 5$ mL. La media estimada del número de partículas por mL es

$$\hat{\lambda} = \frac{47}{5} = 9.4$$

La incertidumbre en la estimación es

$$\begin{aligned}\sigma_{\hat{\lambda}} &= \sqrt{\frac{\lambda}{t}} \\ &= \sqrt{\frac{9.4}{5}} \quad \text{aproximando } \lambda \text{ con } \hat{\lambda} = 9.4 \\ &= 1.4\end{aligned}$$

Ejemplo 4.25

La masa de cierta sustancia radiactiva emite partículas alfa a razón de λ partículas por segundo. Un físico cuenta 1 594 emisiones en 100 segundos. Estime λ y determine la incertidumbre en la estimación.

Solución

La estimación de λ es $\hat{\lambda} = 1\,594/100 = 15.94$ emisiones por segundo. La incertidumbre es

$$\begin{aligned}\sigma_{\hat{\lambda}} &= \sqrt{\frac{\lambda}{t}} \\ &= \sqrt{\frac{15.94}{100}} \quad \text{aproximando } \lambda \text{ con } \hat{\lambda} = 15.94 \\ &= 0.40\end{aligned}$$

Ejemplo 4.26

En el ejemplo 4.25, ¿durante cuántos segundos deben contarse las emisiones para reducir la incertidumbre a 0.3 emisiones por segundo?

Solución

Se desea encontrar el tiempo t para el cual $\sigma_{\hat{\lambda}} = \sqrt{\lambda/t} = 0.3$. Del ejemplo 4.25, $\hat{\lambda} = 15.94$. Al sustituir este valor por λ , se obtiene

$$\sigma_{\hat{\lambda}} = \sqrt{\frac{15.94}{t}} = 0.3$$

Al despejar t se tiene que $t = 177$ segundos.

Algunas veces se desea estimar una función $f(\lambda)$ de una razón de Poisson λ . Se estima $f(\lambda)$ con $f(\hat{\lambda})$. Después se puede emplear el método de propagación de errores (ecuación 3.10 de la sección 3.3) para encontrar la incertidumbre en $f(\hat{\lambda})$.

Ejemplo 4.27

El número de imperfecciones en una lámina de aluminio fabricada por determinado proceso sigue una distribución de Poisson. En una muestra de 100 m^2 de aluminio, se encuentran 200 imperfecciones. Estime la probabilidad de que un metro cuadrado de aluminio no tenga imperfecciones y determine la incertidumbre en la estimación.

Solución

Sea λ el número promedio de imperfecciones por metro cuadrado. Se iniciará calculando $\hat{\lambda}$ y su incertidumbre. Se ha observado que $X = 200$ imperfecciones en $t = 100 \text{ m}^2$ de aluminio. Por tanto, $\hat{\lambda} = 200/100 = 2.00$. La incertidumbre en $\hat{\lambda}$ es

$$\sigma_{\hat{\lambda}} = \sqrt{\frac{\lambda}{t}}$$

$$\begin{aligned}
 &= \sqrt{\frac{2}{100}} \quad \text{aproximando } \lambda \text{ con } \hat{\lambda} = 2 \\
 &= 0.1414
 \end{aligned}$$

Lo que se desea estimar es la probabilidad de que un metro cuadrado de aluminio no tenga imperfecciones. Primero se expresa la probabilidad como una función de λ . Con este propósito, Y representa el número de imperfecciones en 1 m² de lámina de aluminio. Entonces, $Y \sim \text{Poisson}(\lambda)$. Se desea estimar $P(Y = 0)$. Utilizando la función de masa de probabilidad de Poisson, esta probabilidad la da

$$P(Y = 0) = \frac{e^{-\lambda} \lambda^0}{0!} = e^{-\lambda}$$

Por tanto, la probabilidad de que un metro cuadrado no tenga imperfecciones se estima con $e^{-\hat{\lambda}} = e^{-2.00} = 0.1353$. Para encontrar la incertidumbre en la estimación, se emplea el método de propagación de errores (ecuación 3.10).

$$\begin{aligned}
 \sigma_{e^{-\hat{\lambda}}} &\approx \left| \frac{d}{d\hat{\lambda}} e^{-\hat{\lambda}} \right| \sigma_{\hat{\lambda}} \\
 &= \left| -e^{-\hat{\lambda}} \right| \sigma_{\hat{\lambda}} \\
 &= e^{-2.00} (0.1414) \\
 &= 0.0191
 \end{aligned}$$

Para el caso de las partículas en suspensión, o eventos de decaimiento radiactivo, se conocen los principios fundamentales de la física que rigen estos procesos, donde basándose en los primeros principios, podría demostrarse que la distribución del número de eventos es de Poisson. Existen muchos casos en los que la evidencia empírica sugiere que la distribución de Poisson es adecuada, pero las leyes que rigen los procesos no son comprendidas bien para realizar una posible deducción rigurosa. Entre los ejemplos está el número de visitas a un sitio web, el número de accidentes de tráfico en una intersección y el número de árboles en una sección del bosque.

Deducción de la media y la varianza de una variable aleatoria de Poisson

Sea $X \sim \text{Poisson}(\lambda)$. Se demostrará que $\mu_X = \lambda$ y $\sigma_X^2 = \lambda$. Utilizando la definición de la media poblacional para una variable aleatoria discreta (ecuación 2.29 de la sección 2.4):

$$\begin{aligned}
 \mu_X &= \sum_{x=0}^{\infty} x P(X = x) \\
 &= \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} \\
 &= (0)(e^{-\lambda}) \left(\frac{\lambda^0}{0!} \right) + \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!}
 \end{aligned}$$

$$\begin{aligned}
&= 0 + \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-1)!} \\
&= \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-1)!} \\
&= \lambda \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!}
\end{aligned}$$

Ahora la sumatoria $\sum_{x=0}^{\infty} e^{-\lambda} \lambda^x / x!$ es la suma de la función de masa de probabilidad de Poisson(λ) sobre todos sus valores posibles. Por tanto, $\sum_{x=0}^{\infty} e^{-\lambda} \lambda^x / x! = 1$, por lo que,

$$\mu_X = \lambda$$

Se emplea la ecuación (2.31) (de la sección 2.4) para demostrar que $\sigma_X^2 = \lambda$.

$$\sigma_X^2 = \sum_{x=0}^{\infty} x^2 e^{-\lambda} \frac{\lambda^x}{x!} - \mu_X^2 \quad (4.13)$$

Al sustituir $x(x-1) + x$ para x^2 y λ para μ_X en la ecuación (4.13), se obtiene

$$\sigma_X^2 = \sum_{x=0}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} - \lambda^2 \quad (4.14)$$

Ahora, $x(x-1) = 0$ si $x = 0$ o 1 , y $\sum_{x=0}^{\infty} x e^{-\lambda} \lambda^x / x! = \mu_X = \lambda$. Por consiguiente, se puede comenzar por sumar el lado derecho de la ecuación (4.14) en $x = 2$, y sustituir λ por $\sum_{x=0}^{\infty} x e^{-\lambda} \lambda^x / x!$. Se obtiene

$$\begin{aligned}
\sigma_X^2 &= \sum_{x=2}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} + \lambda - \lambda^2 \\
&= \sum_{x=2}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-2)!} + \lambda - \lambda^2 \\
&= \lambda^2 \sum_{x=2}^{\infty} e^{-\lambda} \frac{\lambda^{x-2}}{(x-2)!} + \lambda - \lambda^2 \\
&= \lambda^2 \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} + \lambda - \lambda^2 \\
&= \lambda^2(1) + \lambda - \lambda^2 \\
&= \lambda
\end{aligned}$$

Ejercicios para la sección 4.3

1. Sea $X \sim \text{Poisson}(4)$. Determine
 - a) $P(X = 1)$
 - b) $P(X = 0)$
 - c) $P(X < 2)$
 - d) $P(X > 1)$
 - e) μ_X
 - f) σ_X
2. La concentración de partículas en una suspensión es 2 por mL. Se agita por completo la concentración, y posteriormente se extraen 3 mL. Sea X el número de partículas que son retiradas. Determine
 - a) $P(X = 5)$
 - b) $P(X \leq 2)$
 - c) $P(X > 1)$
 - d) μ_X
 - e) σ_X
3. Suponga que 0.03% de los contenedores plásticos producidos en cierto proceso tiene pequeños agujeros que los dejan inservibles. X representa el número de contenedores en una muestra aleatoria de 10 000 que tienen este defecto. Determine
 - a) $P(X = 3)$
 - b) $P(X \leq 2)$
 - c) $P(1 \leq X < 4)$
 - d) μ_X
 - e) σ_X
4. Uno de cada 5 000 individuos en una población porta cierto gen defectuoso. Se estudia una muestra aleatoria de 1 000 individuos.
 - a) ¿Cuál es la probabilidad de que sólo uno de los individuos de la muestra porte el gen?
 - b) ¿Cuál es la probabilidad de que ninguno sea portador?
 - c) ¿Cuál es la probabilidad de que más de dos individuos porte el gen?
 - d) ¿Cuál es la media del número de individuos de la muestra que porta el gen?
 - e) ¿Cuál es la desviación estándar del número de individuos portadores de gen?
5. El número de mensajes recibidos por el tablero computado de anuncios es una variable aleatoria de Poisson con una razón media de ocho mensajes por hora.
 - a) ¿Cuál es la probabilidad de que se reciban cinco mensajes en una hora?
 - b) ¿Cuál es la probabilidad de que se reciban diez mensajes en 1.5 horas?
 - c) ¿Cuál es la probabilidad de que se reciban menos de tres mensajes en $1\frac{1}{2}$ horas?
6. Cierta tipo de tablero de circuitos contiene 300 diodos. Cada uno tiene una probabilidad $p = 0.002$ de fallar.
 - a) ¿Cuál es la probabilidad de que fallen exactamente dos diodos?
 - b) ¿Cuál es la media del número de diodos que falla?
 - c) ¿Cuál es la desviación estándar del número de diodos que falla?
 - d) Un tablero funciona si ninguno de sus diodos falla. ¿Cuál es la probabilidad de que funcione un tablero?
 - e) Se envían cinco tableros a un cliente. ¿Cuál es la probabilidad de que cuatro o más de ellos funcione?
7. Una variable aleatoria X tiene una distribución binomial y una variable aleatoria Y tiene una distribución de Poisson. Tanto X como Y tienen medias iguales a 3. ¿Es posible determinar qué variable aleatoria tiene la varianza más grande? Elija una de las siguientes respuestas:
 - i) Sí, X tiene la varianza más grande.
 - ii) Sí, Y tiene la varianza más grande.
 - iii) No, se necesita conocer el número de ensayos, n , para X .
 - iv) No, se necesita conocer la probabilidad de éxito, p , para X .
 - v) No, se necesita conocer el valor de λ para Y .
8. Una química desea estimar la concentración de partículas que hay en determinada suspensión. Ella extrae 3 mL de la suspensión y cuenta 48 partículas. Estime la concentración de partículas por mL y determine la incertidumbre en la estimación.
9. Una microbióloga quiere estimar la concentración de cierto tipo de bacteria en una muestra de agua tratada. Ella pone una muestra de 0.5 mL de agua tratada en el vidrio del microscopio y descubre 39 bacterias. Estime la concentración

de bacterias por mL, en esta agua tratada, y determine la incertidumbre en la estimación.

10. La abuela está probando una nueva receta de pan de pasas. En cada hornada de la masa de pan salen tres hogazas, y cada una tiene 20 rebanadas de pan.
 - a) Si ella agrega 100 pasas a una hornada de masa, ¿cuál es la probabilidad de que una rebanada de pan elegida aleatoriamente no tenga pasas?
 - b) Si ella agrega 200 pasas a una hornada de masa, ¿cuál es la probabilidad de que una rebanada de pan elegida aleatoriamente tenga cinco pasas?
 - c) ¿Cuántas pasas debe agregar para que la probabilidad de que una rebanada elegida de forma aleatoria no tenga pasas sea 0.01?
11. Mamá y la abuela están horneando, cada una, galletas de chispas de chocolate. Cada una le da dos galletas. Una de las galletas de mamá tiene 14 chispas de chocolate y la otra tiene 11. Las galletas de la abuela tienen seis y ocho chispas.
 - a) Estime la media del número de chispas en una de las galletas de mamá.
 - b) Estime la media del número de chispas en una de las galletas de la abuela.
 - c) Determine la incertidumbre en la estimación de las galletas de mamá.
 - d) Determine la incertidumbre en la estimación de las galletas de la abuela.
 - e) Estime cuántas chispas más en promedio tiene una galleta de mamá en comparación con una galleta de la abuela. Determine la incertidumbre en la estimación.
12. Usted ha recibido una masa radiactiva de la que se afirma tiene una media de la razón de decaimiento de al menos una partícula por segundo. Si la media de la razón de decaimiento es menor a una por segundo, usted puede regresar el producto para un reembolso. Sea X el número de eventos de decaimiento que se produce en diez segundos.
 - a) Si la media de la razón de decaimiento es exactamente de una por segundo (de tal forma que la afirmación es verdad, pero apenas), ¿a qué es igual $P(X \leq 1)$?
 - b) Con base en la respuesta del inciso a), si la razón de decaimiento promedio es de una partícula por segundo, ¿un evento en diez segundos sería un número inusualmente pequeño?
 - c) Si usted encuentra un evento de decaimiento en diez segundos, ¿esto sería una evidencia de que debe regresarse el producto? Explique.
 - d) Si la media de la razón de decaimiento es sólo de una por segundo, ¿a qué es igual $P(X \leq 8)$?
 - e) Con base en la respuesta del inciso (d), si la razón de decaimiento promedio es de una partícula por segundo, ¿ocho eventos en diez segundos sería un número inusualmente pequeño?
 - f) Si cuenta ocho eventos de decaimiento en diez segundos, ¿esto sería una evidencia de que debe regresarse el producto? Explique.
13. Alguien afirma que cierta suspensión contiene al menos siete partículas por mL. Extrae una muestra de 1 mL de la solución. Sea X el número de partículas en la muestra.
 - a) Si el número promedio de partículas es exactamente siete por mL (de manera que la afirmación es verdad, pero apenas), ¿a qué es igual $P(X \leq 1)$?
 - b) Con base en la respuesta del inciso (a), si la suspensión contiene siete partículas por mL, ¿una partícula en una muestra de 1 mL sería un número inusualmente pequeño?
 - c) Si encuentra una partícula en la muestra, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
 - d) Si la media del número de partículas es exactamente 7 por mL, ¿a qué es igual $P(X \leq 6)$?
 - e) Con base en la respuesta del inciso (d), si la suspensión contiene siete partículas por mL, ¿seis partículas en una muestra de 1 mL sería un número inusualmente pequeño?
 - f) Si cuenta seis partículas en la muestra, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
14. Un físico desea estimar la razón de emisiones de partículas alfa provenientes de cierta fuente. Él hace dos cuentas. Primero mide la razón fondo contando el número de partículas que hay durante 100 segundos en ausencia de la fuente. Cuenta 36 emisiones de fondo. Después, con la fuente presente, cuenta 324 emisiones en 100 segundos. Esto último representa la suma de las emisiones de la fuente más las emisiones de fondo.
 - a) Estime la razón de fondo, en emisiones por segundo, y determine la incertidumbre en la estimación.
 - b) Estime la suma de la fuente más la razón de fondo, en emisiones por segundo, y determine la incertidumbre en la estimación.
 - c) Estime la razón de emisiones provenientes de la fuente en partículas por segundo, y determine la incertidumbre en la estimación.
 - d) ¿Qué da como resultado una menor incertidumbre al estimar la razón de emisiones provenientes de la fuente: (1) contar las partículas de fondo sólo durante 150 se-

gundos, así como las partículas de fondo, más las de la fuente durante 150 segundos, o (2) contar las partículas de fondo durante 100 segundos, así como las de la fuente más las de fondo durante 200 segundos? Calcule la incertidumbre en cada caso.

- e) ¿Es posible reducir la incertidumbre a 0.03 partículas por segundo si se mide la razón de fondo sólo durante

100 segundos? Si es así, ¿cuánto tiempo puede medirse las partículas de la fuente más las de fondo? Si no, explique por qué.

15. Con referencia al ejemplo 4.27, estime la probabilidad de que en 1 m² de aluminio haya una imperfección y determine la incertidumbre en la estimación.

4.4 Algunas otras distribuciones discretas

En esta sección se analizan diversas distribuciones discretas que son útiles en varias situaciones.

La distribución hipergeométrica

Cuando una población finita contiene dos tipos de unidades, que pueden ser denominados como éxitos y fracasos, y se extrae una muestra aleatoria simple de la población, cada unidad representa un ensayo de Bernoulli. A medida que se selecciona cada unidad, la proporción de éxitos en la población restante disminuye o aumenta, dependiendo si la unidad extraída es un éxito o fracaso. Por esta razón, los ensayos no son independientes, de ahí que el número de éxitos en la muestra no siga una distribución binomial. En su lugar, la distribución que describe adecuadamente el número de éxitos en esta situación se llama **distribución hipergeométrica**.

Como ejemplo, suponga que se tiene un lote de 20 unidades que contiene seis que están defectuosas, y que se extrae aleatoriamente cinco unidades de este lote. Sea X el número de unidades defectuosas en la muestra. Se calculará $P(X = 2)$. Con este propósito, primero se cuenta el número total de los grupos diferentes de cinco unidades que puede extraerse de la población de 20. (Se hará referencia a cada grupo de cinco unidades como combinación.) El número de combinaciones de cinco unidades es el número de muestras diferentes que se puede extraer, y cada una es igualmente probable. Después se determinará cuántas de estas combinaciones contienen exactamente dos defectuosas. La probabilidad de que una combinación de cinco unidades contenga sólo dos defectuosas es el cociente

$$P(X = 2) = \frac{\text{número de combinaciones de cinco unidades que contienen dos defectuosas}}{\text{número de combinaciones de cinco unidades que pueden seleccionarse entre 20}}$$

En general, el número de combinaciones de k unidades que se puede elegir de un grupo de n unidades se denota por $\binom{n}{k}$ y es igual a (véase la deducción de la ecuación 2.12 en la sección 2.2)

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Por tanto, el número de combinaciones de cinco unidades que se puede elegir entre 20 es

$$\binom{20}{5} = \frac{20!}{5!(20-5)!} = 15\,504$$

Para determinar el número de combinaciones de cinco que contienen sólo dos defectuosas, se describe la construcción de dicha combinación como una secuencia de dos operaciones. Primero, se seleccionan dos unidades de las seis defectuosas; segundo, se seleccionan tres unidades de las 14 no defectuosas. El número de combinaciones de dos unidades seleccionadas entre seis es

$$\binom{6}{2} = \frac{6!}{2!(6-2)!} = 15$$

y el número de combinaciones de tres unidades elegidas de 14 es

$$\binom{14}{3} = \frac{14!}{3!(14-3)!} = 364$$

Por tanto, el número total de combinaciones de cinco unidades que puede componerse de dos defectuosas y tres no defectuosas es el producto: $\binom{6}{2} \binom{14}{3} = (15)(364) = 5\,460$ (ésta es una aplicación del principio fundamental de conteo; véase la sección 2.2 para un análisis más detallado). Se concluye que

$$\begin{aligned} P(X=2) &= \frac{\binom{6}{2} \binom{14}{3}}{\binom{20}{5}} \\ &= \frac{5\,460}{15\,504} \\ &= 0.3522 \end{aligned}$$

Para calcular $P(X=2)$ en el ejemplo anterior, fue necesario conocer el número de unidades en la población (20), el de unidades defectuosas en la población (6) y el de unidades extraídas (5). La función de masa de probabilidad de la variable aleatoria X se determina al utilizar estos tres parámetros. Específicamente, X tiene una distribución hipergeométrica con los parámetros 20, 6 y 5, ello se denota mediante $X \sim H(20, 6, 5)$. Ahora se generalizará esta idea.

Resumen

Suponga una población finita que contiene N unidades, de ellas R son clasificadas como éxitos y $N - R$ como fracasos. Suponga que se extrae n unidades de esta población, y sea X el número de éxitos en la muestra. Entonces X sigue la distribución hipergeométrica con los parámetros N , R y n , que se puede denotar como $X \sim H(N, R, n)$.

La función de masa de probabilidad de X es

$$p(x) = P(X=x) = \begin{cases} \frac{\binom{R}{x} \binom{N-R}{n-x}}{\binom{N}{n}} & \text{máx}(0, R+n-N) \leq x \leq \text{mín}(n, R) \\ 0 & \text{de otro modo} \end{cases} \quad (4.15)$$

Ejemplo

4.28

De 50 edificios en un parque industrial, 12 no cumplen el código eléctrico. Si se seleccionan aleatoriamente diez edificios para inspeccionarlos, ¿cuál es la probabilidad de que exactamente tres de los diez no cumplan el código?

Solución

Sea X el número de edificios seleccionados que violan el código. Entonces, $X \sim H(50, 12, 10)$. Se debe determinar $P(X = 3)$. Utilizando la ecuación (4.15),

$$\begin{aligned} P(X = 3) &= \frac{\binom{12}{3} \binom{38}{7}}{\binom{50}{10}} \\ &= \frac{(220)(12\,620\,256)}{10\,272\,278\,170} \\ &= 0.2703 \end{aligned}$$

Media y varianza de la distribución hipergeométrica

En el recuadro siguiente se presenta la media y varianza de la distribución hipergeométrica. Se omiten sus deducciones.

Si $X \sim H(N, R, n)$, entonces

$$\mu_X = \frac{nR}{N} \quad (4.16)$$

$$\sigma_X^2 = n \left(\frac{R}{N} \right) \left(1 - \frac{R}{N} \right) \left(\frac{N-n}{N-1} \right) \quad (4.17)$$

Ejemplo

4.29

Con referencia al ejemplo 4.28 encuentre la media y la varianza de X .

Solución

$X \sim H(50, 12, 10)$, por lo que

$$\mu_X = \frac{(10)(12)}{50}$$

$$= 2.4000$$

$$\sigma_X^2 = (10) \left(\frac{12}{50} \right) \left(1 - \frac{12}{50} \right) \left(\frac{50-10}{50-1} \right)$$

$$= 1.4890$$

Comparación con la distribución binomial

Una población de tamaño N contiene R éxitos y $N - R$ fracasos. Imagine que se toma una muestra de n unidades de esta población con reemplazo; es decir, cada unidad de la muestra se regresará a la población después de ser extraída. Entonces, las unidades de la muestra son resultado de una secuencia de ensayos de Bernoulli independientes, y el número de los éxitos X en la muestra tiene una distribución binomial con n ensayos y probabilidad de éxito $p = R/N$.

En la práctica, rara vez se extraen muestras con reemplazo, debido a que no es necesario extraer la misma unidad dos veces. En su lugar, el muestreo se realiza sin reemplazo, en el cual cada unidad es eliminada de la población después de que es extraída. Entonces, las unidades de la muestra son resultado de ensayos de Bernoulli dependientes, ya que la población cambia conforme se extrae cada unidad. Por esta razón, la distribución del número de éxitos, X , es $H(N, R, n)$ en vez de $\text{Bin}(n, R/N)$.

Cuando el tamaño muestral n es pequeño en comparación con el tamaño de la población N (es decir, no mayor a 5%), la diferencia entre el muestreo con o sin reemplazo es poca, y la distribución binomial $\text{Bin}(n, R/N)$ es una buena aproximación de la distribución hipergeométrica $H(N, R, n)$. Observe que la media de $H(N, R, n)$ es nR/N , la misma que la de $\text{Bin}(n, R/N)$. Esto último indica que sea que realice la muestra con o sin reemplazo, la proporción de éxitos de la muestra en promedio es la misma que la de éxitos de la población. La varianza de $\text{Bin}(n, R/N)$ es $n(R/N)(1 - R/N)$ y la varianza de $H(N, R, n)$ se obtiene al multiplicar esto por el factor $(N - n)/(N - 1)$. Observe que cuando n es pequeña en relación con N , este factor se aproxima a 1.

Distribución geométrica

Suponga que se lleva a cabo una secuencia de ensayos de Bernoulli independientes, cada uno con la misma probabilidad de éxito p . Sea X el número de experimentos hasta incluir el primer éxito. Por tanto, X es una variable aleatoria discreta, la cual tiene una **distribución geométrica** con parámetro p . Se expresa como $X \sim \text{Geom}(p)$.

Ejemplo 4.30

Una prueba de resistencia de soldadura consiste en poner carga en uniones soldadas hasta que se dé una ruptura. Para cierto tipo de soldadura, 80% de las rupturas ocurre en la propia soldadura, mientras que otro 20% se da en las vigas. Se prueba cierto número de soldaduras. Sea X el número de pruebas, incluyendo la primera prueba que da como resultado la ruptura de la viga. ¿Cuál es la distribución de X ?

Solución

Cada prueba es un ensayo de Bernoulli, con un éxito definido como la ruptura de una viga. Por consiguiente, la probabilidad de éxito es $p = 0.2$. El número de ensayos incluyendo al primer éxito tiene una distribución geométrica con parámetro $p = 0.2$. Por consecuencia, $X \sim \text{Geom}(0.2)$.

Ejemplo**4.31**

Con referencia al ejemplo 4.30, determine $P(X = 3)$.

Solución

El evento $X = 3$ ocurre cuando los primeros dos ensayos resultan en fracaso y el tercer experimento en éxito. De ahí que

$$\begin{aligned} P(X = 3) &= P(\text{FFS}) \\ &= (0.8)(0.8)(0.2) \\ &= 0.128 \end{aligned}$$

El resultado del ejemplo 4.31 se puede generalizar para generar la función de masa de probabilidad de una variable aleatoria geométrica.

Si $X \sim \text{Geom}(p)$, entonces la función de masa de probabilidad de X es

$$p(x) = P(X = x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \dots \\ 0 & \text{de otro modo} \end{cases}$$

Media y varianza de una distribución geométrica

En el siguiente recuadro se muestra la media y varianza de la distribución geométrica. Sus deducciones requieren de la manipulación de series infinitas y se omiten.

Si $X \sim \text{Geom}(p)$, entonces

$$\mu_X = \frac{1}{p} \quad (4.18)$$

$$\sigma_X^2 = \frac{1-p}{p^2} \quad (4.19)$$

Ejemplo**4.32**

Con referencia al ejemplo 4.30, sea X el número de pruebas hechas incluyendo la primera ruptura de la viga. Determine la media y la varianza de X .

Solución

Puesto que $X \sim \text{Geom}(p)$, $\mu_X = 1/0.2 = 5$, y $\sigma_X^2 = (1 - 0.2)/(0.2^2) = 20$.

La distribución binomial negativa

La distribución binomial negativa constituye una extensión de la distribución geométrica. Sea r un entero positivo. Suponga que se realizan ensayos de Bernoulli independientes, cada uno

con probabilidad de éxito p , y X representa el número de ensayos hasta incluir al r -ésimo éxito. Por consecuencia, X tiene una **distribución binomial negativa** con parámetros r y p . Se expresa como $X \sim \text{NB}(r, p)$.

Ejemplo

4.33

(Continuación del ejemplo 4.30.) En una prueba de fuerza de soldadura, 80% de las pruebas da como resultado ruptura de la soldadura, mientras que otro 20% propicia ruptura de la viga. Sea X el número de pruebas incluyendo la tercera ruptura de la viga. ¿Cuál es la distribución de X ? Determine $P(X = 8)$.

Solución

Dado que X representa el número de ensayos incluyendo al tercer éxito, y puesto que la probabilidad de éxito es $p = 0.2$, $X \sim \text{NB}(3, 0.2)$. Se calculará $P(X = 8)$ y el método de cálculo conducirá a una deducción de la función de masa de probabilidad de una variable aleatoria binomial negativa. Debido a que $X \sim \text{NB}(3, 0.2)$, el evento $X = 8$ significa que el tercer éxito se presentó en el octavo ensayo. Otra forma de decir esto es que hubo exactamente dos éxitos en los primeros siete ensayos y el octavo ensayo fue un éxito. En virtud de que todos los experimentos son independientes, se tiene que

$$P(X = 8) = P(\text{sólo dos éxitos en los primeros siete ensayos})P(\text{éxitos en el octavo ensayo})$$

Ahora el número de éxitos en los primeros siete ensayos tiene una distribución $\text{Bin}(7, 0.2)$, por lo que

$$P(\text{exactamente dos éxitos en los primeros siete ensayos}) = \binom{7}{2} (0.2)^2 (0.8)^5$$

La probabilidad de que el octavo ensayo (o cualquier otro experimento) dé como resultado un éxito es 0.2. Por tanto,

$$\begin{aligned} P(X = 8) &= \binom{7}{2} (0.2)^2 (0.8)^5 (0.2) \\ &= \binom{7}{2} (0.2)^3 (0.8)^5 \\ &= 0.05505 \end{aligned}$$

Se generaliza el resultado del ejemplo 4.33 para construir la función de masa de probabilidad de una variable aleatoria binomial negativa.

Si $X \sim \text{NB}(r, p)$, entonces la función de masa de probabilidad de X es

$$p(x) = P(X = x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & x = r, r+1, \dots \\ 0 & \text{de otro modo} \end{cases}$$

Observe que el valor posible más pequeño de X es r , ya que se requiere al menos r ensayos para producir r éxitos. Observe que cuando $r = 1$, la distribución binomial negativa es la misma que la distribución geométrica. En símbolos, $NB(1, p) = \text{Geom}(p)$.

Una variable aleatoria binomial negativa constituye una suma de variables aleatorias geométricas

Suponga que una secuencia de ocho ensayos de Bernoulli independientes, cada uno con probabilidad de éxito p , aparece de la siguiente manera:

F F S F S F F S

Si X es el número de ensayos incluyendo al tercer éxito, entonces $X \sim NB(3, p)$, y para esta secuencia de experimentos, $X = 8$. Se denota el número de ensayos incluyendo al primer éxito por Y_1 . Para esta secuencia, $Y_1 = 3$, pero en general $Y_1 \sim \text{Geom}(p)$. Ahora se comienza a contar, empezando desde el primer ensayo después del primer éxito, incluyendo al segundo éxito. Este número de ensayos se representa con Y_2 . Para esta secuencia $Y_2 = 2$, pero en general $Y_2 \sim \text{Geom}(p)$. Finalmente, se cuenta el número de ensayos, a partir del primer ensayo después del segundo éxito, incluyendo al tercero. Y_3 denota el número de ensayos. Para esta secuencia $Y_3 = 3$, pero nuevamente, en general, $Y_3 \sim \text{Geom}(p)$. Es claro que $X = Y_1 + Y_2 + Y_3$. Más aún, puesto que los experimentos son independientes, Y_1 , Y_2 y Y_3 también lo son. Esto demuestra que si $X \sim NB(3, p)$, entonces X es la suma de tres variables aleatorias independientes $\text{Geom}(p)$. Este resultado se puede generalizar para cualquier entero positivo r .

Resumen

Si $X \sim NB(r, p)$, entonces

$$X = Y_1 + \cdots + Y_r$$

donde $Y_1, \dots, Y_r$ son variables aleatorias independientes, cada una con distribución $\text{Geom}(p)$.

Media y varianza de la distribución negativa binomial

Si $X \sim NB(r, p)$, entonces $X = Y_1 + \cdots + Y_r$, donde $Y_1, \dots, Y_r$ son variables aleatorias independientes, cada una con distribución $\text{Geom}(p)$. De ahí que la media de X sea la suma de las medias de las Y , y la varianza de X sea la suma de las varianzas. Cada Y_i tiene una media $1/p$ y varianza $(1 - p)/p^2$. Por consiguiente, $\mu_X = r/p$ y $\sigma_X^2 = r(1 - p)/p^2$.

Resumen

Si $X \sim NB(r, p)$, entonces

$$\mu_X = \frac{r}{p} \quad (4.20)$$

$$\sigma_X^2 = \frac{r(1 - p)}{p^2} \quad (4.21)$$

Ejemplo**4.34**

Con referencia al ejemplo 4.33, determine la media y la varianza de X , donde X representa el número de pruebas incluyendo la tercera ruptura de la viga.

Solución

Dado que $X \sim \text{NB}(3, 0.2)$, se tiene que

$$\mu_X = \frac{3}{0.2} = 15$$

$$\sigma_X^2 = \frac{3(1 - 0.2)}{0.2^2} = 60$$

Distribución multinomial

Un ensayo de Bernoulli representa un proceso que tiene dos posibles resultados. Una generalización del ensayo de Bernoulli constituye el **experimento multinomial**, que es un proceso con k resultados, donde $k \geq 2$. Por ejemplo, el lanzamiento de un dado es un experimento multinomial, con seis posibles resultados: 1, 2, 3, 4, 5, 6. Cada resultado de un experimento multinomial tiene una probabilidad de ocurrir. Las probabilidades de los k resultados se representan como $p_1, \dots, p_k$. Por ejemplo, en el lanzamiento de un dado, $p_1 = p_2 = \dots = p_6 = 1/6$.

Ahora suponga que se realizan n experimentos independientes multinomiales, cada uno con los mismos k resultados posibles y con las mismas probabilidades $p_1, \dots, p_k$. Se enumeran los resultados 1, 2, $\dots$, k . Para cada resultado i , X_i representa el número de ensayos que tiene ese resultado. Entonces $X_1, \dots, X_k$ son variables aleatorias discretas. La colección $X_1, \dots, X_k$ tiene una **distribución multinomial** con parámetros $n, p_1, \dots, p_k$. Se expresa como $X_1, \dots, X_k \sim \text{MN}(n, p_1, \dots, p_k)$. Observe que toda la colección $X_1, \dots, X_k$ sigue la distribución multinomial, en vez de sólo una X_i .

Ejemplo**4.35**

Se inspeccionan las unidades producidas por una línea de ensamble, y se clasifica cada una como correspondiente (aceptable), inferior o rechazada. En general, 70% de las unidades son correspondientes, 20% son inferiores y 10% rechazadas. Suponga que se selecciona de forma independiente y aleatoria cuatro unidades. Sean X_1, X_2, X_3 los números entre los cuatro eventos correspondientes inferiores y rechazados, respectivamente. ¿Cuál es la distribución de X_1, X_2, X_3 ?

Solución

Cada unidad es un ensayo multinomial con tres posibles resultados: correspondiente, inferior y rechazada. Las probabilidades asociadas con los resultados son $p_1 = 0.7, p_2 = 0.2$ y $p_3 = 0.1$. Las variables aleatorias X_1, X_2, X_3 se refieren a los números de cada resultado en los cuatro ensayos independientes. Por tanto, $X_1, X_2, X_3 \sim \text{MN}(4, 0.7, 0.2, 0.1)$.

Para mostrar cómo se calculan las probabilidades de variables aleatorias multinomiales, se calculará $P(X_1 = 2, X_2 = 1 \text{ y } X_3 = 1)$, en la que X_1, X_2, X_3 se definieron en el ejemplo 4.35.

Esto último conduce a una deducción de la función de masa de probabilidad multinomial. Se comenzará por observar que hay 12 arreglos de dos correspondientes (C), una inferior (D) y una rechazada (R) entre los cuatro resultados. Aquí se enlistan.

CCDR	CCRD	CDCR	CDRC	CRCD	CRDC
DCCR	DCRC	DRCC	RCCD	RCDC	RDCC

Cada uno de estos 12 arreglos es igualmente probable. Se calcula la probabilidad de CCDD. El evento CCDD es una secuencia de cuatro resultados: C en el primer ensayo, C en el segundo, D en el tercero y R en el cuarto. Dado que los ensayos son independientes, la probabilidad de la secuencia de resultados es igual al producto de cada una de sus probabilidades.

$$P(\text{CCDD}) = (0.7)(0.7)(0.2)(0.1) = (0.7)^2(0.2)(0.1)$$

Puesto que cada uno de los 12 arreglos tiene la misma probabilidad,

$$P(X_1 = 2, X_2 = 1, X_3 = 1) = (12)(0.7)^2(0.2)(0.1) = 0.1176$$

En este cálculo, el número de arreglos era pequeño para contarlos al enlistarlos a todos. En general, para calcular probabilidades como ésta, se necesita una fórmula. Esta última se encuentra en el siguiente recuadro. En la sección 2.2 se presenta una deducción.

Suponga que se llevan a cabo n experimentos independientes, cada uno tiene k resultados posibles. Sean $x_1, \dots, x_k$ el número de ensayos con resultados $1, 2, \dots, k$, respectivamente. El número de arreglos de los resultados entre los n ensayos es

$$\frac{n!}{x_1! x_2! \cdots x_k!}$$

Ahora, se puede especificar la función de masa de probabilidad multinomial.

Si $X_1, \dots, X_k \sim \text{MN}(n, p_1, \dots, p_k)$, entonces la función de masa de probabilidad de $X_1, \dots, X_k$ es

$$p(x_1, \dots, x_k) = P(X_1 = x_1, \dots, X_k = x_k) = \begin{cases} \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} & x_i = 0, 1, 2, \dots, n \\ & \text{y } \sum x_i = n \\ 0 & \text{de otro modo} \end{cases}$$

Observe que la distribución multinomial difiere de las otras distribuciones en que estudia diversas variables aleatorias simultáneamente. Este hecho se expresa al afirmar que $p(x_1, \dots, x_k)$ es la **función de masa de probabilidad conjunta** de $X_1, \dots, X_k$. La sección 2.6 analiza con más profundidad las funciones de masa de probabilidad conjunta.

Ejemplo

4.36

La alcaptonuria es una enfermedad genética que se caracteriza por carencia de una enzima necesaria para metabolizar al ácido homogentísico. Algunas personas son portadores de aquélla, lo cual significa que no tienen la enfermedad, sino que pueden transmitirla potencialmente a sus hijos. De acuerdo con las leyes de la herencia genética, un hijo, cuyos padres son portadores de alcaptonuria, tiene probabilidad de 0.25 de no tener la enfermedad, 0.5 de ser portador y 0.25 de padecer la enfermedad. En una muestra de diez hijos de portadores de la alcaptonuria, ¿cuál es la probabilidad de que tres no la tengan, cinco sean portadores y dos la padezcan?

Solución

Sean X_1, X_2, X_3 los números entre los diez hijos que no tienen la enfermedad, son portadores y la padecen, respectivamente. Entonces $X_1, X_2, X_3 \sim \text{MN}(10, 0.25, 0.50, 0.25)$, por lo que de la ecuación (4.22) se tiene que

$$\begin{aligned} P(X_1 = 3, X_2 = 5, X_3 = 2) &= \frac{10!}{3! 5! 2!} (0.25)^3 (0.50)^5 (0.25)^2 \\ &= (2\,520)(0.015625)(0.03125)(0.0625) \\ &= 0.07690 \end{aligned}$$

Algunas veces sólo se desea concentrar en uno de los posibles resultados de un experimento multinomial. En esta situación, se puede considerar el resultado de interés como “éxito”, y cualquier otro resultado como “fracaso”. De este modo, es posible ver que el número de ocurrencias de cualquier resultado particular tiene una distribución binomial.

Si $X_1, \dots, X_k \sim \text{MN}(n, p_1, \dots, p_k)$, entonces para cada i

$$X_i \sim \text{Bin}(n, p_i)$$

Ejemplo

4.37

Con referencia al ejemplo 4.36, determine la probabilidad de que sólo cuatro de los diez hijos no padezcan la enfermedad.

Solución

Sea X el número de hijos que no tiene la enfermedad en una muestra de diez. Entonces $X \sim \text{Bin}(10, 0.25)$; por tanto,

$$\begin{aligned} P(X = 4) &= \frac{10!}{4! 6!} (0.25)^4 (0.75)^6 \\ &= 0.1460 \end{aligned}$$

Ejercicios para la sección 4.4

1. Quince automóviles son llevados a una concesionaria para validar su garantía. Suponga que cinco presentan graves problemas de motor, mientras que diez tienen problemas sin importancia. Se eligen aleatoriamente seis automóviles para componerlos. ¿Cuál es la probabilidad de que dos tengan graves problemas?
 - i) 0
 - ii) $1/p$
 - iii) p
 - iv) 1
 - v) $(1 - p)/p^2$
2. Un cargamento contiene 40 elementos. Se seleccionará de forma aleatoria y se probará cinco elementos. Si dos o más están defectuosos, se regresará el cargamento.
 - a) Si, de hecho, el cargamento contiene cinco elementos defectuosos, ¿cuál es la probabilidad de que sean aceptados (no regresados)?
 - b) Si, de hecho, el cargamento contiene diez elementos defectuosos, ¿cuál es la probabilidad de que se regresen?
3. La probabilidad de que una computadora que corre cierto sistema operativo se descomponga en determinado día es de 0.1. Determine la probabilidad de que la máquina se descomponga por primera vez en el duodécimo día, después de la instalación del sistema operativo.
4. Un semáforo localizado en cierta intersección está en verde 50% de las veces, en ámbar 10% y en rojo 40%. Un automóvil pasa por esta intersección una vez al día. Sea X el número de días que ha transcurrido, incluyendo la primera vez que el automóvil se topa con una luz roja. Suponga que cada día representa un experimento independiente.
 - a) Determine $P(X = 3)$.
 - b) Determine $P(X \leq 3)$.
 - c) Determine μ_X .
 - d) Determine σ_X^2 .
5. Con referencia al ejercicio 4. Sea Y el número de días incluyendo el tercer día en que se topa con una luz roja.
 - a) Determine $P(X = 7)$.
 - b) Determine μ_X .
 - c) Determine σ_X^2 .
6. Con referencia al ejercicio 4, ¿cuál es la probabilidad de que en una secuencia de diez días, se haya encontrado con cuatro luces verdes, una ámbar y cinco rojas?
7. Si $X \sim \text{Geom}(p)$, ¿cuál es el valor más probable de X ?
 - i) 0
 - ii) $1/p$
 - iii) p
 - iv) 1
 - v) $(1 - p)/p^2$
8. Un proceso que llena paquetes se detiene cada vez que se detecta uno cuyo peso no cumple la especificación. Suponga que cada paquete tiene probabilidad de 0.01 de no cumplir con la especificación y que los pesos de los paquetes son independientes.
 - a) Determine el número promedio de paquetes que será llenado antes de que se detenga el proceso.
 - b) Determine la varianza del número de paquetes que será llenado antes de que se detenga el proceso.
 - c) Suponga que el proceso no se detendrá hasta que se detecten cuatro paquetes, cuyo peso no cumpla con la especificación. Determine la media y la varianza del número de paquetes que será llenado antes de que se detenga el proceso.
9. En un lote de diez microcircuitos, tres están defectuosos. Se elige aleatoriamente cuatro microcircuitos para ser probados. Sea X el número de circuitos probados que son defectuosos.
 - a) Determine $P(X = 2)$.
 - b) Determine μ_X .
 - c) Determine σ_X .
10. De los clientes que ordenan cierto tipo de computadora personal, 20% ordena una tarjeta gráfica actualizada, 30% memoria extra, 15% ordena tanto una tarjeta gráfica actualizada como memoria extendida, y 35% no ordena ninguna. Se eligen de forma aleatoria quince órdenes. Sea X_1, X_2, X_3, X_4 los respectivos números de órdenes en las cuatro categorías dadas.
 - a) Determine $P(X_1 = 3, X_2 = 4, X_3 = 2 \text{ y } X_4 = 6)$.
 - b) Determine $P(X_1 = 3)$.
11. Cierta marca de automóvil viene equipada con un motor en uno de cuatro tamaños (en litros): 2.8, 3.0, 3.3, o 3.8. El 10% de los clientes ordena el motor de 2.8 litros, 40% de 3.0, 30% de 3.3 y 20% de 3.8. Se selecciona una muestra aleatoria de 20 órdenes para una auditoría.

- a) ¿Cuál es la probabilidad de que el número de órdenes para los motores de 2.8, 3.0, 3.3 y 3.8 litros sean 3, 7, 6 y 4, respectivamente?
- b) ¿Cuál es la probabilidad de que haya más de diez órdenes de los motores de 3.0 litros?
12. Un termopar localizado dentro de cierto medio produce lecturas con margen de 0.1°C de la temperatura real 70% de las veces, lecturas mayores a 0.1°C por encima de la temperatura real 10% de las veces, y lecturas mayores a 0.1°C por debajo de la temperatura real 20% de las veces.
- a) En una serie de diez lecturas independientes, ¿cuál es la probabilidad de que cinco se encuentren dentro de 0.1°C de la temperatura real, dos a más de 0.1°C por encima de ella, y tres a más de 0.1°C debajo de dicho parámetro?
- b) ¿Cuál es la probabilidad de que más de ocho lecturas se encuentren dentro de 0.1°C de la temperatura real?

4.5 Distribución normal

La **distribución normal** (también conocida como **distribución de Gauss**) es la distribución más utilizada en la estadística. Constituye un buen modelo para muchas, aunque no para todas las poblaciones continuas. Parte de esto último se debe al teorema del límite central, que se analizará en la sección 4.10.

La distribución normal es continua en vez de discreta. La media de una variable aleatoria normal puede tener cualquier valor y la varianza cualquier valor positivo. La función de densidad de probabilidad de una variable aleatoria normal con media μ y varianza σ^2 está dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (4.23)$$

Al final de esta sección se comprueba el hecho de que μ y σ^2 son la media y la varianza, respectivamente. Si X es una variable aleatoria cuya función de densidad de probabilidad es normal con media μ y varianza σ^2 , se expresa como $X \sim N(\mu, \sigma^2)$.

Resumen

Si $X \sim N(\mu, \sigma^2)$, entonces la media y la varianza de X están dadas por

$$\begin{aligned} \mu_X &= \mu \\ \sigma_X^2 &= \sigma^2 \end{aligned}$$

La figura 4.4 presenta una gráfica de la función de densidad de probabilidad normal con media μ y desviación estándar σ . Algunas veces a la función de densidad de probabilidad normal se le llama **curva normal**. Observe que ésta es simétrica alrededor de μ , de tal forma que μ representa la mediana, así como la media. También, toda población normal se caracteriza por

- Aproximadamente 68% de la población se encuentra en el intervalo $\mu \pm \sigma$.
- Aproximadamente 95% de la población se encuentra en el intervalo $\mu \pm 2\sigma$.
- Aproximadamente 99.7% de la población se encuentra en el intervalo $\mu \pm 3\sigma$.

La proporción de una población normal que se encuentra a cierto número de desviaciones estándar de la media es la misma en cualquier población normal. Por esta razón, cuando se trabaja con poblaciones normales, se convierte las unidades en las cuales se midió originalmente las unidades de la población a **unidades estándar**. Estas últimas indican a cuántas desviaciones estándar se encuentra un dato de la media poblacional.

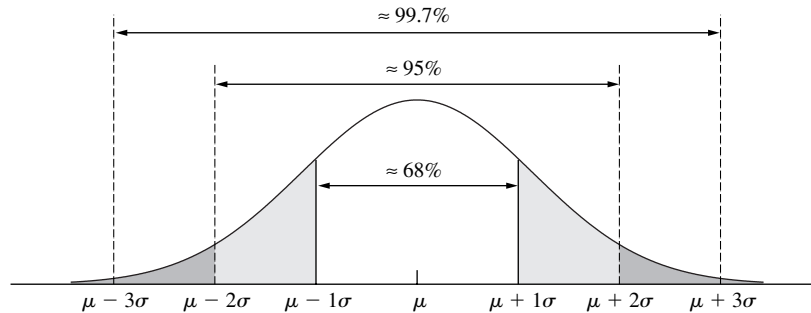


FIGURA 4.4 Función de densidad de probabilidad de una variable aleatoria normal con media μ y varianza σ^2 .

Ejemplo

4.38

Suponga que la estatura en una población de mujeres sigue una curva normal con media $\mu = 64$ pulgadas y desviación estándar $\sigma = 3$ pulgadas. La estatura de dos mujeres elegidas aleatoriamente es de 67 y 62 pulgadas, respectivamente. Convierta estas estaturas a unidades estándares.

Solución

Una estatura de 67 pulgadas es tres pulgadas mayor que la media de 64, y tres pulgadas es igual a una desviación estándar. Por tanto, 67 pulgadas es una desviación estándar mayor que la media y equivalente a una unidad estándar. Una estatura de 62 pulgadas es 0.67 desviaciones estándar menor que la media, por lo que 62 pulgadas es equivalente a -0.67 unidades estándar.

En general, se convierte a unidades estándar al restar la media y dividir entre la desviación estándar. Por consiguiente, si x es una unidad seleccionada de una población normal con media μ y varianza σ^2 , la unidad estándar equivalente a x es el número z , donde

$$z = \frac{x - \mu}{\sigma} \quad (4.24)$$

Algunas veces, al número z se le denomina “puntaje z ” de x , que representa un elemento extraído de una población normal con media 0 y desviación estándar 1. A aquélla se le llama **población normal estándar**.

Ejemplo

4.39

Las láminas de aluminio utilizadas para fabricar latas de bebida tienen un espesor (en milésimas de pulgada) que se distribuye normalmente con una media de 10 y desviación estándar de 1.3. Una lámina particular tiene un espesor de 10.8 milésimas de pulgadas. Determine el puntaje z .

Solución

La cantidad 10.8 es un dato de una población normal con media $\mu = 10$ y desviación estándar $\sigma = 1.3$. Por consecuencia,

$$z = \frac{10.8 - 10}{1.3} \\ = 0.62$$

Ejemplo

4.40

Con referencia al ejemplo 4.39, el espesor de cierta lámina tiene un puntaje z de -1.7 . Determine el espesor de la lámina en las unidades originales en milésimas de pulgada.

Solución

Usando la ecuación (4.24), sustituyendo -1.7 para z y despejando x . Se obtiene

$$-1.7 = \frac{x - 10}{1.3}$$

Al despejar x se obtiene $x = 7.8$. La lámina tiene un espesor de 7.8 milésimas de pulgada.

La proporción de una población normal que se encuentra dentro de un intervalo específico es igual al área que se encuentra debajo de la densidad de probabilidad normal en dicho intervalo. Esto último sugiere que dichas proporciones se calculan al integrar la densidad de probabilidad normal dada en la ecuación (4.23). Lo que es muy interesante es que las áreas debajo de esta curva no pueden determinarse mediante el método, enseñado en cálculo elemental, de encontrar la integral de la función y colocar los límites de integración. Lo anterior se debe a que la integral de esta función es una serie infinita y no puede escribirse con exactitud. En su lugar, las áreas debajo de esta curva deben aproximarse numéricamente.

Las áreas debajo de la curva normal estándar (media 0, varianza 1) se han tabulado extensivamente. Una tabla común, denominada **tabla normal estándar**, o **tabla z** , es la tabla A.2 (en el Apéndice A). Para determinar las áreas debajo de una curva normal con diferentes media y varianza, se convierten a unidades estándares y se utiliza la tabla z . La tabla A.2 proporciona las áreas en la cola izquierda de la curva para valores de z . Es posible calcular otras áreas al sustraer o usando el hecho de que el área total debajo de la curva es igual a 1. Ahora se presentan diversos ejemplos para mostrar el uso de la tabla z .

Ejemplo

4.41

Determine el área debajo de la curva normal a la izquierda de $z = 0.47$.

Solución

De la tabla z , el área es de 0.6808. Véase la figura 4.5.

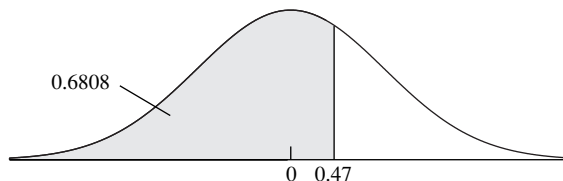


FIGURA 4.5 Solución al ejemplo 4.41.

Ejemplo**4.42**

Determine el área debajo de la curva normal y a la derecha de $z = 1.38$.

Solución

De la tabla z , el área a la izquierda de $z = 1.38$ es de 0.9162. Por tanto, el área a la derecha es $1 - 0.9162 = 0.0838$. Véase la figura 4.6.

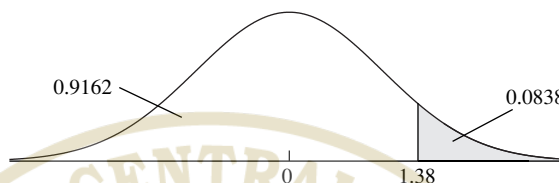


FIGURA 4.6 Solución al ejemplo 4.42.

Ejemplo**4.43**

Determine el área debajo de la curva normal entre $z = 0.71$ y $z = 1.28$.

Solución

De la tabla z , el área a la izquierda de $z = 1.28$ es de 0.8997. El área a la izquierda de $z = 0.71$ es de 0.7611. Por consiguiente, el área entre $z = 0.71$ y $z = 1.28$ es $0.8997 - 0.7611 = 0.1386$. Véase la figura 4.7.

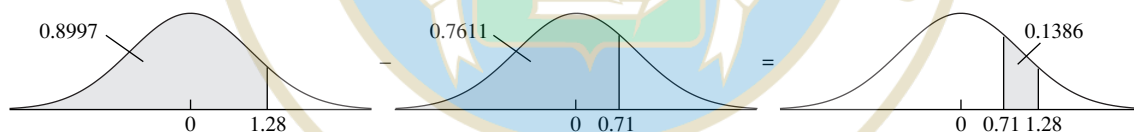


FIGURA 4.7 Solución al ejemplo 4.43.

Ejemplo**4.44**

¿A qué puntaje z corresponde el 75o. percentil de una curva normal? ¿El 25o. percentil? ¿La mediana?

Solución

Para responder esta pregunta se utiliza al revés la tabla z . Se necesita encontrar el puntaje z para el que 75% del área de la curva está a la izquierda. A partir del contenido de la tabla, el área más cercana a 75% es 0.7486, correspondiente al puntaje z de 0.67. Por tanto, el 75o. percentil es aproximadamente de 0.67. Por simetría de la curva, el 25o. percentil es $z = -0.67$ (esto también puede verse directamente en la tabla). Véase la figura 4.8. La mediana es $z = 0$.

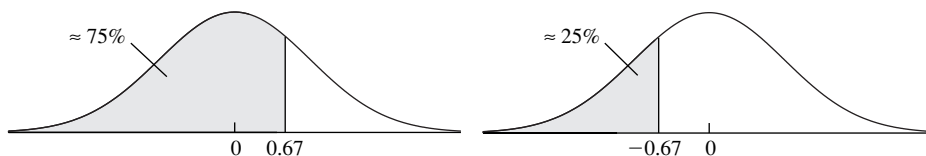


FIGURA 4.8 Solución al ejemplo 4.44.

Ejemplo**4.45**

Los tiempos de vida de las baterías en cierta aplicación se distribuyen normalmente con media de 50 horas y desviación estándar de cinco horas. Determine la probabilidad de que se elija aleatoriamente una batería que dure entre 42 y 52 horas.

Solución

Sea X el tiempo de vida de una batería elegida aleatoriamente. Entonces $X \sim N(50, 5^2)$. La figura 4.49 muestra la función de densidad de probabilidad de la población $N(50, 5^2)$. El área sombreada representa $P(42 < X < 52)$, la probabilidad de que una batería seleccionada de forma aleatoria tenga una duración entre 42 y 52 horas. Para calcular esta área, se hará uso de la tabla z . Primero se necesita convertir las cantidades 42 y 52 a unidades estándar. Se tiene

$$z = \frac{42 - 50}{5} = -1.60 \quad z = \frac{52 - 50}{5} = 0.40$$

De la tabla z , el área a la izquierda de $z = -1.60$ es 0.0548, y el área a la izquierda de $z = 0.40$ es 0.6554. La probabilidad de que una batería tenga tiempo de vida entre 42 y 52 horas es $0.6554 - 0.0548 = 0.6006$.

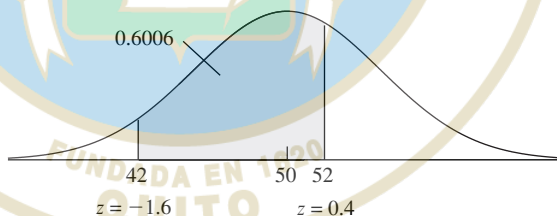


FIGURA 4.9 Solución al ejemplo 4.45.

Ejemplo**4.46**

Con referencia al ejemplo 4.45, determine el 40o. percentil de los tiempos de vida de las baterías.

Solución

De la tabla z , el área más cercana a 0.4000 es 0.4013, correspondiente al puntaje z de -0.25 . La población de los tiempos de vida tiene una media de 50 y una desviación estándar de 5. El 40o. percentil es el punto 0.25 desviaciones estándar menor a la media. Este valor se determina al convertir el puntaje z en uno nuevo, utilizando la ecuación (4.24):

$$-0.25 = \frac{x - 50}{5}$$

Al despejar x se tiene que $x = 48.75$. El 40o. percentil de los tiempos de vida de las baterías es de 48.75 horas. Véase la figura 4.10.

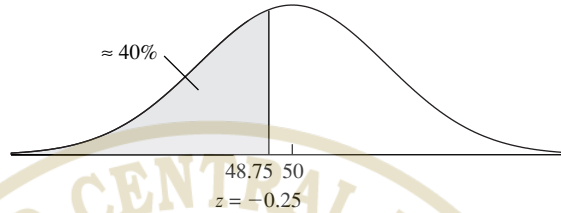


FIGURA 4.10 Solución al ejemplo 4.46.

Ejemplo 4.47

Un proceso fabrica cojinetes de bolas cuyos diámetros se distribuye normalmente con media de 2.505 cm y desviación estándar de 0.008 cm. Las especificaciones requieren que el diámetro esté dentro del intervalo 2.5 ± 0.01 cm. ¿Qué proporción de cojinetes de bolas cumple con la especificación?

Solución

Sea X el diámetro de un cojinete de bolas seleccionado aleatoriamente. Entonces $X \sim N(2.505, 0.008^2)$. La figura 4.11 presenta la función de densidad de probabilidad de la población $N(2.505, 0.008^2)$. El área sombreada representa $P(2.49 < X < 2.51)$, que es la proporción de cojinetes de bolas que cumplen con la especificación.

Se calcula los puntajes z de 2.49 y 2.51:

$$z = \frac{2.49 - 2.505}{0.008} = -1.88 \quad z = \frac{2.51 - 2.505}{0.008} = 0.63$$

El área a la izquierda de $z = -1.88$ es 0.0301. El área a la izquierda de $z = 0.63$ es 0.7357. El área entre $z = 0.63$ y $z = -1.88$ es $0.7357 - 0.0301 = 0.7056$. Aproximadamente 70.56% de los diámetros satisface la especificación.

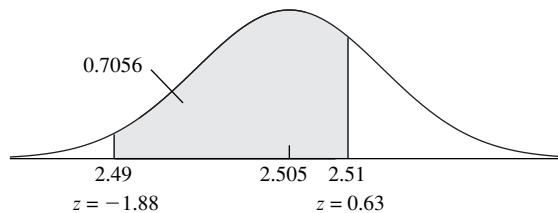


FIGURA 4.11 Solución al ejemplo 4.47.

Ejemplo 4.48

Con referencia al ejemplo 4.47, el proceso puede recalibrarse para que la media sea igual a 2.5 cm, el centro del intervalo de la especificación. La desviación estándar del proceso sigue siendo de 0.008 cm. ¿Qué proporción de los diámetros satisface la especificación?

Solución

El método de solución es el mismo que en el ejemplo 4.47. La media es de 2.500 en vez de 2.505. Los cálculos se realizan de la siguiente manera:

$$z = \frac{2.49 - 2.50}{0.008} = -1.25 \quad z = \frac{2.51 - 2.50}{0.008} = 1.25$$

El área a la izquierda de $z = -1.25$ es 0.1056. El área a la izquierda de $z = 1.25$ es 0.8944. El área entre $z = 1.25$ y $z = -1.25$ es $0.8944 - 0.1056 = 0.7888$. Véase la figura 4.12. El recalibrado aumenta a 78.88% la proporción de diámetros que satisface la especificación.

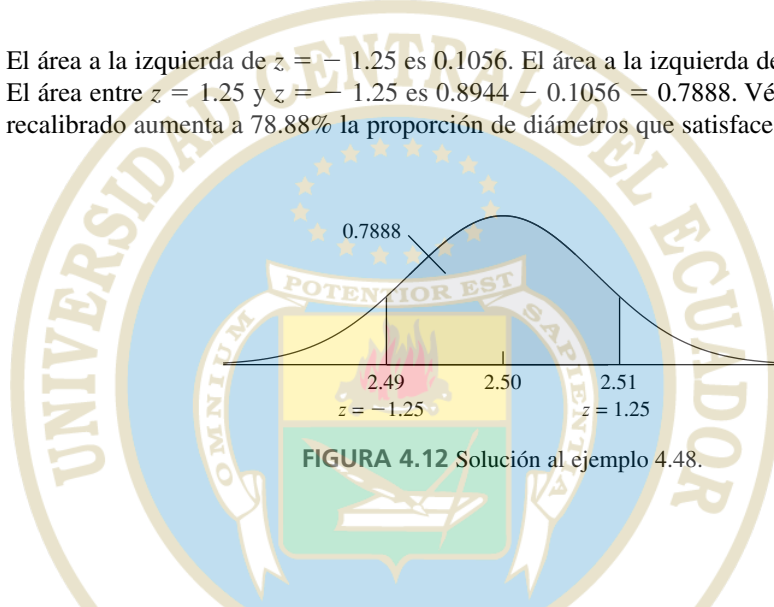


FIGURA 4.12 Solución al ejemplo 4.48.

Ejemplo 4.49

Con referencia a los ejemplos 4.47 y 4.48, suponga que se ha recalibrado el proceso de tal forma que la media del diámetro mide ahora 2.5 cm. ¿A qué valor debe reducirse la desviación estándar para que 95% de los diámetros satisfaga la especificación?

Solución

El intervalo de especificación es 2.49 – 2.51 cm. Se debe encontrar un valor de σ para que este intervalo abarque 95% de la población de diámetros de cojinetes de bolas. Véase la figura 4.13. El puntaje z que tiene 2.5% del área a la izquierda es $z = -1.96$. El puntaje z que tiene 2.5% del área a su derecha es $z = 1.96$ (esto último se obtiene de la simetría de la curva). De ahí que el límite menor de la especificación, 2.49, tenga un puntaje z de -1.96 , mientras que el límite superior de 2.51 tiene un puntaje z de 1.96. Cualesquiera de estos hechos se puede utilizar para encontrar a σ . De la ecuación (4.24),

$$1.96 = \frac{2.51 - 2.50}{\sigma}$$

Al despejar σ se tiene que $\sigma = 0.0051$ cm.

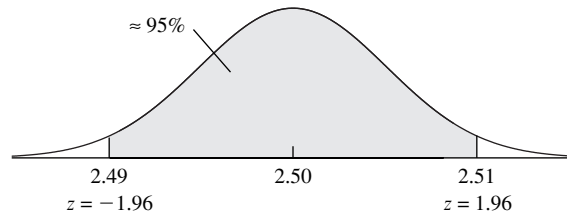


FIGURA 4.13 Solución al ejemplo 4.49. Si $\sigma = 0.0051$, entonces 95% de la población estará entre 2.49 y 2.51.

Estimación de los parámetros de una distribución normal

Los parámetros μ y σ de una distribución normal representan su media y varianza, respectivamente. Por tanto, si $X_1, \dots, X_n$ son una muestra aleatoria de una distribución $N(\mu, \sigma^2)$, μ se estima con la media muestral $\bar{X}$ y σ^2 se estima con la varianza muestral s^2 . Al igual que con cualquier otra media, la incertidumbre en $\bar{X}$ es $\sigma/\sqrt{n}$, que será reemplazada con $s/\sqrt{n}$ si no se conoce a σ . Además, $\mu_{\bar{X}} = \mu$, por lo que $\bar{X}$ es un estimador no sesgado de μ .

Combinaciones lineales de variables aleatorias independientes normales

Una de las características más notables de la distribución normal consiste en que las combinaciones lineales de variables aleatorias independientes normales son en sí mismas variables aleatorias. Para ser más específico, suponga que $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $\dots$, $X_n \sim N(\mu_n, \sigma_n^2)$ son variables aleatorias independientes normales. Observe que las medias y varianzas de estas variables aleatorias pueden diferir entre sí. Sean $c_1, c_2, \dots, c_n$ constantes. Entonces, la combinación lineal $c_1X_1 + c_2X_2 + \dots + c_nX_n$ es una variable aleatoria distribuida normalmente. La media y la varianza de la combinación lineal son $c_1\mu_1 + c_2\mu_2 + \dots + c_n\mu_n$ y $c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \dots + c_n^2\sigma_n^2$, respectivamente (véanse las ecuaciones 2.49 y 2.53 en la sección 2.5).

Resumen

Sean $X_1, X_2, \dots, X_n$ independientes y distribuidas normalmente con medias $\mu_1, \mu_2, \dots, \mu_n$ y varianzas $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$. Sean $c_1, c_2, \dots, c_n$ constantes, y sea $c_1X_1 + c_2X_2 + \dots + c_nX_n$ una combinación lineal, entonces

$$c_1X_1 + c_2X_2 + \dots + c_nX_n \sim N(c_1\mu_1 + c_2\mu_2 + \dots + c_n\mu_n, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \dots + c_n^2\sigma_n^2) \quad (4.25)$$

Ejemplo

4.50

En el artículo “Advances in Oxygen Equivalent Equations for Predicting the Properties of Titanium Welds” (D. Harwig, W. Ittiwattana y H. Castner, en *The Welding Journal*, 2001: 126s-136s), los autores proponen una ecuación de equivalencia al oxígeno para predecir la re-

sistencia, ductilidad y dureza de soldaduras hechas casi de titanio puro. La ecuación es $E = 2C + 3.5N + O$, donde E es la equivalencia del oxígeno, y C , N y O las proporciones de peso, en partes por millón, de carbono, nitrógeno y oxígeno, respectivamente (se omitió un término constante que consiste en el contenido de hierro). Suponga que para un nivel particular de titanio puro comercial, las cantidades C , N y O son aproximadamente independientes y se distribuyen normalmente con medias $\mu_C = 150$, $\mu_N = 200$, $\mu_O = 1\,500$ y desviaciones estándar $\sigma_C = 30$, $\sigma_N = 60$, $\sigma_O = 100$. Determine la distribución de E . Determine $P(E > 3\,000)$.

Solución

Puesto que E es una combinación lineal de variables aleatorias normales independientes, su distribución es normal. Ahora se debe determinar la media y la varianza de E . Utilizando la ecuación (4.25) se calcula

$$\begin{aligned}\mu_E &= 2\mu_C + 3.5\mu_N + 1\mu_O \\ &= 2(150) + 3.5(200) + 1(1\,500) \\ &= 2\,500 \\ \sigma_E^2 &= 2^2\sigma_C^2 + 3.5^2\sigma_N^2 + 1^2\sigma_O^2 \\ &= 2^2(30^2) + 3.5^2(60^2) + 1^2(100^2) \\ &= 57\,700\end{aligned}$$

Se concluye que $E \sim N(2\,500, 57\,700)$.

Para calcular $P(E > 3\,000)$, se calcula el puntaje z : $z = (3\,000 - 2\,500)/\sqrt{57\,700} = 2.08$. El área a la derecha de $z = 2.08$ debajo de la curva normal es 0.0188. Por tanto, $P(E > 3\,000) = 0.0188$.

Si $X_1, X_2, \dots, X_n$ es una muestra aleatoria de cualquier población con media μ y varianza σ^2 , entonces la media muestral $\bar{X}$ tiene una media $\mu_{\bar{X}} = \mu$ y varianza $\sigma_{\bar{X}}^2 = \sigma^2/n$. Si la población es normal, entonces $\bar{X}$ también es normal, ya que es una combinación lineal de $X_1, \dots, X_n$ con coeficientes $c_1 = \dots = c_n = 1/n$.

Resumen

Sean $X_1, \dots, X_n$ independientes y distribuidas normalmente con media μ y varianza σ^2 , entonces

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad (4.26)$$

Otras combinaciones lineales importantes son la suma y resta de dos variables aleatorias. Si X y Y son variables aleatorias independientes, la suma $X + Y$ y la resta $X - Y$ son combinaciones lineales. Las distribuciones $X + Y$ y $X - Y$ se puede determinar utilizando la ecuación (4.25) con $c_1 = 1$, $c_2 = 1$ para $X + Y$ y $c_1 = 1$, $c_2 = -1$ para $X - Y$.

Resumen

Sean X y Y independientes, con $X \sim N(\mu_X, \sigma_X^2)$ y $Y \sim N(\mu_Y, \sigma_Y^2)$, entonces

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2) \quad (4.27)$$

$$X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2) \quad (4.28)$$

¿Cómo puede saberse si los datos que se tiene provienen de una población normal?

En la práctica, a menudo se tiene una muestra de alguna población y se le debe utilizar para decidir si la distribución de la población se aproxima a la normal. Si la muestra es razonablemente grande, su histograma es buena indicación. Las muestras grandes de poblaciones normales tienen histogramas que se parecen a la función de densidad normal: con un pico en el centro, y decrecientes más o menos simétricamente en cada lado. Las gráficas de probabilidad, que se analizarán en la sección 4.9, son otra buena forma de determinar si una muestra grande proviene de una población que es aproximadamente normal. Para muestras pequeñas, es difícil determinar si la distribución normal es la apropiada. Un hecho importante es éste: *las muestras de poblaciones normales raramente tienen datos atípicos*. Por tanto, no se utiliza la distribución normal en conjuntos de datos que tengan valores atípicos. Esto es cierto cuando el tamaño muestral es pequeño. Infortunadamente, para conjuntos pequeños de datos que no tienen datos atípicos, es difícil determinar si la población es aproximadamente normal. En general, se requiere de cierto conocimiento del proceso que generó los datos.

Deducción de la media y varianza de una variable aleatoria normal

Sea $X \sim N(\mu, \sigma^2)$. Se demuestra que $\mu_X = \mu$ y $\sigma_X^2 = \sigma^2$. Empleando la definición de la media poblacional de una variable aleatoria continua (ecuación 2.35 de la sección 2.4),

$$\mu_X = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} x e^{-(x-\mu)^2/2\sigma^2} dx$$

Se hace la sustitución de $z = (x - \mu)/\sigma$. Entonces $x = \sigma z + \mu$ y $dx = \sigma dz$. Se obtiene

$$\begin{aligned} \mu_X &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} (\sigma z + \mu) \sigma e^{-z^2/2} dz \\ &= \sigma \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z e^{-z^2/2} dz + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \end{aligned}$$

El cálculo directo demuestra que

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z e^{-z^2/2} dz = 0$$

Asimismo,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1$$

debido a que es la integral de la función de densidad de probabilidad $N(0, 1)$ sobre todos sus posibles valores.

Por tanto,

$$\mu_X = \sigma(0) + \mu(1) = \mu$$

Para demostrar que $\sigma_X^2 = \sigma^2$ se utiliza la ecuación (2.36) (de la sección 2.4):

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx$$

Se hace la sustitución de $z = (x - \mu)/\sigma$. Recuerde que $\mu_X = \mu$. Entonces

$$\begin{aligned} \sigma_X^2 &= \int_{-\infty}^{\infty} \sigma^2 z^2 \frac{1}{\sigma\sqrt{2\pi}} \sigma e^{-z^2/2} dz \\ &= \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^2 e^{-z^2/2} dz \end{aligned}$$

Al integrar por partes dos veces se demuestra que

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^2 e^{-z^2/2} dz = 1$$

Por consiguiente, $\sigma_X^2 = \sigma^2$.

Ejercicios para la sección 4.5

- Determine el área bajo la curva normal
 - A la derecha de $z = -0.85$.
 - Entre $z = 0.40$ y $z = 1.30$.
 - Entre $z = -0.30$ y $z = 0.90$.
 - Desde $z = -1.50$ hasta $z = -0.45$.
 - ¿Cuál es la proporción de puntuaciones mayores a 700?
 - ¿Cuál es el 25o. percentil de las puntuaciones?
 - Si la puntuación de alguien es de 600, ¿en qué percentil se encuentra?
 - ¿Qué proporción de las puntuaciones se encuentra entre 420 y 520?
- Determine el área bajo la curva normal
 - A la izquierda de $z = 0.56$.
 - Entre $z = -2.93$ y $z = -2.06$.
 - Entre $z = -1.08$ y $z = 0.70$.
 - Desde $z = 0.96$ hasta $z = 1.62$.
- Las puntuaciones de una prueba estandarizada se distribuyen normalmente con media de 480 y desviación estándar de 90.
 - ¿Cuánto mide una mujer cuya estatura se encuentra en el 90o. percentil?
- Suponga que la estatura de mujeres en una población sigue la curva normal con media de 64.3 pulgadas y desviación estándar de 2.6 pulgadas.
 - ¿Qué proporción de mujeres tiene estatura entre 60 y 66 pulgadas?
 - La estatura de una mujer es 0.5 de desviación estándar mayor a la media. ¿Qué proporción de mujeres mide más que ella?
 - ¿Cuánto mide una mujer cuya estatura se encuentra en el 90o. percentil?

- d) Se elige aleatoriamente una mujer de esta población. ¿Cuál es la probabilidad de que ella mida más de 67 pulgadas?
- e) Se elige aleatoriamente a cinco mujeres de esta población. ¿Cuál es la probabilidad de que sólo una de ellas mida más de 67 pulgadas?
5. La resistencia de una aleación de aluminio se distribuye normalmente con media de 10 gigapascuales (GPa) y desviación estándar de 1.4 GPa.
- a) ¿Cuál es la probabilidad de que una muestra de esta aleación tenga resistencia mayor a 12 GPa?
- b) Determine el primer cuartil de la resistencia de esta aleación.
- c) Determine el 95o. percentil de la resistencia de esta aleación.
6. En una universidad, las puntuaciones del SAT en matemáticas de una clase de primer año fue de, en promedio, 650 y tuvo desviación estándar de 100. El máximo puntaje posible es de 800. ¿Es posible que el histograma de las puntuaciones de estos alumnos siga una curva normal? Explique.
7. La penicilina es producida por el hongo *Penicillium*, que crece en un caldo, cuyo contenido de azúcar debe controlarse con cuidado. La concentración óptima de azúcar es de 4.9 mg/mL. Si la concentración excede los 6 mg/mL, el hongo muere y el proceso debe suspenderse todo el día.
- a) Si la concentración de azúcar en tandas de caldo se distribuye normalmente con media 4.9 mg/mL y desviación estándar 0.6 mg/mL, ¿en qué proporción de días se suspenderá el proceso?
- b) El distribuidor ofrece vender caldo con una concentración de azúcar que se distribuye normalmente con media de 5.2 mg/mL y desviación estándar de 0.4 mg/mL. ¿Este caldo surtirá efectos con menos días de producción perdida? Explique.
8. Un método de cromatografía utilizado para purificar a una proteína también destruye parte de ésta, en un proceso denominado desnaturación. Un método particular recupera una media de 55% (0.55) de la proteína y tiene desviación estándar de 0.15. La cantidad recuperada se distribuye normalmente.
- a) En cierto proceso industrial, no es posible obtener una recuperación menor a 0.30 más de 5% de las veces. ¿Este proceso cumple con este requisito? Explique.
- b) En otro proceso, la recuperación debe ser mayor a 0.50 al menos 95% de las veces. Si la media de la recuperación se distribuye normalmente con una media de 0.60, ¿cuál es el valor más grande que puede tener la desviación estándar para cumplir con este requisito?
9. Se hace una perforación cilíndrica en un molde y se coloca un pistón cilíndrico en la perforación. La holgura es igual a la mitad de la diferencia entre los diámetros de la perforación y el pistón. El diámetro de la perforación se distribuye normalmente con media de 15 cm y desviación estándar de 0.025 cm, y el diámetro del pistón se distribuye con media 14.88 cm y desviación estándar 0.015 cm.
- a) Determine la media de la holgura.
- b) Determine la desviación estándar de la holgura.
- c) ¿Cuál es la probabilidad de que la holgura mida menos de 0.05 cm?
- d) Determine el 25o. percentil de la holgura.
- e) Las especificaciones requieren que la holgura mida entre 0.05 y 0.09 cm. ¿Cuál es la probabilidad de que la holgura satisfaga la especificación?
- f) Se puede ajustar la media del diámetro de la perforación. ¿A qué valor debe ajustarse para maximizar la probabilidad de que la holgura esté entre 0.05 y 0.09 cm?
10. Los ejes fabricados para el uso de dispositivos de almacenamiento óptico tienen diámetros que se distribuyen normalmente con media $\mu = 0.652$ y desviación estándar $\sigma = 0.003$ cm. La especificación para el diámetro del eje mide entre 0.650 ± 0.005 cm.
- a) ¿Qué proporción de los ejes fabricados por este proceso cumple con la especificación?
- b) La media del proceso puede ajustarse utilizando calibración. Si se establece que la media mide 0.650 cm, ¿qué proporción de los ejes cumplirá con la especificación?
- c) Si se establece que la media mide 0.650 cm, ¿cuál debe ser la desviación estándar para que 99% de los ejes cumpla con la especificación?
11. El volumen de latas llenadas por cierta máquina se distribuye con media de 12.05 onzas y desviación estándar de 0.03 onzas.
- a) ¿Qué proporción de latas contiene menos de 12 onzas?
- b) La media del proceso se puede ajustar utilizando calibración. ¿En qué valor debe fijarse la media para que 99% de las latas contenga 12 onzas o más?
- c) Si la media del proceso sigue siendo de 12.05 onzas, ¿en qué valor debe fijarse la media para que 99% de las latas contenga 12 onzas o más?

12. Un proceso de recubrimiento de películas genera filmes cuyo espesor se distribuye con media de 110 micrones y desviación estándar de 10 micrones. En cierta aplicación, el espesor mínimo aceptable es de 90 micrones.
- ¿Qué proporción de películas estarán demasiado delgadas?
 - ¿A qué valor debe establecerse la media para que sólo 1% de las películas esté muy delgado?
 - Si la media sigue siendo 110, ¿cuál debe ser la desviación estándar para que sólo 1% de las películas sea muy delgado?
13. Un proceso hilador de fibras produce una fibra cuya resistencia se distribuye con media de 75 N/m^2 . La resistencia mínima aceptable es de 65 N/m^2 .
- 10% de las fibras producidas mediante el método actual no cumple con la especificación mínima. ¿Cuál es la desviación estándar de la resistencia de las fibras en el proceso actual?
 - Si la media sigue siendo de 75 N/m^2 , ¿cuál debe ser la desviación estándar para que sólo 1% de las fibras no satisfaga la especificación?
 - Si la desviación estándar es de 5 N/m^2 , ¿en qué valor debe fijarse la media para que sólo 1% de las fibras no satisfaga la especificación?
14. El programa de garantía de calidad de cierto proceso de formulación de un adhesivo consiste en medir qué tanto el adhesivo pega un pedazo de plástico a una superficie de vidrio. Cuando el proceso funciona correctamente, la fuerza del adhesivo X se distribuye con media de 200 N y desviación estándar de 10 N. Cada hora, usted hace una medición de la fuerza del adhesivo. Usted debe informar a su supervisor si su medición indica que el proceso se ha desviado de su distribución objetivo.
- Calcule $P(X \leq 160)$ bajo el supuesto de que el proceso está funcionando correctamente.
 - Con base en su respuesta al inciso a), si el proceso funciona bien, ¿una fuerza de 160 N sería inusualmente pequeña? Explique.
 - Si usted observa una fuerza adhesiva de 160 N, ¿esto último sería una evidencia de que el proceso ya no funciona correctamente? Explique.
 - Encuentre $P(X \geq 203)$, bajo la suposición de que el proceso está funcionando bien.
 - Con base en su respuesta del inciso d), si el proceso funciona correctamente, ¿sería una fuerza de 203 N inusualmente grande? Explique.
- f) Si usted observa una fuerza adhesiva de 203 N, ¿lo anterior sería una evidencia de que el proceso ya no funciona correctamente? Explique.
- g) Encuentre $P(X \leq 195)$, bajo la suposición de que el proceso está funcionando bien.
- h) Con base en su respuesta del inciso g), si el proceso está funcionando correctamente, ¿sería una fuerza de 195 N inusualmente pequeña? Explique.
- i) Si usted observa una fuerza adhesiva de 195 N, ¿esto sería una evidencia de que el proceso ya no funciona correctamente? Explique.
15. Una instalación de luz tiene dos focos. El A es de un tipo cuya duración se distribuye con media de 800 horas y desviación estándar de 100 horas. El B tiene una duración que se distribuye con media de 900 horas y desviación estándar de 150 horas. Suponga que las duraciones de los focos son independientes.
- ¿Cuál es la probabilidad de que el foco B dure más que el A?
 - ¿Cuál es la probabilidad de que el foco B dure 200 horas más que el A?
 - Otra instalación de luz tiene sólo un foco. Se pone uno del tipo A y cuando se funde se instala otro de tipo B. ¿Cuál es la probabilidad de que la duración total de ambos sea mayor a 2 000 horas?
16. La molaridad de un soluto en solución se define como el número de moles del soluto por litro de solución ($1 \text{ mol} = 6.02 \times 10^{23}$ moléculas). Si X es la molaridad de una solución de cloruro de sodio (NaCl) y Y es la molaridad de una solución de carbonato de sodio (Na_2CO_3), la molaridad del ion de sodio (Na^+) en una solución hecha de partes iguales NaCl y Na_2CO_3 está dada por $M = 0.5X + Y$. Suponga que X y Y son independientes y se distribuyen normalmente y que X tiene media de 0.450 y desviación estándar de 0.050, y Y tiene media de 0.250 y desviación estándar de 0.025.
- ¿Cuál es la distribución de M ?
 - Determine $P(M > 0.5)$.
17. Una compañía recibe importante cargamento de pernos. Éstos se utilizarán en una aplicación que necesita de una torsión de 100 J. Antes de que se acepte el cargamento, un ingeniero especialista en control de calidad sacará una muestra de 12 pernos y medirá la torsión necesaria para romper a cada uno de ellos. El cargamento será aceptado si el ingeniero concluye que menos de 1% de los pernos tiene torsión de ruptura menor a 100 J.

- a) Si los 12 valores son 107, 109, 111, 113, 113, 114, 114, 115, 117, 119, 122, 124, calcule la media y la desviación estándar muestral.
- b) Suponga que se saca una muestra de 12 valores de una población normal, y suponga que la media y la desviación estándar muestrales calculadas en el inciso a) son realmente la media y la desviación estándar de la población. Calcule la proporción de pernos cuya torsión de ruptura es menor a 100 J. ¿Será aceptado el cargamento?
- c) ¿Qué pasará si los 12 valores hubieran sido 108, 110, 112, 114, 114, 115, 115, 116, 118, 120, 123, 140? Utilice el método descrito en los incisos a) y b) para determinar si el cargamento hubiera sido aceptado.
- d) Compare los conjuntos de 12 valores en los incisos a) y c). ¿En qué muestra los pernos son más resistentes?
- e) ¿El método es válido para ambas muestras? ¿Por qué sí o por qué no?

4.6 Distribución lognormal

Para datos que tienen valores atípicos, la distribución normal no es apropiada. La distribución **lognormal**, que tiene relación con la distribución normal, es, a menudo, buena opción para estos conjuntos de datos. La distribución lognormal se deriva de la distribución normal de la siguiente manera: Si X es una variable aleatoria normal con media μ y varianza σ^2 , entonces la variable aleatoria $Y = e^X$ tiene **distribución lognormal** con parámetros μ y σ^2 . Observe que si Y tiene una distribución normal con parámetros μ y σ^2 , entonces $X = \ln Y$ tiene una distribución normal con media μ y varianza σ^2 .

Resumen

- Si $X \sim N(\mu, \sigma^2)$, entonces la variable aleatoria $Y = e^X$ tiene distribución lognormal con parámetros μ y σ^2 .
- Si Y tiene distribución lognormal con parámetros μ y σ^2 , entonces la variable aleatoria $X = \ln Y$ tiene la distribución $N(\mu, \sigma^2)$.

La función de densidad de probabilidad de una variable aleatoria lognormal con parámetros μ y σ^2 es

$$f(x) = \begin{cases} \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[-\frac{1}{2\sigma^2} (\ln x - \mu)^2 \right] & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad (4.29)$$

La figura 4.14 presenta una gráfica de la función de densidad lognormal con parámetros $\mu = 0$ y $\sigma = 1$. Observe que la función de densidad está sesgada. Ésta es la razón por la que se utiliza la distribución lognormal para modelar procesos que tienden a producir ocasionalmente valores grandes o atípicos.

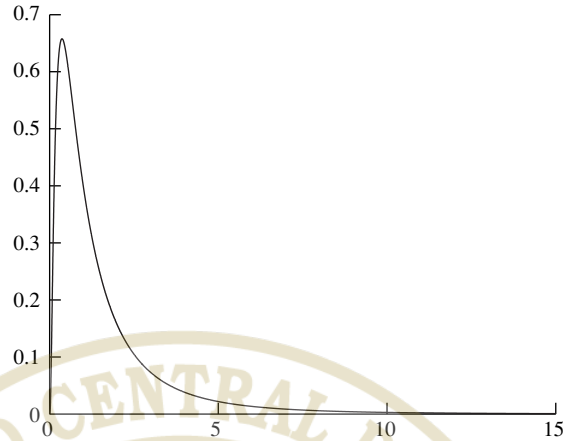


FIGURA 4.14 Función de densidad de probabilidad de la distribución lognormal con parámetros $\mu = 0$ y $\sigma = 1$.

Puede demostrarse mediante métodos avanzados que si Y es una variable aleatoria lognormal con parámetros μ y σ^2 , entonces la media $E(Y)$ y la varianza $V(Y)$ están dadas por

$$E(Y) = e^{\mu + \sigma^2/2} \quad V(Y) = e^{2\mu + 2\sigma^2} - e^{2\mu + 2\sigma^2} \quad (4.30)$$

Observe que si Y tiene una distribución lognormal, los parámetros μ y σ^2 no se refieren a la media y varianza de Y , sino que lo hacen a la media y varianza de la variable aleatoria normal $\ln Y$. En la ecuación (4.30) se utiliza la notación $E(Y)$ en vez de μ_Y y $V(Y)$ lugar de σ_Y^2 , para evitar confusiones entre μ y σ .

Ejemplo 4.51

El tiempo de vida de cierto componente sigue una distribución lognormal con parámetros $\mu = 1$ día y $\sigma = 0.5$ días. Determine la media del tiempo de vida de estos componentes. Encuentre la desviación estándar de los tiempos de vida.

Solución

Sea Y el tiempo de vida de un componente escogido aleatoriamente. La media de Y determinada mediante la ecuación (4.30) es $e^{1 + 0.5^2/2} = 3.08$ días. La varianza es $e^{2(1) + 2(0.5)^2} - e^{2(1) + (0.5)^2} = 2.6948$. Por tanto, la desviación estándar es $\sqrt{2.6948} = 1.64$ días.

Para calcular las probabilidades de variables aleatorias lognormales, se saca el logaritmo y se utiliza la tabla z (tabla A.2). Los ejemplos 4.52 y 4.53 ilustran el método.

Ejemplo

4.52

Con referencia al ejemplo 4.51, encuentre la probabilidad de que un componente dure más de cuatro días.

Solución

Sea Y el tiempo de vida de un componente seleccionado de forma aleatoria. Se necesita calcular $P(Y > 4)$. No se puede utilizar la tabla z para Y , debido a que Y no proviene de una población normal. Sin embargo, $\ln Y$ proviene de una población normal; específicamente, $\ln Y \sim N(1, 0.5^2)$. Se expresa $P(Y > 4)$ como una probabilidad que implica a $\ln Y$:

$$P(Y > 4) = P(\ln Y > \ln 4) = P(\ln Y > 1.386)$$

El puntaje z de 1.386 es

$$\begin{aligned} z &= \frac{1.386 - 1.000}{0.5} \\ &= 0.77 \end{aligned}$$

De la tabla z se tiene que $P(\ln Y > 1.386) = 0.2206$. (Véase la figura 4.15.) Se concluye que aproximadamente 22% de los componentes durará más de cuatro días.

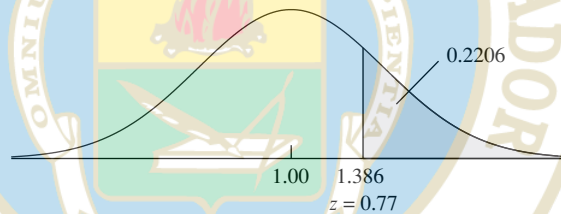


FIGURA 4.15 Solución al ejemplo 4.52.

Ejemplo

4.53

Con referencia al ejemplo 4.51 determine la mediana de los tiempos de vida. Encuentre el 80o. percentil de los tiempos de vida.

Solución

Sea Y el tiempo de vida de un componente escogido aleatoriamente. Sea M la mediana de los tiempos de vida. Entonces $P(Y \leq m) = 0.5$. En logaritmos, se tiene que $P(\ln Y \leq \ln m) = 0.5$. Esto significa que $\ln m$ es la mediana de $\ln Y$. Ahora, $\ln Y \sim N(1, 0.5^2)$. Por consecuencia, $\ln m = 1$, por lo que $m = e^1 = 2.718$.

Para encontrar al 80o. percentil, p_{80} , se hace $P(Y \leq p_{80}) = 0.80$. Por lo que $P(\ln Y \leq \ln p_{80}) = 0.80$. Esto significa que $\ln p_{80}$ es el 80o. percentil de $\ln Y$. Ahora $\ln Y \sim N(1, 0.5^2)$. De la tabla z , el puntaje z del 80o. percentil es 0.84. Por consiguiente, $\ln p_{80} = 1 + (0.84)(0.5) = 1.42$, de tal forma que $p_{80} = e^{1.42} = 4.14$.

Estimación de los parámetros de una distribución lognormal

Si Y es una variable aleatoria cuya distribución es lognormal con parámetros μ y σ^2 , entonces μ y σ^2 son la media y la varianza, respectivamente, de $\ln Y$. Por tanto, si $Y_1, \dots, Y_n$ es una muestra aleatoria de una población lognormal, primero se transforma a la escala logarítmica, definiendo $X_1 = \ln Y_1, \dots, X_n = \ln Y_n$. Ahora $X_1, \dots, X_n$ es una muestra aleatoria de $N(\mu, \sigma^2)$. Se estima μ con $\bar{X}$ y σ^2 con la varianza muestral s_X^2 . Al igual que con la media muestral, la incertidumbre en $\bar{X}$ es $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, y si no se conoce σ , se le estima con la desviación estándar muestral s_X .

Ejemplo 4.54

Los diámetros (en mm) de las semillas de cierta planta siguen una distribución normal. Una muestra aleatoria de cinco semillas tiene los diámetros 1.52, 2.22, 2.64, 2.00 y 1.69. Estime los parámetros μ y σ .

Solución

Para estimar μ y σ , se sacan los logaritmos de los cinco valores de la muestra, para obtener 0.419, 0.798, 0.971, 0.693 y 0.525. La media muestral es 0.681 y la desviación estándar muestral es 0.218. Por consecuencia, se estima que $\hat{\mu} = 0.681$, $\hat{\sigma} = 0.218$.

¿Cómo puede saberse si los datos provienen de una población lognormal?

Como se afirmó anteriormente, es muy raro que las muestras provenientes de poblaciones normales contengan datos atípicos. En contraste, las muestras provenientes de poblaciones lognormales a menudo tienen datos atípicos en la cola derecha. Es decir, las muestras contienen pocos valores que son más grandes que el resto de los datos. Obviamente, esto último se refleja en la larga cola derecha de la función de densidad lognormal (figura 4.14). Para las muestras con datos atípicos a la derecha, se *transforman* los datos, sacando el logaritmo natural (o cualquier logaritmo) de cada valor. Luego se intenta determinar si estos logaritmos vienen de una población normal, graficándolos en un histograma o en una gráfica de probabilidad. En la sección 4.9 se analizarán las gráficas de probabilidad.

Observe que la densidad lognormal tiene sólo una cola larga, a la derecha. Por esta razón, las muestras de poblaciones lognormales tienen datos atípicos a la derecha, pero no a la izquierda. Por consiguiente, no debe utilizarse la distribución lognormal en muestras con inusualmente muy pocos datos. Además, las poblaciones lognormales sólo tienen valores positivos, por lo que no puede emplearse la distribución lognormal en muestras que contengan ceros o valores negativos. Por último, es importante observar que la transformación log no siempre genera una muestra que se aproxima a la normal. Para verificar lo anterior, se tiene que graficar un histograma o un gráfico de probabilidad (véase la sección 4.9).

La figura 4.16 presenta dos histogramas. El primero muestra la producción mensual de 255 pozos de gas, en unidades de miles de pies cúbicos. El histograma claramente tiene una larga cola derecha, por lo que se concluye que los datos no provienen de una población normal. El segundo muestra los logaritmos naturales de las producciones mensuales. Este histograma se aproxima más a la curva normal, aunque se percibe cierta diferencia con la normalidad.

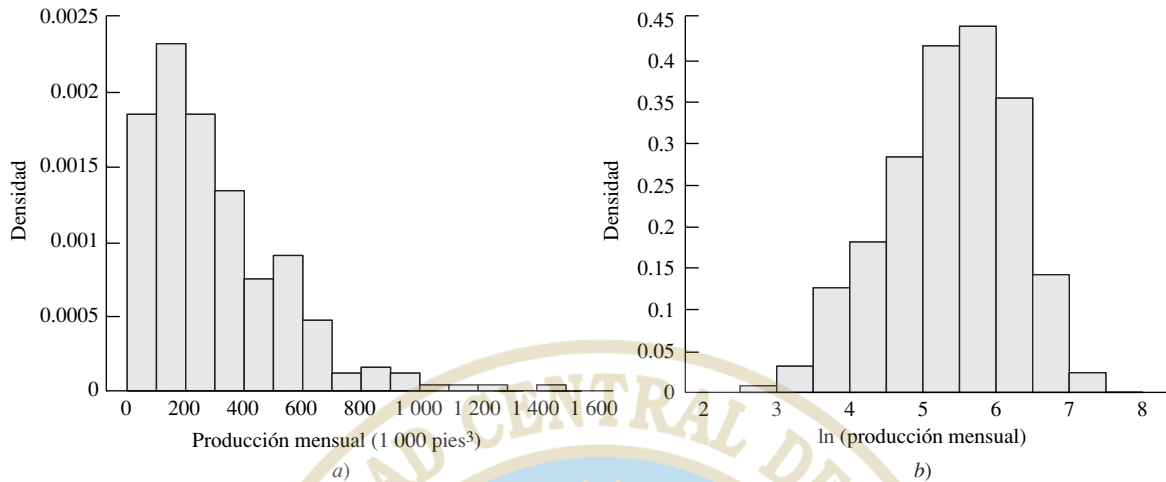


FIGURA 4.16 a) Histograma que muestra la producción mensual de 255 pozos de gas. Tiene una larga cola derecha. b) Histograma que muestra los logaritmos naturales de las producciones mensuales. La distribución de los datos logarítmicos se aproxima mucho más a la normal.

Ejercicios para la sección 4.6

- El tiempo de vida (en días) de cierto componente electrónico que opera en un ambiente a alta temperatura sigue una distribución lognormal con $\mu = 1.2$ y $\sigma = 0.4$.
 - Determine la media del tiempo de vida.
 - Determine la probabilidad de que un componente dure entre tres y seis días.
 - Determine la mediana del tiempo de vida.
 - Determine al 90o. percentil de los tiempos de vida.
- Cuando un pesticida entra en contacto con la piel, se absorbe cierto porcentaje de éste. El porcentaje del pesticida que será absorbido durante cierto espacio de tiempo puede modelarse con una distribución lognormal. Suponga que para cierto pesticida, la cantidad que es absorbida (en porcentaje) durante un periodo de dos horas sigue una distribución lognormal con $\mu = 1.5$ y $\sigma = 0.5$.
 - Determine la media del porcentaje absorbido.
 - Determine la mediana del porcentaje absorbido.
 - Determine la probabilidad de que el porcentaje absorbido sea mayor que 10.
 - Determine la probabilidad de que el porcentaje absorbido sea menor que 5.
- El índice de masa corporal (IMC) de una persona se define como la masa corporal de una persona dividido entre el cuadrado del peso de la persona. El artículo "Influences of Parameter Uncertainties within the ICRP 66 Respiratory Tract Model: Particle Deposition" (W. Bolch, E. Farfan y colaboradores, en *Health Physics*, 2001:378-394) establece que el índice de masa corporal (en kg/m^2) en hombres entre 25-34 años sigue una distribución lognormal con parámetros $\mu = 3.215$ y $\sigma = 0.157$.
 - Determine la media del IMC para hombres entre 25-34 años.
 - Determine la desviación estándar del IMC para hombres entre 25-34 años.
 - Determine la mediana del IMC para hombres entre 25-34 años.
 - ¿Qué proporción de hombres entre 25-34 años tiene un IMC menor a 22?
 - Encuentre el 75o. percentil de IMC para hombres entre 25-34 años.

4. El artículo “Stochastic Estimates of Exposure and Cancer Risk from Carbon Tetrachloride Released to the Air from the Rocky Flats Plant” (A. Rood, P. McGavran y colaboradores, en *Risk Analysis*, 2001:675-695) modela el aumento en el riesgo de cáncer debido a la exposición al tetracloruro de carbono como una lognormal con $\mu = -15.65$ y $\sigma = 0.79$.
- Determine la media del riesgo.
 - Determine la mediana del riesgo.
 - Determine la desviación estándar del riesgo.
 - Determine el 5o. percentil.
 - Determine el 95o. percentil.
5. El artículo “Withdrawal Strength of Threaded Nails” (D. Rammer, S. Winistorfer y D. Bender, en *Journal of Structural Engineering*, 2001:442-449) describe un experimento que compara la resistencia final a ser retirados (en N/mm) de varios tipos de clavos. Para un clavo con rosca anular y un vástago de diámetro de 3.76 mm introducido en una madera de pinabete, pino y abeto, la resistencia final de retiro fue modelada como una lognormal con $\mu = 3.82$ y $\sigma = 0.219$. Para un clavo con rosca en espiral, bajo las mismas condiciones, la resistencia se modeló como una lognormal con $\mu = 3.47$ y $\sigma = 0.272$.
- ¿Cuál es la media de la resistencia de retiro para los clavos con rosca anular?
 - ¿Cuál es la media de la resistencia de retiro para los clavos con rosca en espiral?
 - ¿Con qué tipo de clavo es más probable que la fuerza de retiro sea mayor a 50 N/mm?
 - ¿Cuál es la probabilidad de que un clavo con rosca en espiral tenga una resistencia mayor a ser retirado que la mediana de los clavos con rosca anular?
 - Se realiza un experimento en el cual se mide las resistencias a ser retirados de diversos clavos de los dos tipos. Se registra que un clavo tiene resistencia a ser retirado de 20 N/mm, pero no se sabe el tipo. ¿Piensa que se trata de un clavo con rosca anular o de un clavo con rosca en espiral? ¿Por qué? ¿Qué tan seguro está?
6. Elija la mejor respuesta, y explique. Si X es una variable aleatoria con una distribución lognormal, entonces _____
- la media de X siempre es mayor que la mediana.
 - la media de X siempre es menor que la mediana.
 - la media puede ser mayor que, menor que, o igual que la mediana, dependiendo del valor de σ .
7. Los precios de acciones u otros instrumentos financieros con frecuencia se modelan como una distribución lognormal. Un inversionista está considerando comprar acciones en una de dos compañías, A o B. Hoy el precio de una acción en ambas compañías es de un dólar. Para la A, el valor de la acción en un año a partir de ahora se modela como una lognormal con parámetros $\mu = 0.05$ y $\sigma = 0.1$. Para la B, el valor de la acción en un año a partir de ahora se modela como una lognormal con parámetros $\mu = 0.02$ y $\sigma = 0.2$.
- Determine la media del precio de una acción de la compañía A en un año a partir de ahora.
 - Determine la probabilidad de que el precio de una acción de la compañía A en un año a partir de ahora sea mayor a \$1.20.
 - Determine la media del precio de una acción de la compañía B en un año a partir de ahora.
 - Determine la probabilidad de que el precio de una acción de la compañía B en un año a partir de ahora sea mayor a \$1.20.
8. Un fabricante afirma que la resistencia a la tensión de cierto compuesto (en MPa) tiene una distribución lognormal con $\mu = 5$ y $\sigma = 0.5$. Sea X la resistencia de una muestra aleatoria y representativa de este compuesto.
- Si la afirmación es cierta, ¿a qué es igual $P(X < 20)$?
 - Con base en la respuesta al inciso (a), si la afirmación es cierta, ¿una resistencia de 20 MPa sería inusualmente pequeña?
 - Si usted observa una resistencia a la tensión de 20 MPa, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
 - Si la afirmación es cierta, ¿a qué es igual $P(X < 130)$?
 - Con base en la respuesta al inciso d), si la afirmación es cierta, ¿una resistencia de 130 MPa sería inusualmente pequeña?
 - Si usted observa una resistencia a la tensión de 130 MPa, ¿esto sería una evidencia de que la afirmación es falsa? Explique.

4.7 La distribución exponencial

La **distribución exponencial** es una distribución continua que algunas veces se utiliza para modelar el tiempo que transcurre antes de que ocurra un evento. A menudo, a aquél se le llama **tiempo de espera**. En algunas ocasiones la distribución exponencial se utiliza para modelar el tiempo de vida de un componente. Asimismo, hay una relación cercana entre la distribución exponencial y la distribución de Poisson.

La función de densidad de probabilidad de la distribución exponencial tiene un parámetro, que representa una constante positiva λ cuyo valor determina la localización y forma de la función.

Definición

La función de densidad de probabilidad de la distribución exponencial con parámetro $\lambda > 0$ es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.31)$$

La figura 4.17 presenta la función de densidad de probabilidad de la distribución exponencial para varios valores de λ . Si X es una variable aleatoria cuya distribución es exponencial con parámetro λ , se expresa como $X \sim \text{Exp}(\lambda)$.

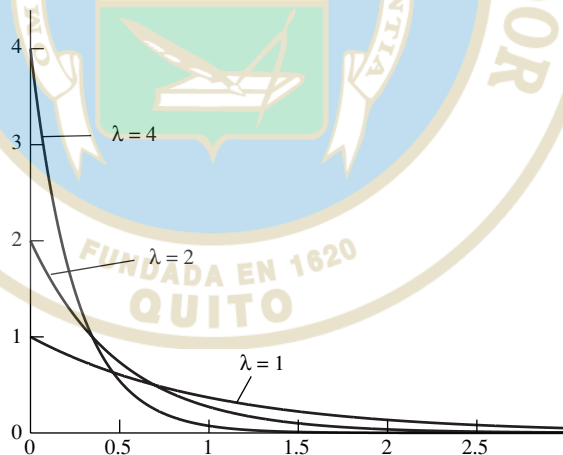


FIGURA 4.17 Gráficas de la función de densidad de probabilidad exponencial para varios valores de λ .

Es fácil calcular la función de distribución acumulativa de la distribución exponencial. Para $x \leq 0$, $F(x) = P(X \leq x) = 0$. Para $x > 0$, la función de distribución acumulativa es

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

Resumen

Si $X \sim \text{Exp}(\lambda)$, la función de distribución acumulativa de X es

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.32)$$

La media y la varianza de una variable aleatoria exponencial se pueden calcular mediante la integración por partes. Al final de la sección se muestra las deducciones.

Si $X \sim \text{Exp}(\lambda)$, entonces

$$\mu_X = \frac{1}{\lambda} \quad (4.33)$$

$$\sigma_X^2 = \frac{1}{\lambda^2} \quad (4.34)$$

Ejemplo

4.55

Si $X \sim \text{Exp}(2)$, encuentre μ_X , σ_X^2 y $P(X \leq 1)$.

Solución

Se calcula μ_X y σ_X^2 a partir de las ecuaciones (4.33) y (4.34), sustituyendo $\lambda = 2$. Se obtiene $\mu_X = 0.5$, $\sigma_X^2 = 0.25$. Utilizando la ecuación (4.32), se tiene que

$$P(X \leq 1) = 1 - e^{-2(1)} = 0.865$$

Ejemplo

4.56

Con referencia al ejemplo 4.55, determine la mediana de X . Encuentre el 30o. percentil de X .

Solución

Sea m la mediana de X . Entonces $P(X \leq m) = 0.5$. Mediante la ecuación (4.32) se tiene que $1 - e^{-2m} = 0.5$. Al despejar m , se obtiene que $m = 0.3466$.

Sea p_{30} el 30o. percentil. Entonces $P(X \leq p_{30}) = 0.30$. Utilizando la ecuación (4.32), se tiene que $1 - e^{-2p_{30}} = 0.30$. Al despejar p_{30} se obtiene que $p_{30} = 0.1783$.

La distribución exponencial y el proceso de Poisson

Se mencionó que algunas veces se utiliza la distribución exponencial para modelar el tiempo de espera de un evento. Resulta que la distribución exponencial es el modelo correcto para los tiempos de espera siempre y cuando los eventos sigan un proceso de Poisson. Recuerde de la sección 4.3 que los eventos que siguen un proceso Poisson con un parámetro de razón λ cuando los números de eventos en intervalos disjuntos son independientes, y el número X de eventos que ocurre en un intervalo con una longitud t tiene una distribución de Poisson con media

λt , es decir, cuando $X \sim \text{Poisson}(\lambda t)$. La conexión entre la distribución exponencial y el proceso de Poisson es la siguiente:

Si los eventos siguen un proceso de Poisson con un parámetro de razón λ , y si T representa el tiempo de espera desde cualquier punto inicial hasta el próximo evento, entonces $T \sim \text{Exp}(\lambda)$.

Al final de esta sección se demuestra este hecho.

Ejemplo 4.57

Una masa radiactiva emite partículas de acuerdo con un proceso de Poisson a una media de razón de 15 partículas por minuto. En algún punto inicia un reloj. ¿Cuál es la probabilidad de que transcurran cinco segundos antes de la siguiente emisión? ¿Cuál es la media del tiempo de espera hasta que se emite la siguiente partícula?

Solución

El tiempo se medirá en segundos. T denota el tiempo en segundos que transcurre antes de que se emita la siguiente partícula. La media de la razón de las emisiones es de 0.25 por segundo, por lo que el parámetro de razón es $\lambda = 0.25$ y $T \sim \text{Exp}(0.25)$. La probabilidad de que transcurran más de cinco segundos antes de la siguiente emisión es igual a

$$\begin{aligned} P(T > 5) &= 1 - P(T \leq 5) \\ &= 1 - (1 - e^{-0.25(5)}) \\ &= e^{-1.25} \\ &= 0.2865 \end{aligned}$$

La media del tiempo de espera es $\mu_T = \frac{1}{0.25} = 4$.

Propiedad de falta de memoria

La distribución exponencial tiene una propiedad conocida como falta de memoria, que se muestra en los ejemplos 4.58 y 4.59.

Ejemplo 4.58

El tiempo de vida de un circuito integrado particular tiene una distribución exponencial con media de dos años. Encuentre la probabilidad de que el circuito dure más de tres años.

Solución

Sea T el tiempo de vida del circuito. Dado que $\mu_T = 2$, $\lambda = 0.5$. Se necesita encontrar $P(T > 3)$.

$$\begin{aligned} P(T > 3) &= 1 - P(T \leq 3) \\ &= 1 - (1 - e^{-0.5(3)}) \\ &= e^{-1.5} \\ &= 0.223 \end{aligned}$$

Ejemplo

4.59

Con referencia al ejemplo 4.58, suponga que actualmente un circuito tiene cuatro años y aún funciona. Determine la probabilidad de que funcione tres años más. Compárela con la probabilidad de que un circuito nuevo funcione tres años más, la cual se calculó en el ejemplo 4.58.

Solución

Se pide que el tiempo de vida de un circuito sea más de cuatro años y se tiene que calcular la probabilidad de que el tiempo de vida sea mayor que $4 + 3 = 7$ años. La probabilidad está dada por

$$P(T > 7 | T > 4) = \frac{P(T > 7 \text{ y } T > 4)}{P(T > 4)}$$

Si $T > 7$, entonces también $T > 4$. Por consecuencia, $P(T > 7 \text{ y } T > 4) = P(T > 7)$. De ahí que

$$\begin{aligned} P(T > 7 | T > 4) &= \frac{P(T > 7)}{P(T > 4)} \\ &= \frac{e^{-0.5(7)}}{e^{-0.5(4)}} \\ &= e^{-0.5(3)} \\ &= e^{-1.5} \\ &= 0.223 \end{aligned}$$

La probabilidad de que un circuito con cuatro años dure tres años más es la misma probabilidad de que el circuito nuevo dure tres años.

Los ejemplos 4.58 y 4.59 muestran la propiedad de falta de memoria. La probabilidad de que se tenga que esperar t unidades adicionales, dado que ya se han esperado s unidades, es la misma que la probabilidad de que se tenga que esperar t unidades desde el inicio. La distribución exponencial no “recuerda” cuánto tiempo se ha esperado. En particular, si el tiempo de vida de un componente sigue una distribución exponencial, entonces la probabilidad de que un componente que tiene s unidades de tiempo dure t unidades de tiempo adicionales es la misma que la probabilidad de que un componente nuevo dure t unidades de tiempo. En otras palabras, un componente cuyo tiempo de vida siga una distribución exponencial no muestra ningún síntoma de los años o del uso.

Los cálculos en los ejemplos 4.58 y 4.59 se pueden repetir para cualesquiera valores s y t en lugar de 4 y 3, y para cualquier valor de λ en lugar de 0.5. Ahora, se establece la propiedad de falta de memoria en su forma general:

Propiedad falta de memoria

Si $T \sim \text{Exp}(\lambda)$, y t y s son números positivos, entonces

$$P(T > t + s | T > s) = P(T > t)$$

Ejemplo

4.60

El número de visitas a un sitio web sigue un proceso de Poisson con una razón de tres por minuto. ¿Cuál es la probabilidad de que transcurra más de un minuto sin recibir una visita? Si transcurren dos minutos sin una visita, ¿cuál es la probabilidad que se dé una visita en el siguiente minuto?

Solución

Sea T el tiempo de espera en minutos hasta la siguiente visita. Entonces $T \sim \text{Exp}(3)$. La probabilidad de que transcurra un minuto sin ninguna visita es $P(T > 1) = e^{-3(1)} = 0.0498$. Debido a la propiedad de falta de memoria, la probabilidad de que pase un minuto adicional sin ninguna visita, dado que han transcurrido dos minutos sin una visita, es también igual a 0.0498. Por tanto, la probabilidad de que ocurra una visita en el siguiente minuto es igual a $1 - 0.0498 = 0.9502$.

Uso de la distribución exponencial para estimar una razón

Si $X \sim \text{Exp}(\lambda)$, entonces $\mu_X = 1/\lambda$, por lo que $\lambda = 1/\mu_X$. De ahí que $X_1, \dots, X_n$ sea una muestra aleatoria de $\text{Exp}(\lambda)$, una estimación razonable de λ es $\hat{\lambda} = 1/\bar{X}$.

Se analizará el sesgo en $\hat{\lambda} = 1/\bar{X}$. Al igual que con cualquier media muestral $\bar{X}$, $\mu_{\bar{X}} = \mu$, por tanto, $\bar{X}$ es un estimador no sesgado de μ . No obstante, $\mu_{1/\bar{X}} \neq 1/\mu$, ya que $1/\mu$ no es una función lineal de μ . Por consecuencia, $\hat{\lambda} = 1/\bar{X}$ es un estimador sesgado de $\lambda = 1/\mu$. Utilizando métodos avanzados puede demostrarse que $\mu_{\hat{\lambda}} \approx \lambda + \lambda/n$; de esta forma el sesgo es aproximadamente λ/n . Por tanto, para una muestra de tamaño n lo suficientemente grande el sesgo es despreciable, pero puede ser importante cuando el tamaño de aquélla es pequeño.

Se puede estimar la incertidumbre en $\hat{\lambda}$ mediante el método de propagación de errores (ecuación 3.10 de la sección 3.3):

$$\sigma_{\hat{\lambda}} \approx \left| \frac{d}{d\bar{X}} \frac{1}{\bar{X}} \right| \sigma_{\bar{X}}$$

Para que esta expresión sea útil, es necesario conocer $\sigma_{\bar{X}}$. Ahora, que la desviación estándar de una distribución $\text{Exp}(\lambda)$ es $\sigma = 1/\lambda$ (lo cual se tiene a partir de la ecuación 4.34; observe que la desviación estándar es la misma que la media). Por consiguiente, $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 1/(\lambda\sqrt{n})$. Se puede sustituir λ con el estimador $1/\bar{X}$ para obtener

$$\sigma_{\bar{X}} \approx \frac{\bar{X}}{\sqrt{n}}$$

Ahora, se puede estimar la incertidumbre $\sigma_{\hat{\lambda}}$:

$$\begin{aligned} \sigma_{\hat{\lambda}} &\approx \left| \frac{d}{d\bar{X}} \frac{1}{\bar{X}} \right| \sigma_{\bar{X}} \\ &= \frac{1}{\bar{X}^2} \frac{\bar{X}}{\sqrt{n}} \\ &= \frac{1}{\bar{X}\sqrt{n}} \end{aligned}$$

Esta estimación con el método de propagación de errores es muy bueno cuando el tamaño muestral es de al menos 20 o algo así. Para tamaños menores de muestra, éste subestima la incertidumbre.

Resumen

Si $X_1, \dots, X_n$ es una muestra aleatoria de $\text{Exp}(\lambda)$, entonces el parámetro λ se estima con

$$\hat{\lambda} = \frac{1}{\bar{X}} \quad (4.35)$$

Este estimador es sesgado. El sesgo es aproximadamente igual a λ/n . La incertidumbre en $\hat{\lambda}$ se estima con

$$\sigma_{\hat{\lambda}} = \frac{1}{\bar{X}\sqrt{n}} \quad (4.36)$$

Este estimador de la incertidumbre es razonablemente bueno cuando el tamaño muestral es mayor a 20.

Corrección del sesgo

Dado que $\mu_{\hat{\lambda}} = \mu_{1/\bar{X}} \approx \lambda + \lambda/n = (n+1)\lambda/n$, se tiene que $\mu_{n/(n+1)\bar{X}} \approx \lambda$. En otras palabras, la cantidad $n/[(n+1)\bar{X}]$ es un estimador menos sesgado de λ que $1/\bar{X}$. A esto último se le conoce como estimador *con corrección de sesgo*.

Ejemplo 4.61

Se toma una muestra aleatoria de tamaño 5 de una distribución $\text{Exp}(\lambda)$. Los valores son 7.71, 1.32, 7.46, 6.53 y 0.44. Encuentre un estimador con corrección de sesgo de λ .

Solución

La media muestral es $\bar{X} = 4.6920$. El tamaño muestral es $n = 5$. El estimador con corrección de sesgo de λ es $5/[6(4.6920)] = 0.178$.

Deducción de la media y la varianza de una variable aleatoria exponencial

Para deducir la ecuación (4.33) se comienza con la ecuación (2.35) (de la sección 2.4):

$$\mu_X = \int_{-\infty}^{\infty} xf(x) dx$$

Al sustituir la función de densidad de probabilidad exponencial (4.31) para $f(x)$, se obtiene

$$\mu_X = \int_0^{\infty} \lambda x e^{-\lambda x} dx$$

Al integrar por partes, haciendo $u = x$ y $dv = \lambda e^{-\lambda x}$ se obtiene

$$\mu_X = -xe^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \quad (4.37)$$

Se evalúa la primera cantidad en el lado derecho de la ecuación (4.37):

$$\begin{aligned}
 -xe^{-\lambda x} \Big|_0^{\infty} &= \lim_{x \rightarrow \infty} -xe^{-\lambda x} - 0 \\
 &= \lim_{x \rightarrow \infty} -\frac{x}{e^{\lambda x}} \\
 &= \lim_{x \rightarrow \infty} -\frac{1}{\lambda e^{\lambda x}} && \text{por la regla de L'Hospital} \\
 &= 0
 \end{aligned}$$

Por consecuencia,

$$\begin{aligned}
 \mu_X &= \int_0^{\infty} e^{-\lambda x} dx \\
 &= \frac{1}{\lambda}
 \end{aligned}$$

Para deducir la ecuación (4.34) se comienza con la ecuación (2.37) (de la sección 2.4):

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu_X^2$$

Al sustituir la función de densidad de probabilidad exponencial (4.31) por $f(x)$ y $1/\lambda$ por μ_X , se obtiene

$$\sigma_X^2 = \int_0^{\infty} \lambda x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} \quad (4.38)$$

Se evalúa la integral $\int_0^{\infty} \lambda x^2 e^{-\lambda x} dx$, utilizando la integración por partes. Al hacer $u = x^2$ y $dv = \lambda e^{-\lambda x}$ se obtiene

$$\int_0^{\infty} \lambda x^2 e^{-\lambda x} dx = -x^2 e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx \quad (4.39)$$

Se evalúa la primera cantidad en el lado derecho de la ecuación (4.39):

$$\begin{aligned}
 -x^2 e^{-\lambda x} \Big|_0^{\infty} &= \lim_{x \rightarrow \infty} -x^2 e^{-\lambda x} - 0 \\
 &= \lim_{x \rightarrow \infty} -\frac{x^2}{e^{\lambda x}} \\
 &= \lim_{x \rightarrow \infty} -\frac{2x}{\lambda e^{\lambda x}} && \text{por la regla de L'Hospital} \\
 &= \lim_{x \rightarrow \infty} -\frac{2}{\lambda^2 e^{\lambda x}} && \text{por la regla de L'Hospital} \\
 &= 0
 \end{aligned}$$

Por tanto,

$$\begin{aligned}
 \int_0^{\infty} \lambda x^2 e^{-\lambda x} dx &= \int_0^{\infty} 2x e^{-\lambda x} dx \\
 &= \frac{2}{\lambda} \int_0^{\infty} \lambda x e^{-\lambda x} dx \\
 &= \frac{2}{\lambda} \mu_X \\
 &= \left(\frac{2}{\lambda}\right) \left(\frac{1}{\lambda}\right) \\
 &= \frac{2}{\lambda^2}
 \end{aligned}$$

Al sustituir en (4.38) se obtiene

$$\begin{aligned}
 \sigma_X^2 &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
 &= \frac{1}{\lambda^2}
 \end{aligned}$$

Deducción de la relación entre la distribución exponencial y el proceso de Poisson

Sea T el tiempo de espera hasta el siguiente evento en un proceso de Poisson con un parámetro de razón λ . Es evidente que $T \sim \text{Exp}(\lambda)$ al demostrar que la función de distribución acumulativa de T es $F(t) = 1 - e^{-\lambda t}$, que es la función de distribución acumulativa de $\text{Exp}(\lambda)$.

Primero, si $t \leq 0$, entonces $F(t) = P(T \leq t) = 0$. Ahora $t > 0$. Se comienza por calcular $P(T > t)$. La clave es considerar que $T > t$, si y sólo si, no ocurre ningún evento durante las siguientes t unidades de tiempo. Sea X el número de eventos que sucede en las siguientes t unidades de tiempo. Ahora $T > t$ si y sólo si $X = 0$, por lo que $P(T > t) = P(X = 0)$.

Puesto que $X \sim \text{Poisson}(\lambda t)$,

$$\begin{aligned}
 P(X = 0) &= e^{-\lambda t} \frac{\lambda^0}{0!} \\
 &= e^{-\lambda t}
 \end{aligned}$$

Por consecuencia, $P(T > t) = e^{-\lambda t}$. La función de distribución acumulativa de T es $F(t) = 0$ para $t \leq 0$, y para $t > 0$

$$\begin{aligned}
 F(t) &= P(T \leq t) \\
 &= 1 - P(T > t) \\
 &= 1 - e^{-\lambda t}
 \end{aligned}$$

Dado que $F(t)$ es la función de distribución acumulativa de $\text{Exp}(\lambda)$, se tiene que $T \sim \text{Exp}(\lambda)$.

Ejercicios para la sección 4.7

- $T \sim \text{Exp}(0.45)$. Determine
 - μ_T
 - σ_T^2
 - $P(T > 3)$
 - La mediana de T
- El tiempo de vida de un fusible en cierta aplicación tiene distribución exponencial con media de dos años.
 - ¿Cuál es el valor del parámetro λ ?
 - ¿Cuál es la mediana del tiempo de vida de dicho fusible?
 - ¿Cuál es la desviación estándar?
 - ¿Cuál es el 60o. percentil?
 - Determine la probabilidad de que un fusible dure más de cinco años.
 - Si un fusible tiene un año y aún sigue funcionando, ¿cuál es la probabilidad de que funcione dos años más?
- Una investigadora de catalizadores afirma que los diámetros, en micrones, de los poros de un nuevo producto que ella ha fabricado sigue una distribución exponencial con parámetro $\lambda = 0.25$.
 - ¿Cuál es la media del diámetro de los poros?
 - ¿Cuál es la desviación estándar de los diámetros de los poros?
 - ¿Qué proporción de los poros tiene un diámetro menor a tres micrones?
 - ¿Qué proporción de los poros tiene un diámetro mayor a 11 micrones?
 - ¿Cuál es la mediana del diámetro de los poros?
 - ¿Cuál es el tercer cuartil de los diámetros de los poros?
 - ¿Cuál es el 99o. percentil de los diámetros de los poros?
- Alguien argumenta que el tiempo de espera, en minutos, entre las visitas a un sitio web tiene una distribución exponencial con parámetro $\lambda = 1$.
 - Sea X el tiempo de espera hasta la siguiente visita. Si la afirmación es verdadera, ¿a qué es igual $P(X \geq 5)$?
 - Con base en la respuesta al inciso (a), si la afirmación es verdadera, ¿esperar cinco minutos es un tiempo de espera inusualmente largo?
- Si usted hubiera esperado cinco minutos para que ocurriera la siguiente visita, ¿aún seguiría creyendo en la afirmación? Explique.
 - Si usted hubiera esperado cinco minutos para que ocurriera la siguiente visita, ¿aún seguiría creyendo en la afirmación? Explique.
- Cierto tipo de componente puede ser comprado nuevo o viejo. El 50% de los componentes nuevos duran más de cinco años, pero sólo 30% de los usados dura más de cinco años. ¿Sería posible que las duraciones de los componentes se distribuyan exponencialmente? Explique.
 - Una masa radiactiva emite partículas de acuerdo con un proceso de Poisson a una razón media de dos por segundo. Sea T el tiempo de espera, en segundos, entre las emisiones.
 - ¿Cuál es la media del tiempo de espera?
 - ¿Cuál es la mediana del tiempo de espera?
 - Determine $P(T > 2)$.
 - Determine $P(T < 0.1)$.
 - Determine $P(0.3 < T < 1.5)$.
 - Si han transcurrido tres segundos sin que haya ninguna emisión, ¿cuál es la probabilidad de que haya una emisión dentro del siguiente segundo?
- Se considera que el número de accidentes de tránsito en cierta intersección sigue el modelo de un proceso de Poisson con una media de tres accidentes al año.
 - Determine la media del tiempo de espera entre los accidentes.
 - Determine la desviación estándar de los tiempos de espera entre los accidentes.
 - Determine la probabilidad de que transcurra más de un año entre un accidente y otro.
 - Determine la probabilidad de que transcurra menos de un mes entre un accidente y otro.
 - Si no ocurre ningún accidente en los últimos seis meses, ¿cuál es la probabilidad de que suceda un accidente durante el siguiente año?
- La distancia entre imperfecciones consecutivas en un rollo de lámina de aluminio se distribuye exponencialmente con una distancia media de 3 m. Sea X la distancia, en metros, entre las imperfecciones.
 - ¿Cuál es la media del número de imperfecciones por metro?
 - ¿Cuál es la probabilidad de que cinco metros de aluminio tengan sólo dos imperfecciones?

9. Una instalación de luz contiene cinco focos. La duración de cada foco se distribuye exponencialmente con una media de 200 horas. Cada vez que se funde un foco, éste se reemplaza. Sea T el momento del primer reemplazo del foco. Sea X_i , $i = 1, \dots, 5$ la duración de cinco focos. Suponga que la duración de los focos son independientes.
- Determine $P(X_1 > 100)$.
 - Determine $P(X_1 > 100 \text{ y } X_2 > 100 \text{ y } \dots \text{ y } X_5 > 100)$.
 - Explique por qué el evento $T > 100$ es el mismo que $\{X_1 > 100 \text{ y } X_2 > 100 \text{ y } \dots \text{ y } X_5 > 100\}$.
 - Determine $P(T \leq 100)$.
 - T es cualquier número positivo. Encuentre $P(T \leq t)$, ¿cuál es la función de distribución acumulativa de T ?
 - ¿ T tiene una distribución exponencial?
 - Determine la media de T .
 - Si hubiera n focos, y la duración de cada uno se distribuyera exponencialmente con parámetro λ , ¿cuál sería la distribución de T ?

4.8 Las distribuciones gamma y de Weibull

Las distribuciones gamma y de Weibull son extensiones de la distribución exponencial. Ambas implican una integral conocida como la **función gamma**. Primero se define la función gamma y se establecen algunas de sus propiedades.

Definición

Para $r > 0$, la función gamma está definida por

$$\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt \quad (4.40)$$

La función gamma tiene las siguientes propiedades:

- Si r es un entero, entonces $\Gamma(r) = (r-1)!$
- Para cualquier r , $\Gamma(r+1) = r\Gamma(r)$.
- $\Gamma(1/2) = \sqrt{\pi}$.

La distribución gamma

La **distribución gamma** es una distribución continua, uno de sus propósitos es ampliar la utilidad de la distribución exponencial en el modelado de tiempos de espera. La función de densidad de probabilidad gamma tiene dos parámetros, r y λ , que son constantes positivas.

Definición

La función de densidad de probabilidad de la distribución gamma con parámetros $r > 0$ y $\lambda > 0$ es

$$f(x) = \begin{cases} \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.41)$$

Si X es una variable aleatoria cuya función de densidad de probabilidades gamma con parámetros r y λ , se expresa como $X \sim \Gamma(r, \lambda)$. Observe que cuando $r = 1$, la distribución gamma es igual a la exponencial. Con símbolos, $\Gamma(1, \lambda) = \text{Exp}(\lambda)$. La figura 4.18 muestra gráficas de la función de densidad de probabilidad gamma para varios valores de r y λ .

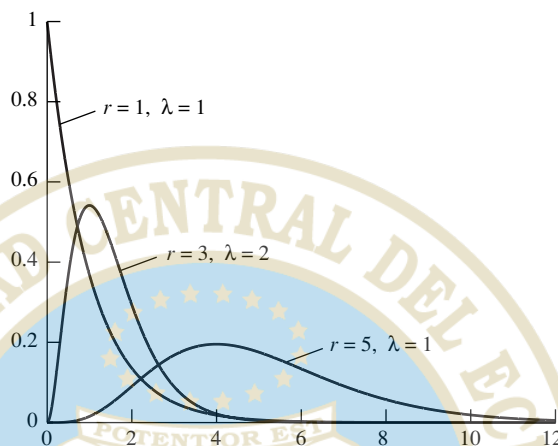


FIGURA 4.18 Función de densidad de probabilidad gamma para varios valores de r y λ .

Cuando el parámetro r es un entero, la distribución gamma es una extensión directa de la distribución exponencial. Para ser más específicos, recuerde que si los eventos seguían un proceso de Poisson con parámetro de razón λ , el tiempo de espera hasta que ocurriera un evento se distribuía como $\text{Exp}(\lambda)$. Si r es cualquier entero positivo, entonces el tiempo de espera hasta que haya ocurrido r eventos se distribuye como $\Gamma(r, \lambda)$. Esto puede decirse de otra manera. Sea X_1 el tiempo de espera hasta el primer evento, y, para $i > 1$, sea X_i el tiempo de espera entre los eventos $i - 1$ e i . El tiempo de espera hasta el r -ésimo evento es la suma de las variables aleatorias independientes $X_1 + \dots + X_r$, cada uno de los cuales se distribuye como $\text{Exp}(\lambda)$.

Resumen

Si $X_1, \dots, X_n$ son variables aleatorias independientes, cada una se distribuye como $\text{Exp}(\lambda)$, entonces la suma $X_1 + \dots + X_r$ se distribuye como $\Gamma(r, \lambda)$.

Dado que la media y la varianza de una variable aleatoria exponencial está dada por $1/\lambda$ y $1/\lambda^2$, respectivamente, se puede utilizar el hecho de que una variable aleatoria gamma es la suma de variables aleatorias exponenciales independientes para calcular la media y la varianza de una variable aleatoria gamma en el caso en que r es un entero. Los resultados se presentan en el siguiente cuadro, y de hecho, son válidos para todos los valores de r y λ .

Si $X \sim \Gamma(r, \lambda)$, entonces

$$\mu_X = \frac{r}{\lambda} \quad (4.42)$$

$$\sigma_X^2 = \frac{r}{\lambda^2} \quad (4.43)$$

Ejemplo 4.62

Suponga que los tiempos de llegada a una ventanilla de toma de órdenes desde el automóvil sigue un proceso de Poisson con media de razón $\lambda = 0.2$ llegadas por minuto. Sea T el tiempo de espera hasta la tercera llegada. Determine la media y varianza de T . Encuentre $P(T \leq 20)$.

Solución

La variable aleatoria T se distribuye $\Gamma(3, 0.2)$. Utilizando las ecuaciones (4.42) y (4.43) se calcula $\mu_T = 3/0.2 = 15$ y $\sigma_T^2 = 3/(0.2^2) = 75$. Para calcular $P(T \leq 20)$ se sigue la siguiente lógica: $T \leq 20$ significa que el tercer evento ocurrirá dentro de 20 minutos. Esto es lo mismo que si se dijera que el número de eventos que ocurren dentro de 20 minutos es mayor que o igual a tres. Ahora, sea X el número de eventos que ocurren dentro de 20 minutos. Lo que se ha dicho es que $P(T \leq 20) = P(X \geq 3)$. Ahora la media de X es $(20)(0.2) = 4$ y X tiene una distribución de Poisson, por lo que $X \sim \text{Poisson}(4)$. De ahí que

$$\begin{aligned} P(T \leq 20) &= P(X \geq 3) \\ &= 1 - P(X \leq 2) \\ &= 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \\ &= 1 - \left(e^{-4} \frac{4^0}{0!} + e^{-4} \frac{4^1}{1!} + e^{-4} \frac{4^2}{2!} \right) \\ &= 1 - (e^{-4} + 4e^{-4} + 8e^{-4}) \\ &= 0.7619 \end{aligned}$$

El método empleado en el ejemplo 4.62 para determinar a $P(T \leq 20)$ se puede utilizar para determinar la función de distribución acumulativa $F(x) = P(T \leq x)$, cuando $T \sim \Gamma(r, \lambda)$ y r es un entero positivo.

Si $T \sim \Gamma(r, \lambda)$ y r es un entero positivo, la función de distribución acumulativa de T está dada por

$$F(x) = P(T \leq x) = \begin{cases} 1 - \sum_{j=0}^{r-1} e^{-\lambda x} \frac{(\lambda x)^j}{j!} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.44)$$

A una distribución gamma en la cual el parámetro r es un entero positivo se le denomina algunas veces **distribución de Erlang**. Si $r = k/2$ donde k es un entero positivo, a la distribución $\Gamma(r, 1/2)$ se le llama **distribución Ji-cuadrada con k grados de libertad**. La distribución Ji-cuadrada es muy importante en la inferencia estadística. Se analizarán algunos de sus usos en la sección 6.10.

Distribución de Weibull

La distribución de Weibull constituye una distribución continua que se utiliza en varias situaciones. Una aplicación común es modelar los tiempos de vida de componentes, como cojinetes, cerámica, capacitores y dieléctricos. La función de densidad de probabilidad de Weibull tiene dos parámetros, ambos constantes positivas, que determinan su localización y forma. Éstos se representan por α y β . La función de densidad de probabilidad de la distribución de Weibull es

$$f(x) = \begin{cases} \alpha\beta^\alpha x^{\alpha-1} e^{-(\beta x)^\alpha} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.45)$$

Si X es una variable aleatoria cuya función de densidad de probabilidad de Weibull con parámetros α y β , se expresa como $X \sim \text{Weibull}(\alpha, \beta)$. Observe que cuando $\alpha = 1$, la distribución de Weibull es igual que la distribución exponencial con parámetro $\lambda = \beta$. Con símbolos, $\text{Weibull}(1, \beta) = \text{Exp}(\beta)$.

La figura 4.19 muestra gráficas de la función de densidad de probabilidad de Weibull(α, β) para diversas opciones de los parámetros α y β . Al variar los valores de α y β se puede generar gran variedad de curvas. Debido a esto último, se puede construir la distribución de Weibull para que se ajuste a gran variedad de conjuntos de datos. Ésta es la principal razón de la utilidad de la distribución de Weibull.

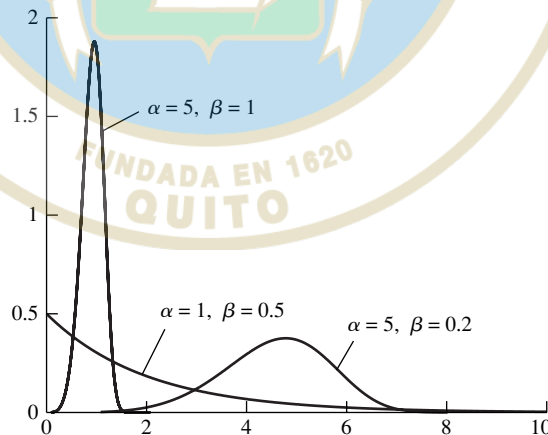


FIGURA 4.19 La función de densidad de probabilidad de Weibull para varias elecciones de α y β .

La función de distribución acumulativa de Weibull se puede calcular al integrar la función de densidad de probabilidad:

$$F(x) = P(X \leq x) = \begin{cases} \int_0^x \alpha \beta^\alpha t^{\alpha-1} e^{-(\beta t)^\alpha} dt = 1 - e^{-(\beta x)^\alpha} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (4.46)$$

Esta integral no es tan difícil como parece. Sólo se sustituyen $u = (\beta t)^\alpha$ y $du = \alpha \beta^\alpha t^{\alpha-1} dt$.

La media y la varianza de la distribución de Weibull se expresan en términos de la función gamma.

Si $X \sim \text{Weibull}(\alpha, \beta)$, entonces

$$\mu_X = \frac{1}{\beta} \Gamma\left(1 + \frac{1}{\alpha}\right) \quad (4.47)$$

$$\sigma_X^2 = \frac{1}{\beta^2} \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^2 \right\} \quad (4.48)$$

Para el caso especial de que $1/\alpha$ sea un entero, entonces

$$\mu_X = \frac{1}{\beta} \left[\left(\frac{1}{\alpha} \right)! \right] \quad \sigma_X^2 = \frac{1}{\beta^2} \left\{ \left(\frac{2}{\alpha} \right)! - \left[\left(\frac{1}{\alpha} \right)! \right]^2 \right\}$$

Si la cantidad $1/\alpha$ es un entero, entonces $1 + 1/\alpha$ y $1 + 2/\alpha$ son enteros, por lo que se puede aplicar la propiedad 1 de la función gamma para calcular exactamente a μ_X y σ_X^2 . Si la cantidad $1/\alpha$ es de la forma $n/2$, en donde n es un entero, entonces, en principio, se puede calcular exactamente μ_X y σ_X^2 mediante aplicaciones repetitivas de las propiedades 2 y 3 de la función gamma. Para otros valores de α , μ_X y σ_X^2 deben ser aproximadas. Muchos programas de computación pueden hacer esto.

Ejemplo

4.63

En el artículo “Snapshot: A Plot Showing Program through a Device Development Laboratory” (D. Lambert, J. Landwehr y M. Shyu, en *Statistical Case Studies for Industrial Process Improvement*, ASA-SIAM, 1997), los autores sugieren utilizar una distribución de Weibull para modelar la duración de un proceso de horneado en la fabricación de un semiconductor. Sea T la duración en horas del proceso de horneado de una muestra elegida aleatoriamente. Si $T \sim \text{Weibull}(0.3, 0.1)$, ¿cuál es la probabilidad de que el proceso de horneado dure más de cuatro horas? ¿Cuál es la probabilidad de que dure entre dos y siete horas?

Solución

Se utiliza la función de distribución acumulativa, ecuación (4.46). Al sustituir 0.3 por α y 0.1 por β , se tiene

$$P(T \leq t) = 1 - e^{-(0.1t)^{0.3}}$$

En consecuencia,

$$\begin{aligned} P(T > 4) &= 1 - P(T \leq 4) \\ &= 1 - (1 - e^{-(0.1(4))^{0.3}}) \end{aligned}$$

$$\begin{aligned}
 &= e^{-(0.4)^{0.3}} \\
 &= e^{-0.7597} \\
 &= 0.468
 \end{aligned}$$

La probabilidad de que el proceso dure entre dos y siete horas es

$$\begin{aligned}
 P(2 < T < 7) &= P(T \leq 7) - P(T \leq 2) \\
 &= (1 - e^{-(0.1)(7)^{0.3}}) - (1 - e^{-(0.1)(2)^{0.3}}) \\
 &= e^{-(0.1)(2)^{0.3}} - e^{-(0.1)(7)^{0.3}} \\
 &= e^{-(0.2)^{0.3}} - e^{-(0.7)^{0.3}} \\
 &= e^{-0.6170} - e^{-0.8985} \\
 &= 0.132
 \end{aligned}$$

Ejercicios para la sección 4.8

1. Sea $T \sim \Gamma(4, 0.5)$.

- Determine μ_T .
- Determine σ_T .
- Determine $P(T \leq 1)$.
- Determine $P(T \geq 4)$.

2. La duración, en años, de un tipo de motor eléctrico pequeño operando en condiciones adversas se distribuye exponencialmente con $\lambda = 3.6$. Cada vez que falla un motor, es reemplazado por otro del mismo tipo. Determine la probabilidad de que menos de seis motores falle dentro de un año.

3. Sea $T \sim \text{Weibull}(0.5, 3)$.

- Determine μ_T .
- Determine σ_T .
- Determine $P(T < 1)$.
- Determine $P(T > 5)$.
- Determine $P(2 < T < 4)$.

4. Si T es una variable aleatoria continua que siempre es positiva (como el tiempo de espera), con una función de densidad de probabilidad $f(t)$ y una función de distribución acumulativa $F(t)$, entonces se define que la **función de riesgo** es la función

$$h(t) = \frac{f(t)}{1 - F(t)}$$

La función de riesgo es la tasa de fallos por unidad de tiempo, expresada como una proporción de los elementos que no ha fallado.

- Si $T \sim \text{Weibull}(\alpha, \beta)$, encuentre $h(t)$.

- ¿En qué valores de α la función de riesgo aumenta junto con el tiempo? ¿En qué valores de α la función de riesgo disminuye?

- Si T tiene una distribución exponencial, demuestre que la función de riesgo es constante.

5. En el artículo “Parameter Estimation with Only One Complete Failure Observation” (W. Pang, P. Leung y colaboradores, en *International Journal of Reliability, Quality, and Safety Engineering*, 2001:109-122), se modela la duración, en horas, de cierto tipo de cojinete con la distribución de Weibull con parámetros $\alpha = 2.25$ y $\beta = 4.474 \times 10^{-4}$.

- Determine la probabilidad de que un cojinete dure más de 1 000 horas.
- Determine la probabilidad de que un cojinete dure menos de 2 000 horas.
- Determine la mediana de la duración de un cojinete.
- La función de riesgo se definió en el ejercicio 4. ¿Cuál es el riesgo en $t = 2\,000$ horas?

6. La duración de cierta batería se modela con la distribución de Weibull con $\alpha = 2$ y $\beta = 0.1$.

- ¿Qué proporción de baterías durará más de diez horas?
- ¿Qué proporción de baterías durará menos de cinco horas?
- ¿Qué proporción de baterías durará más de 20 horas?
- La función de riesgo se definió en el ejercicio 4. ¿Cuál es el riesgo en $t = 10$ horas?

7. La duración de un ventilador, en horas, que se usa en un sistema computacional tiene una distribución de Weibull con $\alpha = 1.5$ y $\beta = 0.0001$.
- ¿Cuál es la probabilidad de que un ventilador dure más de 10 000 horas?
 - ¿Cuál es la probabilidad de que un ventilador dure menos de 5 000 horas?
 - ¿Cuál es la probabilidad de que un ventilador dure entre 3 000 y 9 000 horas?
8. Alguien sugiere que la duración T (en días) de cierta componente se puede modelar con una distribución de Weibull con parámetros $\alpha = 3$ y $\beta = 0.01$.
- Si este modelo es correcto, ¿a qué es igual $P(T \leq 1)$?
 - Con base en su respuesta al inciso (a), si el modelo es correcto, ¿un día sería una duración inusualmente corta? Explique.
 - Si usted observó que un componente duró un día, ¿consideraría que este modelo es plausible? Explique.
 - Si este modelo es correcto, ¿a qué es igual $P(T \leq 90)$?
- e) Con base en su respuesta al inciso (d), si el modelo es correcto, ¿90 días sería una duración inusualmente corta?, ¿una duración inusualmente larga? Explique.
- f) Si usted observó que un componente duró 90 días, ¿consideraría que este modelo es plausible? Explique.
9. Un sistema consiste de dos componentes conectados en serie. El sistema fallará cuando alguno de los componentes falle. Sea T el momento en el que el sistema falla. Sean X_1 y X_2 las duraciones de los dos componentes. Suponga que X_1 y X_2 son independientes y que cada uno sigue una distribución de Weibull con $\alpha = 2$ y $\beta = 0.2$.
- Determine $P(X_1 > 5)$.
 - Determine $P(X_1 > 5 \text{ y } X_2 > 5)$.
 - Explique por qué el evento $T > 5$ es el mismo que el evento $\{X_1 > 5 \text{ y } X_2 > 5\}$.
 - Determine $P(T \leq 5)$.
 - Sea t cualquier número positivo. Encuentre $P(T \leq t)$, ¿cuál es la función de distribución acumulativa de T ?
 - ¿ T tiene una distribución de Weibull? Si es así, ¿cuáles son sus parámetros?

4.9 Gráficas de probabilidad

Los científicos e ingenieros trabajan con frecuencia con datos que se pueden considerar como una muestra aleatoria de cierta población. En muchos de dichos casos, es importante determinar una distribución de probabilidad que describa aproximadamente la población. En algunos casos, el conocimiento del proceso que generó los datos puede orientar la decisión. No obstante, con mayor frecuencia, la única forma de determinar una distribución apropiada es examinar la muestra para encontrar una distribución de probabilidad que se ajuste.

Las gráficas de probabilidad son una buena forma para lograrlo. Dada una muestra aleatoria $X_1, \dots, X_n$, una gráfica de probabilidad puede determinar si es posible que la muestra provenga de alguna población específica. Se presentará la idea que está detrás de las gráficas de probabilidad con un ejemplo simple. Se extrae una muestra aleatoria de tamaño 5 y se desea determinar si la población de que proviene es normal. La muestra, en orden ascendente, es

3.01, 3.35, 4.79, 5.96, 7.89

Los valores, en orden ascendente, se denotan por $X_1, \dots, X_n$ ($n = 5$ en este caso). Lo primero que debe hacerse es asignar valores crecientes espaciados uniformemente entre 0 y 1 a X_i . Hay muchas maneras aceptables de hacer esto; probablemente la más simple es asignar el valor $(i - 0.5)/n$ a X_i . La siguiente tabla muestra la asignación para determinada muestra.

i	X_i	$(i - 0.5)/5$
1	3.01	0.1
2	3.35	0.3
3	4.79	0.5
4	5.96	0.7
5	7.89	0.9

Se elige al valor $(i - 0.5)/n$ para que refleje la posición de X_i en la muestra ordenada. Existen $i - 1$ valores menores que X_i , además de i valores menores que o iguales a X_i . La cantidad $(i - 0.5)/n$ es un compromiso entre las proporciones $(i - 1)/n$ e i/n .

El objetivo es determinar si es posible que la muestra provenga de una población normal. La distribución normal más factible es aquella donde la media y desviación estándar son iguales a la media y desviación estándar muestrales. La media muestral es $\bar{X} = 5.00$ y la desviación estándar es $s = 2.00$. Por tanto, se determinará si es posible que esta muestra provenga de una distribución $N(5, 2^2)$. La figura 4.20 es una gráfica de los cinco puntos $(X_i, (i - 0.5)/5)$. Esta curva es la función de distribución acumulativa (fda) $F(x)$ de la distribución $N(5, 2^2)$. Recuerde que $F(x) = P(X \leq x)$ donde $X \sim N(5, 2^2)$.

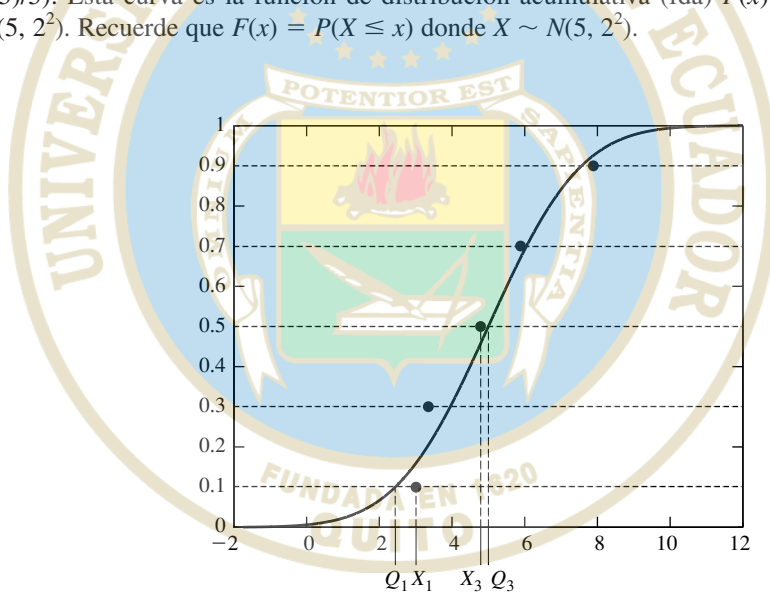


FIGURA 4.20 La curva es la fda de $N(5, 2^2)$. Si los puntos muestrales $X_1, \dots, X_n$ provienen de esta distribución, es probable que estén cerca de la curva.

Se han trazado rectas horizontales a través de los puntos de la muestra. Se denotan los valores x de los puntos de la fda que son atravesados por rectas, en orden ascendente, por $Q_1, \dots, Q_5$. Ahora, la recta horizontal que pasa por $(X_1, 0.1)$ intercepta a la fda en el punto $(Q_1, 0.1)$. Lo anterior significa que la proporción de valores en la población $N(5, 2^2)$ que es menor que o igual a Q_1 es 0.1. Otra forma de decir esto es que Q_1 es el 10o. percentil de la distribución $N(5, 2^2)$. Si la muestra $X_1, \dots, X_n$ verdaderamente proviene de una distribución $N(5, 2^2)$,

entonces es razonable pensar que el valor más bajo de la muestra, X_1 , estará muy cerca del 10o. percentil de la población. De forma intuitiva, la razón de esto último es que se esperaría que el más bajo de los cinco puntos probablemente proviniera del quinto más bajo, o 20%, de la población, y que el 10o. percentil se encontrará a la mitad de ese 20% más bajo. Al aplicar la misma lógica a los puntos restantes, se esperaría que cada Q_i se aproximara a su X_i correspondiente.

La **gráfica de probabilidad** está compuesta por los puntos (X_i, Q_i) . Puesto que la distribución que generó los Q_i era una distribución normal, se le llama **gráfica de probabilidad normal**. Si, de hecho, $X_1, \dots, X_n$ provienen de la distribución que generó los Q_i , los puntos se deben concentrar en una línea recta. Para construir la gráfica, se deben calcular los Q_i . Éstos son los percentiles $100(i - 0.5)/n$ de la distribución que se supone generó la muestra. En este ejemplo, los Q_i son el 10, 30, 50, 70 y 90o. percentiles de la distribución $N(5, 2^2)$. Es posible aproximar estos valores al buscar los puntajes z correspondientes a estos percentiles, y después convertirlos en nuevos puntajes. En la práctica, los Q_i se calculan siempre utilizando un software computacional. La siguiente tabla muestra los X_i y los Q_i para este ejemplo.

i	X_i	Q_i
1	3.01	2.44
2	3.35	3.95
3	4.79	5.00
4	5.96	6.05
5	7.89	7.56

La figura 4.21 presenta una gráfica de probabilidad normal para la muestra $X_1, \dots, X_5$. Se traza una línea recta en la gráfica, para que sea más fácil decir si los puntos se aproximan

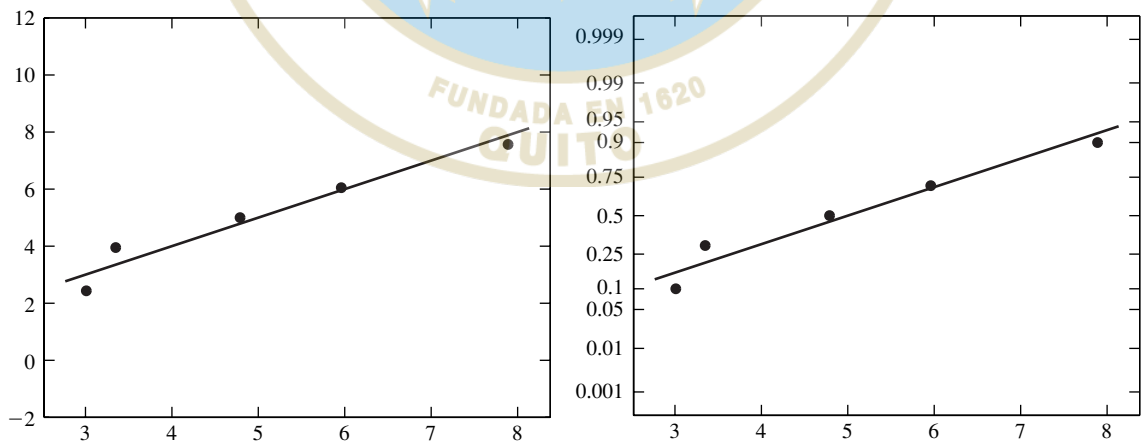


FIGURA 4.21 Gráficas de probabilidad normal para la muestra $X_1, \dots, X_5$. Las gráficas son idénticas, excepto por la escala del eje vertical. Los puntos de la muestra se aproximan a una línea recta, por lo que es factible que provengan de una población normal.

o no a una línea recta. Se presenta dos versiones de la gráfica; son idénticas excepto por la escala del eje vertical. En la gráfica de la izquierda, los valores del eje vertical representan los Q_i . En la gráfica de la derecha, los valores del eje vertical representan los percentiles (como decimales, por lo que 0.1 es el 10o. percentil) de los Q_i . Por ejemplo, el 10o. percentil de $N(5, 2^2)$ es 2.44, de tal forma que el valor 0.1 en la gráfica de la derecha corresponde al valor 2.44 de la gráfica de la izquierda. El 50o. percentil, o mediana, es 5, por lo que el valor 0.5 en la gráfica de la derecha corresponde al valor 5 de la gráfica de la izquierda. A menudo, los software escalan al eje vertical como en la gráfica de la derecha. En la figura 4.21, los puntos de la muestra se aproximan a la recta; por tanto, es muy factible que la muestra provenga de una distribución normal.

Es importante mencionar que a los puntos $Q_1, \dots, Q_n$ se les denomina **cuantiles** de la distribución de la que son generados. Algunas veces, a los puntos $X_1, \dots, X_n$ se les llama **cuantiles empíricos**. Por esta razón, algunas veces se hace referencia a la gráfica de probabilidad como gráfica cuantil-cuantil, o gráfica QQ.

En este ejemplo, se utiliza una muestra de sólo cinco puntos para que los cálculos sean claros. En la práctica, las gráficas de probabilidad funcionan mejor con muestras más grandes. Una buena regla general es que se necesiten al menos 30 puntos antes de confiar en una gráfica de probabilidad. No obstante, las gráficas de probabilidad se pueden utilizar en muestras más pequeñas, pero sólo detectarán desviaciones muy pronunciadas de la normalidad.

La figura 4.22 muestra dos gráficas de probabilidad. La gráfica en la figura 4.22a es de las producciones mensuales de 225 pozos de gas. Estos datos no se aproximan a una línea recta; por consiguiente, no provienen de una población que se aproxima a la normal. La gráfica en la figura 4.22b es de los logaritmos naturales de las producciones mensuales. Los datos se aproximan mucho más a una línea recta, aunque se puede ver cierta desviación de la normalidad. (La figura 4.16 de la sección 4.6 presenta los histogramas de estos datos.)

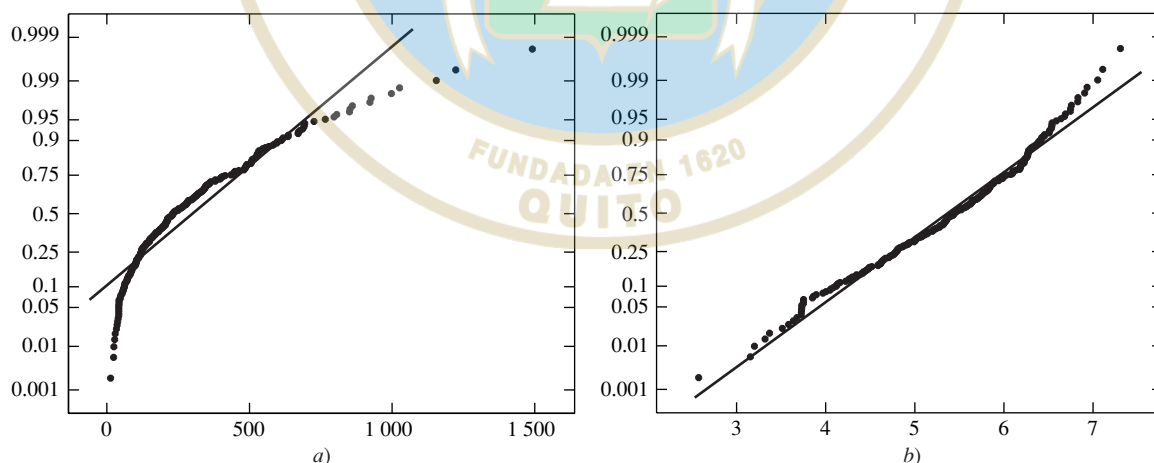


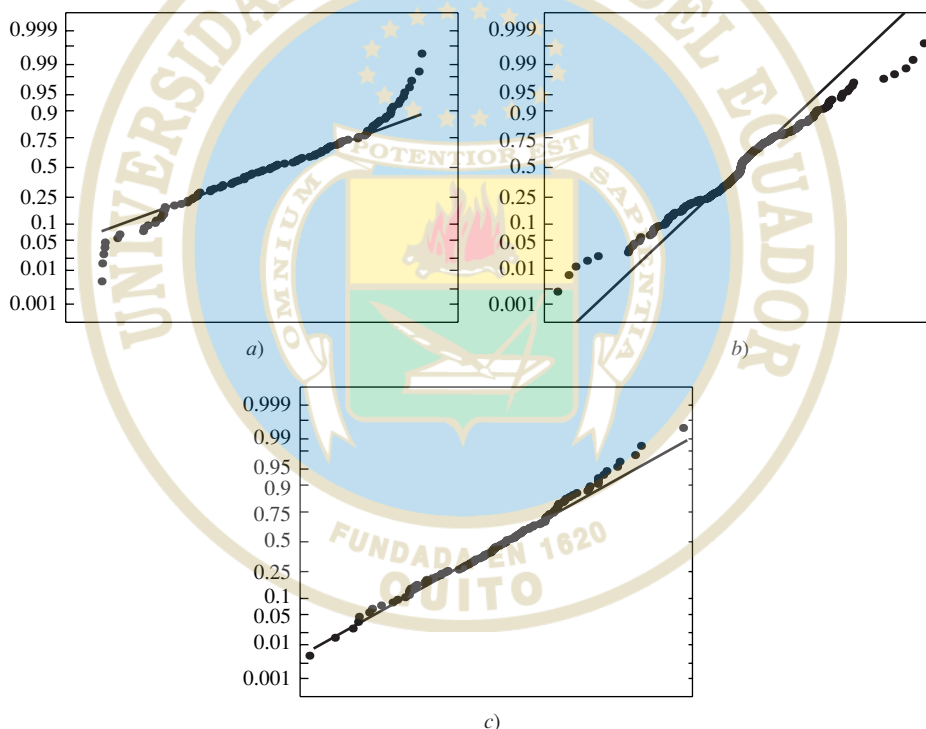
FIGURA 4.22 Dos gráficas de probabilidad normal. (a) Gráfica de producciones mensuales de 225 pozos de gas. Estos datos no se aproximan a una línea recta y, por tanto, no provienen de una población que se aproxima a la normal. (b) Gráfica de los logaritmos naturales de las producciones mensuales. Estos datos se aproximan mucho más a una línea recta, aunque se puede ver cierta desviación de la normalidad. Véase en la figura 4.16 los histogramas de estos datos.

Interpretación de las gráficas de probabilidad

Es mejor no utilizar métodos estrictos e inamovibles en el momento de interpretar una gráfica de probabilidad. La rectitud de la gráfica se puede aproximar a simple vista. Cuando se decide si los puntos de una gráfica de probabilidad se aproximan o no a una línea recta, no preste mucha atención a los puntos en los extremos (superior o inferior) de la muestra, a menos que estén muy alejados de la recta. Es común que algunos puntos en cualesquiera de los extremos se desvíen un poco de la recta. Sin embargo, un punto que se aleja demasiado de la recta, mientras que la mayoría de los otros puntos están cerca, representa un dato atípico y merece atención.

Ejercicios para la sección 4.9

1. Cada una de las tres muestras fueron graficadas sobre una gráfica de probabilidad normal. Para cada una, diga si la muestra parece provenir de una población aproximadamente normal.



2. Construya una gráfica de probabilidad normal para los datos de las barras de jabón del ejercicio 1 en la sección 1.3. ¿Parece ser que estos datos provienen de una distribución aproximadamente normal?

4.1	1.8	3.2	1.9	4.6	2.0	4.5	3.9	4.3	2.3
3.8	1.9	4.6	1.8	4.7	1.8	4.6	1.9	3.5	4.0
3.7	3.7	4.3	3.6	3.8	3.8	3.8	2.5	4.5	4.1
3.7	3.8	3.4	4.0	2.3	4.4	4.1	4.3	3.3	2.0

3. A continuación se muestran las duraciones (en minutos) de 40 erupciones del geiser Old Faithful en el Parque Nacional Yellowstone.

Construya una gráfica de probabilidad normal para estos datos. ¿Parece ser que estos datos provienen de una distribución aproximadamente normal?

4. A continuación se muestran las duraciones (en minutos) de 40 intervalos entre las erupciones del geiser Old Faithful en el Parque Nacional Yellowstone.
- | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|
| 91 | 51 | 79 | 53 | 82 | 51 | 76 | 82 | 84 | 53 |
| 86 | 51 | 85 | 45 | 88 | 51 | 80 | 49 | 82 | 75 |
| 73 | 67 | 68 | 86 | 72 | 75 | 75 | 66 | 84 | 70 |
| 79 | 60 | 86 | 71 | 67 | 81 | 76 | 83 | 76 | 55 |
5. Construya una gráfica de probabilidad normal para los datos MP de la tabla 1.2. ¿Parece ser que las MP vienen de una distribución aproximadamente normal?
6. Construya una gráfica de probabilidad normal para los logaritmos de los datos MP en la tabla 1.2. ¿Parece ser que los logaritmos de los datos MP vienen de una distribución aproximadamente normal?
7. ¿Se puede utilizar la gráfica en el ejercicio 6 para determinar si parece ser que los datos MP vienen de una población lognormal? Explique.
- Construya una gráfica de probabilidad normal para estos datos. ¿Parece ser que estos datos provienen de una distribución aproximadamente normal?

4.10 El teorema del límite central

El **teorema del límite central** es, por mucho, el resultado más importante en estadística. Muchos de los métodos estadísticos comúnmente empleados basan su validez en este teorema. El teorema del límite central establece que si se extrae una muestra lo suficientemente grande de una población, entonces la distribución de la media muestral es aproximadamente normal, sin importar de qué tipo de población haya sido extraída la muestra. Esto último permite calcular las probabilidades de medias muestrales haciendo uso de la tabla z , a pesar de que la población de la cual se sacó la muestra no sea normal. A continuación se explica lo anterior con más detalle.

Sea $X_1, \dots, X_n$ una muestra aleatoria simple de una población con media μ y varianza σ^2 . Sea $\bar{X} = (X_1 + \dots + X_n)/n$ la media muestral. Ahora imagine que se extraen muchas de esas muestras y se calculan sus medias muestrales. Si se pudiera extraer cada muestra posible de tamaño n de la población original, y calcular la media muestral para cada una, la colección resultante sería la población de medias muestrales. Se puede construir la función de densidad de probabilidad de esta población. También es factible observar que la forma de esta función de densidad de probabilidad dependería de la forma de la población de la cual se extrajo la muestra. Lo sorprendente es que si el tamaño de ésta es suficientemente grande esto no es así. Si

El teorema del límite central

Sea $X_1, \dots, X_n$ una muestra aleatoria simple de una población con media μ y varianza σ^2 .

Sea $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ la media muestral.

Sea $S_n = X_1 + \dots + X_n$ la suma de las observaciones muestrales.

Entonces si n es suficientemente grande

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{aproximadamente} \quad (4.49)$$

y

$$S_n \sim N(n\mu, n\sigma^2) \quad \text{aproximadamente} \quad (4.50)$$

el tamaño de la población fuera muy grande, la distribución de la media muestral sería aproximadamente normal, sin importar la distribución de la población de la cual se sacó la muestra.

Observe que el teorema del límite central especifica que $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}}^2 = \sigma^2/n$, lo cual es válido para cualquier media muestral. La suma de las unidades de la muestra es igual a la media multiplicada por el tamaño muestral, esto es, $S_n = n\bar{X}$. De ahí que $\mu_{S_n} = n\mu$ y $\sigma_{S_n}^2 = n^2\sigma^2/n = n\sigma^2$ (véanse las ecuaciones 2.41 y 2.42 de la sección 2.5).

El teorema del límite central establece que $\bar{X}$ y S_n se distribuyen aproximadamente como normales, si el tamaño muestral n es suficientemente grande. La cuestión natural por preguntar es: ¿Qué tan grande es suficientemente grande? La respuesta depende de la forma de la población principal. Si se extrae la muestra de una distribución aproximadamente simétrica, la aproximación normal puede ser buena, incluso para un valor muy pequeño de n . No obstante, si la población está demasiado sesgada, puede ser necesaria una n muy grande. La evidencia empírica sugiere que para la mayoría de las poblaciones, un tamaño de muestra de 30 o más es lo suficientemente grande para que la aproximación normal sea adecuada (véase la figura 4.23).

Para la mayoría de las poblaciones, si el tamaño muestral es mayor a 30, la aproximación del teorema del límite central es buena.

Ejemplo

4.64

Sea X el número de imperfecciones en una pulgada de un alambre de cobre. La función de masa de probabilidad de X se muestra en la siguiente tabla.

x	$P(X = x)$
0	0.48
1	0.39
2	0.12
3	0.01

Se toma una muestra de 100 alambres de esta población. ¿Cuál es la probabilidad de que el número promedio de imperfecciones por alambre en esta muestra sea menor a 0.5?

Solución

La media del número de imperfecciones en la población es $\mu = 0.66$ y la varianza poblacional es $\sigma^2 = 0.5244$. Véase en los ejemplos 2.35 y 2.36 (en la sección 2.4) el cálculo de estas cantidades. $X_1, \dots, X_{100}$ denota el número de imperfecciones en los 100 alambres extraídos de esta población. Se necesita encontrar $P(\bar{X} < 0.5)$. Ahora, el tamaño muestral es $n = 100$, que es una muestra grande. Por el teorema del límite central (expresión 4.49) se tiene que $\bar{X} \sim N(0.66, 0.005244)$. Por tanto, el puntaje z es

$$z = \frac{0.5 - 0.66}{\sqrt{0.005244}} = -2.21$$

De la tabla z , el área a la izquierda de -2.21 es 0.0136. Por consecuencia, $P(\bar{X} < 0.5) = 0.0136$, por lo que sólo 1.36% de las muestras de tamaño 100 tendrá menos de 0.5 imperfecciones por alambre. Véase la figura 4.24.

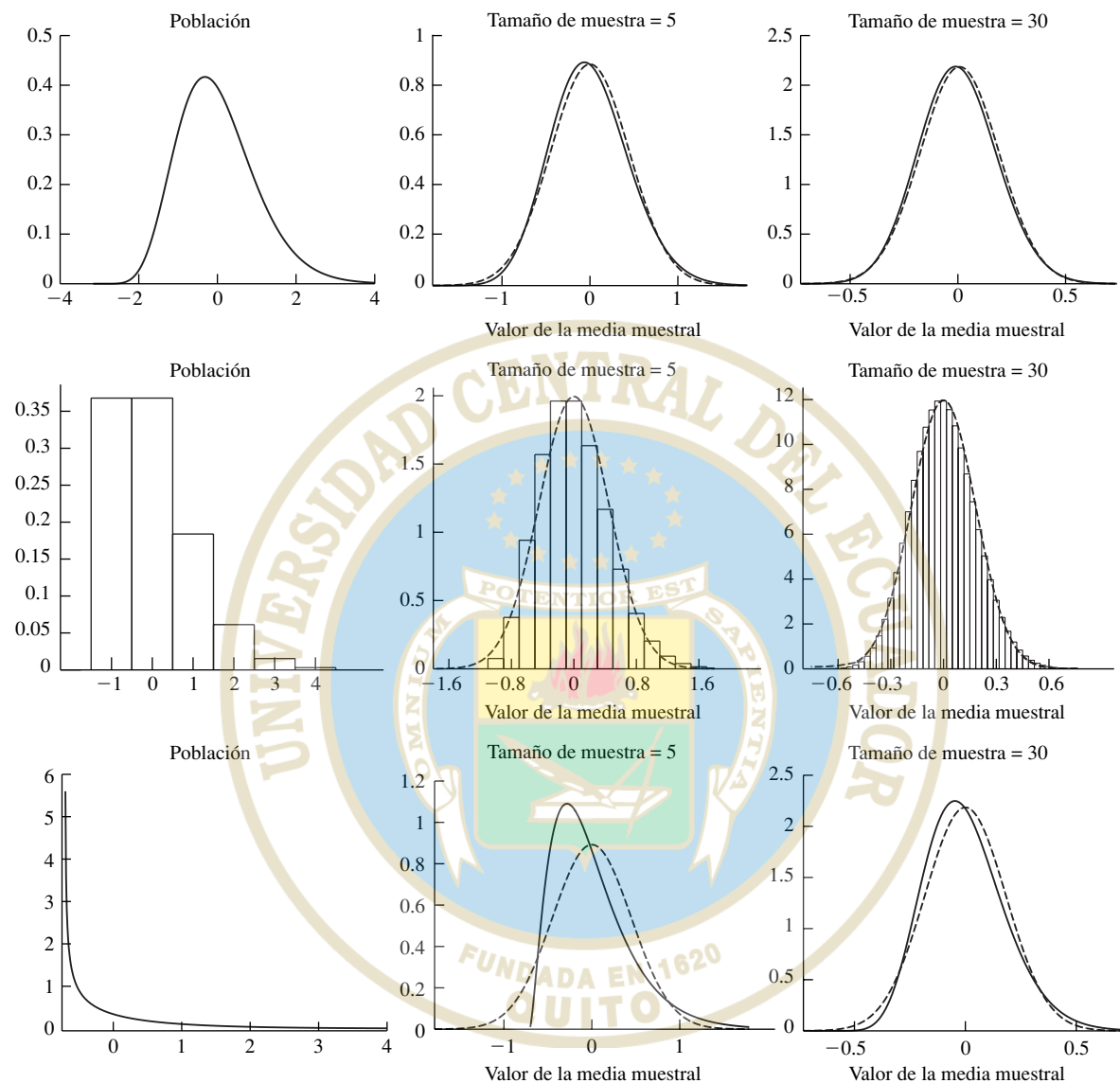


FIGURA 4.23 La gráfica de la extrema izquierda en cada fila es la distribución (función de densidad de probabilidad o función de masa de probabilidad) de una variable aleatoria. Las dos gráficas a la derecha son las distribuciones de la media muestral (línea continua) para muestras de tamaños 5 y 30, respectivamente, con la curva normal (línea discontinua) superpuesta. **Fila superior:** Dado que la distribución original es aproximadamente simétrica, la aproximación normal es buena incluso para un tamaño de muestra tan pequeño como cinco. **Fila de en medio:** La distribución original está algo sesgada. Aun así, la aproximación normal es razonablemente buena incluso para una muestra de tamaño 5, y muy buena para una muestra de tamaño 30. **Fila inferior:** La distribución original está muy sesgada. La aproximación normal no es buena para una muestra de tamaño 5, pero es razonablemente buena para una muestra de tamaño 30. Observe que dos de las distribuciones originales son continuas y una discreta. El teorema del límite central es válido tanto para las distribuciones continuas como para las discretas.

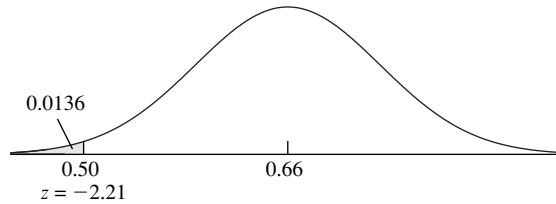


FIGURA 4.24 Solución al ejemplo 4.64.

Observe que en el ejemplo 4.64 se necesita conocer sólo la media y la varianza de la población, no la función de masa de probabilidad.

Ejemplo 4.65

En una universidad grande, la media de la edad de los estudiantes es 22.3 años y la desviación estándar es de cuatro años. Se toma una muestra aleatoria de 64 estudiantes. ¿Cuál es la probabilidad de que la edad promedio de estos estudiantes sea mayor a 23 años?

Solución

Sean $X_1, \dots, X_{64}$ las edades de los 64 estudiantes en la muestra. Se desea determinar $P(\bar{X} > 23)$. Ahora la población de la cual se sacó la muestra tiene una media $\mu = 22.3$ y varianza $\sigma^2 = 16$. El tamaño muestral es $n = 64$. Por el teorema del límite central (expresión 4.49) se tiene que $\bar{X} \sim N(22.3, 0.25)$. El puntaje z para 23 es

$$z = \frac{23 - 22.3}{\sqrt{0.25}} = 1.40$$

De la tabla z , el área a la derecha de 1.40 es 0.0808. Por tanto, $P(\bar{X} > 23) = 0.0808$. Véase la figura 4.25.

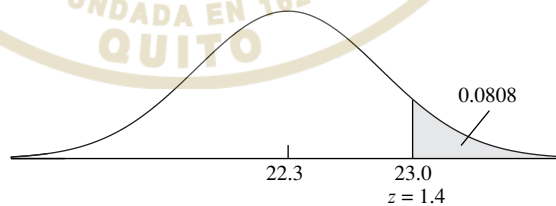


FIGURA 4.25 Solución al ejemplo 4.65.

Aproximación normal a la binomial

Recuerde de la sección 4.2, que si $X \sim \text{Bin}(n, p)$, entonces $X = Y_1 + \dots + Y_n$, donde $Y_1, \dots, Y_n$ es una muestra de una población de Bernoulli(p). Por consiguiente, X es la suma de los datos de la muestra. La proporción muestral es

$$\hat{p} = \frac{X}{n} = \frac{Y_1 + \cdots + Y_n}{n}$$

que es también la media muestral $\overline{Y}$. La población de Bernoulli(p) tiene una media $\mu = p$ y una varianza $\sigma^2 = p(1 - p)$. Por el teorema del límite central se tiene que si el número de experimentos n es grande, entonces $X \sim N(np, np(1 - p))$ y $\hat{p} \sim N(p, p(1 - p)/n)$.

Nuevamente surge la cuestión, ¿qué tan grande es lo suficientemente grande? En el caso binomial, la exactitud de la aproximación normal depende del número promedio de éxitos np y del número promedio de fracasos $n(1 - p)$. Entre más grandes sean los valores de np y $n(1 - p)$, mejor será la aproximación. Una regla general es utilizar la aproximación normal cada vez que $np > 5$ y $n(1 - p) > 5$. Una regla mejor y más prudente es utilizar la aproximación normal cada vez que $np > 10$ y $n(1 - p) > 10$.

Resumen			
Si $X \sim \text{Bin}(n, p)$ y si $np > 10$ y $n(1 - p) > 10$, entonces			
$X \sim N(np, np(1 - p))$	aproximadamente	(4.51)	
$\hat{p} \sim N\left(p, \frac{p(1 - p)}{n}\right)$	aproximadamente	(4.52)	

Para mostrar la exactitud de la aproximación normal a la binomial, la figura 4.26 muestra el histograma de probabilidad $\text{Bin}(100, 0.2)$ con la función de densidad de probabilidad $N(20, 16)$ sobrepuesta. Aunque se puede detectar cierto grado de asimetría en la distribución binomial, la aproximación normal es muy buena.

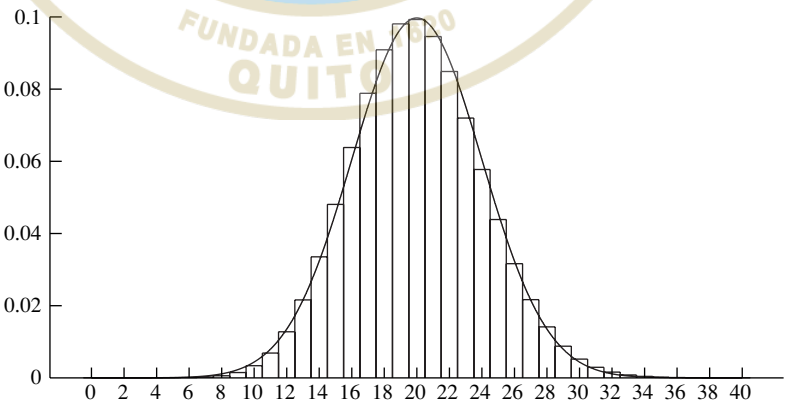


FIGURA 4.26 El histograma de probabilidad $\text{Bin}(100, 0.2)$, con la función de densidad de probabilidad $N(20, 16)$ sobrepuesta.

Corrección por continuidad

La distribución binomial es discreta, mientras que la distribución normal es continua. La **corrección por continuidad** es un ajuste, hecho cuando se aproxima una distribución discreta con una continua, que puede mejorar la exactitud de la aproximación. Para ver cómo funciona, imagine que se lanza al aire una moneda 100 veces. Sea X el número de caras. Entonces $X \sim \text{Bin}(100, 0.5)$. Imagine que se desea calcular la probabilidad de que X esté entre 45 y 55. Esta probabilidad diferirá dependiendo si se incluye o excluye los puntos finales, 45 y 55. La figura 4.27 ilustra el caso en el que se incluye los puntos finales; es decir, en el que se quiere estimar $P(45 \leq X \leq 55)$. La probabilidad exacta está dada por el área total de los rectángulos del histograma de probabilidad binomial correspondientes, incluyendo a los enteros 45 y 55. Se sobrepone la curva normal aproximada. Para obtener una mejor aproximación se debe calcular el área bajo la curva normal entre 44.5 y 55.5. En contraste, la figura 4.28 muestra el caso en el que se desea calcular $P(45 < X < 55)$. Aquí se incluye los puntos finales. La probabilidad exacta está dada por el área total de los rectángulos del histograma de probabilidad binomial correspondientes a los enteros 46 al 54. La mejor aproximación normal se encuentra al calcular el área bajo la curva normal entre 45.5 y 54.5.

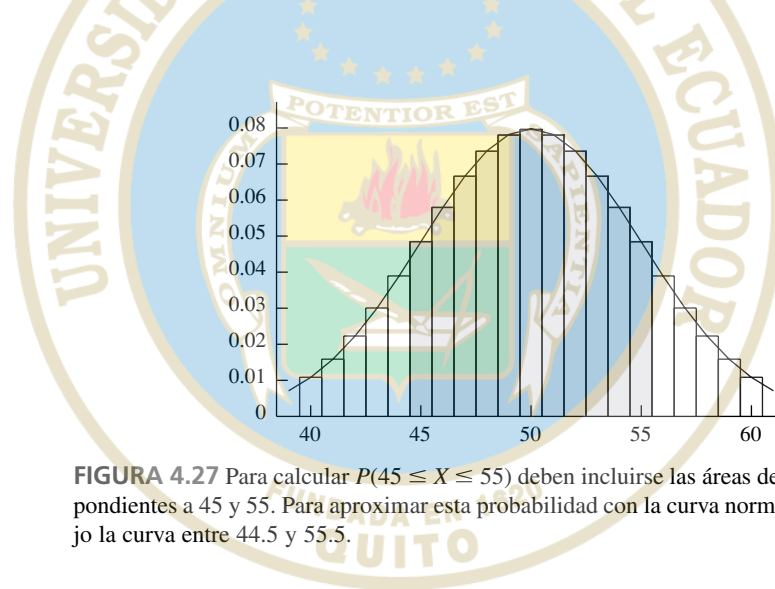


FIGURA 4.27 Para calcular $P(45 \leq X \leq 55)$ deben incluirse las áreas de los rectángulos correspondientes a 45 y 55. Para aproximar esta probabilidad con la curva normal, se calcula el área bajo la curva entre 44.5 y 55.5.

En resumen, para aplicar la corrección por continuidad, determine qué rectángulos del histograma de probabilidad discreta desea incluir, y después calcule el área bajo la curva normal que corresponde a estos rectángulos.

Ejemplo 4.66

Si se lanza al aire una moneda 100 veces, utilice la curva normal para aproximar la probabilidad de que el número de caras esté entre, e *incluyendo* a, 45 y 55.

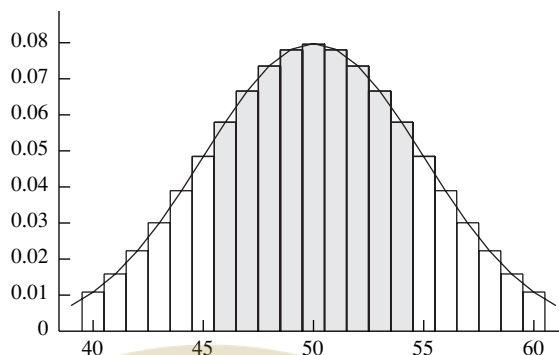


FIGURA 4.28 Para calcular $P(45 < X < 55)$ se debe excluir las áreas de los rectángulos correspondientes a 45 y a 55. Para aproximar esta probabilidad con la curva normal, se calcula el área bajo la curva entre 45.5 y 54.5.

Solución

La figura 4.27 muestra esta situación. Sea X el número de caras obtenidas. Entonces $X \sim \text{Bin}(100, 0.5)$. Al sustituir $n = 100$ y $p = 0.5$ en la ecuación (4.51) se obtiene la aproximación normal $X \sim N(50, 25)$. Puesto que se incluyen los puntos finales 45 y 55, debe calcularse el área bajo la curva normal entre 44.5 y 55.5. Los puntajes z para 44.5 y 55.5 son

$$z = \frac{44.5 - 50}{5} = -1.1, \quad z = \frac{55.5 - 50}{5} = 1.1$$

De la tabla z se determina que la probabilidad es 0.7286. Véase la figura 4.29.

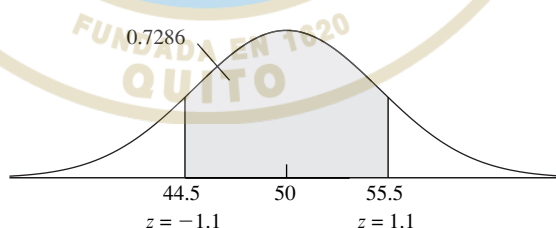


FIGURA 4.29 Solución al ejemplo 4.66.

Ejemplo 4.67

Si se lanza al aire 100 veces una moneda, utilice la curva normal para aproximar la probabilidad de que el número de caras está entre, y *excluyendo* a, 45 y 55.

Solución

La figura 4.28 muestra esta situación. Sea X el número de “caras” obtenidas. Al igual que en el ejemplo 4.66, $X \sim \text{Bin}(100, 0.5)$, y la aproximación normal es $X \sim N(50, 25)$. Como consecuencia de que se excluyen los puntos finales 45 y 55, se debe calcular el área bajo la curva normal entre 45.5 y 54.5. Los puntajes z para 45.5 y 54.5 son

$$z = \frac{45.5 - 50}{5} = -0.9, \quad z = \frac{54.5 - 50}{5} = 0.9$$

De la tabla z se determina que la probabilidad es 0.6318. Véase la figura 4.30.

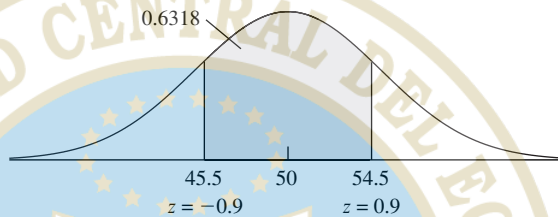


FIGURA 4.30 Solución al ejemplo 4.67.

Ejemplo

4.68

En cierta universidad grande, 25% de los estudiantes tiene más de 21 años de edad. En una muestra de 400 estudiantes, ¿cuál es la probabilidad de que más de 110 supere los 21 años?

Solución

Sea X el número de estudiantes que tiene más de 21 años. Entonces $X \sim \text{Bin}(400, 0.25)$. Puede utilizarse la aproximación normal, la cual es $X \sim N(100, 75)$. Ya que se desea determinar la probabilidad de que el número de estudiantes sea *mayor que* 110, se excluye al valor 110. Por tanto, se determina $P(X > 110.5)$. Se calcula el puntaje z para 110.5, el cual es

$$z = \frac{110.5 - 100}{\sqrt{75}} = 1.21$$

Utilizando la tabla z se determina que $P(X > 110.5) = 0.1131$. Véase la figura 4.31.

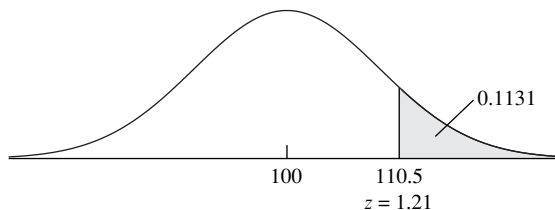


FIGURA 4.31 Solución al ejemplo 4.68.

Exactitud de la corrección por continuidad

La corrección por continuidad mejora, en la mayoría de los casos, la exactitud de la aproximación normal a la distribución binomial. Sin embargo, para distribuciones binomiales con n grande y p pequeña, cuando se calcula una probabilidad que corresponda a un área en la cola de la distribución, la corrección por continuidad puede, en algunos casos, reducir en algo la exactitud de la aproximación normal. Esto resulta del hecho de que la aproximación normal no es perfecta; no puede explicar un pequeño grado de asimetría en estas distribuciones. En suma, el uso de la corrección por continuidad hace que la aproximación normal a la distribución binomial sea mejor en la mayoría de los casos, pero no en todos.

Aproximación normal a la de Poisson

Recuerde que si $X \sim \text{Poisson}(\lambda)$, entonces X es aproximadamente binomial con n grande y $np = \lambda$. Recuerde también que $\mu_X = \lambda$ y $\sigma_X^2 = \lambda$. De ahí que si λ es suficientemente grande, es decir, $\lambda > 10$, entonces X es aproximadamente binomial, con $np > 10$. Por el teorema del límite central se tiene que X es también aproximadamente normal, con media y varianza igual a λ . Por tanto, puede utilizarse la distribución normal para aproximar a la de Poisson.

Resumen			
Si $X \sim \text{Poisson}(\lambda)$, donde $\lambda > 10$, entonces			
$X \sim N(\lambda, \lambda)$,		aproximadamente	(4.53)

Corrección por continuidad para la distribución de Poisson

Dado que una distribución de Poisson es discreta, la corrección por continuidad puede, en principio, aplicarse cuando se utiliza la aproximación normal. Para las áreas que incluyen la parte central de la curva, la corrección por continuidad generalmente mejora la aproximación normal, pero para las áreas de las colas la corrección por continuidad algunas veces empeora la aproximación. No se utilizará la corrección por continuidad para la distribución de Poisson.

Ejemplo
4.69

El número de visitas a un sitio web sigue una distribución de Poisson, con una media de 27 visitas por hora. Encuentre la probabilidad de que haya 90 o más visitas durante tres horas.

Solución

Sea X el número de visitas al sitio web en tres horas. La media del número de visitas en tres horas es 81, por lo que $X \sim \text{Poisson}(81)$. Utilizando la aproximación normal $X \sim N(81, 81)$. Se desea encontrar $P(X \geq 90)$. Se calcula el puntaje z de 90, que es

$$z = \frac{90 - 81}{\sqrt{81}} = 1.00$$

Haciendo uso de la tabla z se determina que $P(X \geq 90) = 0.1587$. Véase la figura 4.32.

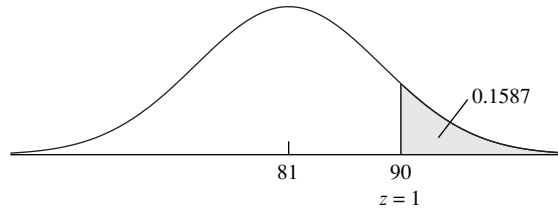


FIGURA 4.32 Solución al ejemplo 4.69.

Ejercicios para la sección 4.10

- Las bolsas de una mezcla de concreto, etiquetadas con un peso de 100 libras, tienen por peso una media poblacional de 100 libras y una desviación estándar poblacional de 0.5 libras.
 - ¿Cuál es la probabilidad de que la media del peso de una muestra aleatoria de 50 bolsas sea menor a 99.9 libras?
 - Si la media del peso de poblacional aumenta a 100.15 libras, ¿cuál es la probabilidad de que la media del peso de una muestra de tamaño 50 sea menor a 100 libras?
- Un libro de 500 páginas tiene 250 hojas de papel. El espesor del papel utilizado para fabricar el libro tiene una media de 0.08 mm y una desviación estándar de 0.01 mm.
 - ¿Cuál es la probabilidad de que un libro seleccionado aleatoriamente tenga un espesor mayor a 20.2 mm (sin incluir las portadas)?
 - ¿Cuál es el 10o. percentil del espesor del libro?
 - Alguien quiere conocer la probabilidad de que una página elegida aleatoriamente tenga un espesor mayor a 0.1 mm. ¿Se tiene la suficiente información para calcular esta probabilidad? Si es así, calcule la probabilidad. Si no, explique por qué.
- Se elige una muestra aleatoria de 100 hombres con una estatura media de 70 pulgadas y una desviación estándar de 2.5 pulgadas. ¿Cuál es la probabilidad de que la estatura promedio de los hombres de la muestra sea mayor a 69.5 pulgadas?
- Entre todas las formas de declaración de ingresos para el pago de impuestos, llenadas en cierto año, la media del impuesto pagado fue de \$2 000 y la desviación estándar fue de \$500. Además, en 10% de las formas, el impuesto pagado fue mayor a \$3 000. Se toma una muestra aleatoria de 625 formas de declaración de ingresos para el pago de impuestos.
 - ¿Cuál es la probabilidad de que el impuesto promedio pagado en las formas de la muestra sea mayor a \$1 980?
 - ¿Cuál es la probabilidad de que más de 60 de las formas de la muestra tenga un impuesto mayor a \$3 000?
- Se extrae una muestra de 225 alambres de la población de alambres descrita en el ejemplo 4.64 (p. 271). Encuentre la probabilidad de que menos de 110 de éstos no tengan imperfección.
- Unos tambores, con una etiqueta de 30 L, son llenados con una solución proveniente de una tina grande. Se agrega una cantidad aleatoriamente de la solución en cada tambor con media de 30.01 L y desviación estándar de 0.1 L.
 - ¿Cuál es la probabilidad de que la cantidad total de la solución contenida en 50 tambores sea mayor a 1 500 L?
 - Si la cantidad total de la solución en la tina es de 2 401 L, ¿cuál es la probabilidad de que puedan llenarse 80 tambores sin que se acabe la solución?
 - ¿Cuánta solución debe contener la tina para que la probabilidad sea 0.9 de que puedan llenarse 80 tambores sin que se acabe la solución?
- Cierto proceso de fabricación de componentes electrónicos produce partes, 20% de las cuales está defectuoso. Las partes son enviadas en unidades de 400. Los envíos que contienen más de 90 partes defectuosas se puede regresar. Usted puede suponer que cada envío constituye una muestra aleatoria simple de partes.
 - ¿Cuál es la probabilidad de que se regrese un envío específico?
 - En un día particular se realizaron 500 envíos. ¿Cuál es la probabilidad de que se regresen 60 o más de éstos?

- c) Se introduce un nuevo proceso de fabricación, el cual se supone reduce los porcentajes de las partes defectuosas. El objetivo de la compañía es reducir la probabilidad de que se regrese un envío a 0.01. ¿Cuál debe ser el porcentaje de partes defectuosas para alcanzar este objetivo?
8. Verdadero o falso:
- Con una muestra grande, el histograma de la muestra se parecerá mucho a la curva normal.
 - Con una muestra grande, la función de densidad de probabilidad de la media muestral se parecerá mucho a la curva normal.
9. La densidad de las partículas en una suspensión es de 50 por mL. Se extrae un volumen de 5 mL de la suspensión.
- ¿Cuál es la probabilidad de que el número de partículas extraídas esté entre 235 y 265?
 - ¿Cuál es la probabilidad de que el número promedio de partículas por mL en la muestra extraída esté entre 48 y 52?
 - Si se toma una muestra de 10 mL, ¿cuál es la probabilidad de que el número promedio por mL de partículas en la muestra extraída esté entre 48 y 52?
 - ¿Qué tan grande debe ser la muestra extraída para que el número promedio de partículas por mL en la muestra esté entre 48 y 52 con probabilidad de 95%?
10. Un productor de baterías afirma que la duración de cierto tipo de batería tiene una media poblacional de 40 horas y desviación estándar de cinco horas. Sea $\bar{X}$ la duración promedio de las baterías en una muestra aleatoria simple de tamaño 100.
- Si la afirmación es cierta, ¿cuál es $P(\bar{X} \leq 36.7)$?
 - Con base en la respuesta al inciso a), si la afirmación es cierta, ¿una media muestral de 36.7 horas sería una duración inusualmente corta?
 - Si la duración media de la muestra de las 100 baterías fuera de 36.7 horas, ¿usted creería en la afirmación del fabricante? Explique.
 - Si la afirmación es cierta, ¿cuál es $P(\bar{X} \leq 39.8)$?
 - Con base en la respuesta al inciso d), si la afirmación es cierta, ¿una media muestral de 39.8 horas sería una duración inusualmente corta?
 - Si la media muestral de la duración de las 100 baterías fuera de 39.8 horas, ¿usted creería en la afirmación del fabricante? Explique.
11. Se ha diseñado un nuevo proceso para fabricar lozas de cerámica. El objetivo es que no haya más de 5% que no sea satisfactorio debido a defectos en la superficie. Se revisa una muestra de 1 000 lozas. Sea X el número de lozas no satisfactorio en la muestra.
- Si 5% de las lozas producidas no es satisfactorio, ¿a qué es igual $P(X \geq 75)$?
 - Con base en la respuesta al inciso a), si 5% de las lozas no es satisfactorio, ¿75 no satisfactorias de las 1 000 sería un número inusualmente grande? Explique.
 - Si 75 de las lozas de la muestra no fueran satisfactorias, ¿sería factible que se haya alcanzado el objetivo? Explique.
 - Si 5% de las lozas producidas no es satisfactorio, ¿a qué es igual $P(X \geq 53)$?
 - Con base en la respuesta al inciso d), si 5% de las lozas no es satisfactorio, ¿53 lozas no satisfactorias de las 1 000 sería un número inusualmente grande?
 - Si 53 de las lozas de la muestra no fueran satisfactorias, ¿sería factible que se haya alcanzado el objetivo? Explique.
12. *Fechado radiactivo:* El carbono-14 es un isótopo radiactivo del carbono que decae al emitir una partícula beta. En la atmósfera terrestre, aproximadamente un átomo de carbono en 10^{12} es carbono-14. Los organismos vivos intercambian carbono con la atmósfera, por lo que esta misma tasa es válida para el tejido vivo. Después de que un organismo muere, éste deja de intercambiar carbono con su ambiente, y su tasa de carbono-14 disminuye exponencialmente con el tiempo. La tasa a la cual se emite partículas beta desde una masa dada de carbono es proporcional a la tasa de carbono-14, por lo que esta tasa también disminuye con el tiempo. Al medir la tasa de emisiones beta en una muestra de tejido, se puede estimar el tiempo transcurrido desde la muerte del organismo. Específicamente, se sabe que t años después de la muerte, el número de emisiones de partículas beta que ocurre en un intervalo a partir de 1 g de carbono sigue una distribución de Poisson con tasa $\lambda = 15.3e^{-0.0001210t}$ eventos por minuto. Por tanto, el número de años t transcurridos desde la muerte de un organismo se puede expresar en términos de λ :
- $$t = \frac{\ln 15.3 - \ln \lambda}{0.0001210}$$
- Un arqueólogo descubre una pequeña pieza de carbón vegetal proveniente de un campamento antiguo. El carbón vegetal contiene 1 g de carbono.

- a) El carbón vegetal tiene 11 000 años de antigüedad, lo que el arqueólogo no sabe. ¿Cuál es el valor verdadero de la tasa de emisión λ ?
- b) El arqueólogo planea contar el número X de emisiones en un intervalo de 25 minutos. Encuentre la media y la desviación estándar de X .
- c) Posteriormente, el arqueólogo planea estimar λ con $\hat{\lambda} = X/25$. ¿Cuáles son la media y la desviación estándar de $\hat{\lambda}$?
- d) ¿Qué valor de $\hat{\lambda}$ daría como resultado una estimación de antigüedad de 10 000 años?
- e) ¿Qué valor de $\hat{\lambda}$ daría como resultado una estimación de antigüedad de 12 000 años?
- f) ¿Cuál es la probabilidad de que la estimación de antigüedad tenga una corrección de $\pm 1\,000$ años?

4.11 Simulación

Cuando nacen gemelos-cuates (no idénticos), ambos pueden ser niños, niñas, o uno de cada uno. Suponga que cada infante tiene la misma probabilidad de ser un niño que una niña, y suponga que los sexos de los gemelos se determinan de manera independiente. ¿Cuál es la probabilidad de que ambos sean niños? Esta probabilidad se estima fácilmente, empleando la regla de la multiplicación de eventos independientes. La respuesta es $(0.5)(0.5) = 0.25$. Pero se supone que no se conocía la regla de la multiplicación. ¿Existe otra forma en la que se pueda estimar esta probabilidad? Se puede realizar un experimento, o estudio, científico. Es posible obtener los registros de nacimientos de gemelos en hospitales, y contar el número en los cuales ambos son niños. Si se obtuviese un número suficientemente grande de registros, la proporción en la cual ambos gemelos fueron niños quizá se aproximaría a 0.25, y se tendría una buena estimación de la probabilidad.

A continuación se presenta un método más sencillo. Hay dos resultados igualmente probables en el nacimiento de un gemelo: niño o niña. También hay dos resultados probables en el lanzamiento al aire de una moneda: “cara” o “cruz”. Por consecuencia, el número de “caras” en el lanzamiento de dos monedas tiene la misma distribución que el número de niños en el nacimiento de un gemelo (ambas son binomiales con $n = 2$ experimentos y una probabilidad de éxito $p = 0.5$). En lugar de tener el problema de dar seguimiento a los nacimientos reales, se puede lanzar al aire dos monedas un gran número de veces. La proporción de lanzamientos en los que en ambas monedas sale “cara” se puede utilizar para estimar la proporción de nacimientos en los que ambos gemelos son niños.

Estimar la probabilidad de que ambos gemelos sean niños al usar la estimación de la probabilidad de que en ambas monedas salga “cara” es un ejemplo de un experimento de **simulación**. Si se designa a los lados de la moneda como “0” y “1”, entonces el lanzamiento al aire de una moneda es un ejemplo de un **generador de números aleatorios**. Este último constituye un procedimiento para obtener un valor que tiene las mismas propiedades estadísticas como una cantidad muestral aleatoria extraída de cierta distribución específica. El número aleatorio generado por el lanzamiento al aire de una moneda viene de una distribución de Bernoulli con probabilidad de éxito $p = 0.5$.

Hoy en día, las computadoras pueden generar miles de números aleatorios en una fracción de segundo, y virtualmente cada software estadístico contiene rutinas que generan muestras aleatorias a partir de una amplia variedad de distribuciones. Cuando un experimento es demasiado costoso, o físicamente difícil o imposible de realizar, y cuando la distribución de la probabilidad de los datos que serían generados por el experimento es aproximadamente co-

nocido, los números aleatorios generados por computadora a partir de la distribución apropiada se puede utilizar en lugar de los verdaderos datos del experimento. A dichos números se les llama datos **simulados** o **sintéticos**.

Resumen

La **simulación** se refiere al proceso que genera números aleatorios y los trata como si hubiesen sido generados por un experimento científico real. A los datos generados de esta forma se les denomina **simulados** o **sintéticos**.

Los métodos de simulación tienen muchos usos, incluyendo la estimación de probabilidades, estimación de medias y varianzas, comprobación de una hipótesis de normalidad, y estimación de sesgos. En el resto de esta sección se describen algunos de estos métodos.

Uso de la simulación para estimar una probabilidad

Con frecuencia se emplea la simulación para estimar probabilidades que son difíciles de calcular directamente. Aquí hay un ejemplo. Un ingeniero eléctrico conectará en paralelo dos resistores, etiquetados como 100 y 25 Ω . Las resistencias verdaderas pueden diferir de los valores etiquetados. Se denota las resistencias reales de los resistores que son elegidos por X y Y . La resistencia total R del montaje está dada por $R = XY/(X + Y)$. Suponga que $X \sim N(100, 10^2)$ y $Y \sim N(25, 2.5^2)$ y que los resistores se seleccionan en forma independiente. Suponga que la especificación para la resistencia del montaje es que $19 < R < 21$. ¿Cuál es la probabilidad de que el montaje cumpla con la especificación? Es decir, ¿a qué es igual $P(19 < R < 21)$?

Se estimará esta probabilidad con una simulación. La idea es generar datos simulados cuya distribución se aproxime en lo posible a los datos que serían generados en un experimento real. En un experimento real, se tomaría una muestra de N resistores etiquetados con 100 Ω , cuyas resistencias verdaderas fuesen $X_1, \dots, X_N$, y después se tomaría una muestra de igual tamaño de resistores etiquetados con 25 Ω , cuyas resistencias verdaderas fuesen $Y_1, \dots, Y_N$. Después se construiría N montajes con resistencias $R_1 = X_1 Y_1 / (X_1 + Y_1), \dots, R_N = X_N Y_N / (X_N + Y_N)$. Los valores $R_1, \dots, R_N$ serían una muestra aleatoria de la población de todos los valores posibles de la resistencia total. La proporción de los valores de la muestra $R_1, \dots, R_N$ que estén entre 19 y 21 serían una estimación de $P(19 < R < 21)$.

En un experimento real, $X_1, \dots, X_N$ sería una muestra aleatoria de $N(100, 10^2)$ y $Y_1, \dots, Y_N$ sería una muestra aleatoria de $N(25, 2.5^2)$. Por consiguiente, en un experimento simulado, se generaría una muestra aleatoria $X_1^*, \dots, X_N^*$ de $N(100, 10^2)$ e, independientemente, una muestra aleatoria $Y_1^*, \dots, Y_N^*$ de $N(25, 2.5^2)$. Luego se calcularán las resistencias totales simuladas $R_1^* = X_1^* Y_1^* / (X_1^* + Y_1^*), \dots, R_N^* = X_N^* Y_N^* / (X_N^* + Y_N^*)$. Se emplea la notación X_i^*, Y_i^* y R_i^* para indicar que estos son valores simulados provenientes de un generador de números aleatorios en vez de datos verdaderos provenientes de un experimento real. Dado que la muestra $X_1^*, \dots, X_N^*$ proviene de la misma distribución que la de una muestra real $X_1, \dots, X_N$, y puesto que la muestra $Y_1^*, \dots, Y_N^*$ proviene de la misma distribución que la de una muestra real $Y_1, \dots, Y_N$, se tiene que la muestra $R_1^*, \dots, R_N^*$ proviene de la misma distribución que la de una muestra real de resistencias totales $R_1, \dots, R_N$. Por tanto, se puede tratar a $R_1^*, \dots, R_N^*$

como si fuera una muestra de resistencias verdaderas, aunque en realidad es una muestra de números aleatorios generados por una computadora.

Los resultados de una simulación con un tamaño de muestra $N = 100$ se muestran en la tabla 4.2. Ésta es la muestra más pequeña que se utilizaría en la práctica. Por lo general, se utilizan con frecuencia muestras de 1 000, 10 000, o más. Las muestras de este tamaño no representan ningún problema para las computadoras modernas y sus programas, y entre más grande sea la muestra, más precisos serán los resultados.

Con el propósito de que los cálculos sean más transparentes, se ordenan los 100 valores de R_i^* , que están en la tabla 4.2, en orden ascendente.

15.37	15.48	15.58	16.66	16.94	17.18	17.44	17.54	17.68	17.69
17.91	17.95	18.01	18.06	18.21	18.31	18.49	18.58	18.60	18.65
18.71	18.80	18.81	18.85	18.91	18.92	18.93	18.99	18.99	19.01
19.02	19.03	19.06	19.11	19.13	19.14	19.20	19.22	19.24	19.30
19.47	19.52	19.56	19.58	19.60	19.60	19.65	19.71	19.77	19.81
19.84	19.90	19.91	19.95	19.97	19.98	20.03	20.14	20.16	20.17
20.17	20.49	20.52	20.54	20.55	20.55	20.58	20.60	20.60	20.64
20.69	20.75	20.76	20.78	20.81	20.90	20.96	21.06	21.13	21.24
21.41	21.49	21.52	21.54	21.58	21.79	21.84	21.87	21.93	21.93
22.02	22.06	22.11	22.13	22.36	22.42	23.19	23.40	23.71	24.01

Para estimar la $P(19 < R < 21)$ se determina que 48 de los 100 valores de la muestra se encuentran en este rango. Por consiguiente, se estima $P(19 < R < 21) = 0.48$. Se observa que con una muestra más grande, se puede emplear un software para hacer este conteo.

Advierta que la importancia del supuesto de que la resistencia X del primer resistor y la Y del segundo fuesen independientes. Debido a este supuesto, es posible simular el experimento al generar muestras independientes X^* y Y^* . Si X y Y hubieran sido dependientes, se tendría que haber generado X^* y Y^* para que tuvieran la misma *distribución conjunta* que X y Y . (En la sección 2.6 se analizaron las distribuciones conjuntas.) Por fortuna, muchos problemas reales de la vida implican muestras independientes.

A continuación se presenta otro ejemplo de una probabilidad estimada con una simulación.

Ejemplo

4.70

Una ingeniera tiene que elegir entre dos tipos de ventiladores para instalarlo en una computadora. Las duraciones, en meses, de los ventiladores del tipo A se distribuyen exponencialmente con una media de 50 meses, y las de los tipo B se distribuyen exponencialmente con una media de 30 meses. Como consecuencia de que los ventiladores del tipo A son más caros, la ingeniera decide que elegirá estos últimos si la probabilidad de que este tipo duplique su tiempo de duración que el de un ventilador B sea mayor a 0.5. Estime esta probabilidad.

Solución

Sea A la duración, en meses, de ventilador de tipo A elegido de forma aleatoria, y B represente la duración, en meses, de otro elegido aleatoriamente. Se necesita calcular a $P(A > 2B)$. Se realiza un experimento de simulación, utilizando muestras de tamaño 1 000. Se genera una muestra aleatoria $A_1^*, \dots, A_{1\,000}^*$ a partir de una distribución exponencial con media 50 ($\lambda = 0.02$) y una muestra aleatoria $B_1^*, \dots, B_{1\,000}^*$ a partir de una distribución exponencial con media 30 ($\lambda = 0.033$). Luego se cuenta el número de veces que $A_i^* > 2B_i^*$. La tabla 4.3 presenta los primeros diez valores y el último valor. La columna etiquetada como " $A^* > 2B^*$ " contiene un "1" si $A_i^* > 2B_i^*$, y un "0" si $A_i^* \leq 2B_i^*$.

TABLA 4.2 Datos simulados de las resistencias en un circuito paralelo

	R*					Y*					X*					R*					Y*					X*					R*					Y*					X*					R*					Y*					X*																																																																																																																																																																																																																																																																																																																																																							
1	102.63	24.30	19.65	26	115.94	24.93	20.52	51	94.20	23.68	18.92	76	113.32	22.54	18.80	2	96.83	21.42	17.54	27	100.65	28.36	22.13	52	82.62	27.82	20.81	77	90.82	23.79	18.85	3	96.46	26.34	20.69	28	89.71	23.00	18.31	53	119.49	22.88	19.20	78	102.88	25.99	20.75	4	88.39	22.10	17.68	29	104.93	24.10	19.60	54	99.43	28.03	21.87	79	93.59	23.04	18.49	5	113.07	29.17	23.19	30	93.74	23.68	18.91	55	108.03	21.69	18.06	80	89.19	27.05	20.76	6	117.66	27.09	22.02	31	104.20	24.02	19.52	56	95.32	20.60	16.94	81	95.04	23.76	19.01	7	108.04	18.20	15.58	32	123.43	26.66	21.93	57	80.70	30.36	22.06	82	109.72	30.25	23.71	8	101.13	28.30	22.11	33	101.38	22.19	18.21	58	91.13	20.38	16.66	83	107.35	27.01	21.58	9	105.43	23.51	19.22	34	88.52	26.85	20.60	59	111.35	27.09	21.79	84	89.59	18.55	15.37	10	103.41	24.64	19.90	35	101.23	26.88	21.24	60	118.75	23.92	19.91	85	101.72	24.65	19.84	11	104.89	22.59	18.58	36	86.96	25.66	19.81	61	103.33	23.99	19.47	86	101.24	25.92	20.64	12	94.91	27.86	21.54	37	95.92	26.16	20.55	62	107.77	18.08	15.48	87	109.67	26.61	21.41	13	92.91	27.06	20.96	38	95.97	26.05	20.49	63	104.86	24.64	19.95	88	100.74	26.18	20.78	14	95.86	24.82	19.71	39	93.76	24.71	19.56	64	84.39	25.52	19.60	89	98.44	23.63	19.06	15	100.06	23.65	19.13	40	113.89	22.79	18.99	65	94.26	25.61	20.14	90	101.05	28.81	22.42	16	90.34	23.75	18.81	41	109.37	26.19	21.13	66	82.16	27.49	20.60	91	88.13	28.43	21.49	17	116.74	24.38	20.17	42	91.13	24.93	19.58	67	108.37	27.35	21.84	92	113.94	29.45	23.40	18	90.45	25.30	19.77	43	101.60	28.66	22.36	68	86.16	21.46	17.18	93	97.42	23.78	19.11	19	97.58	23.05	18.65	44	102.69	21.37	17.69	69	105.97	23.59	19.30	94	109.05	23.04	19.02	20	101.19	23.60	19.14	45	108.50	25.34	20.54	70	92.69	23.88	18.99	95	100.65	26.63	21.06	21	101.77	31.42	24.01	46	80.86	27.55	20.55	71	97.48	25.43	20.17	96	105.64	21.57	17.91	22	100.53	24.93	19.98	47	83.80	24.80	19.24	72	110.45	20.70	17.44	97	78.82	23.25	17.95	23	98.00	27.57	21.52	48	105.96	23.20	19.03	73	89.92	27.23	20.90	98	112.31	22.77	18.93	24	108.10	27.51	21.93	49	103.98	21.78	18.01	74	103.78	25.67	20.58	99	100.14	24.95	19.97	25	91.07	23.38	18.60	50	97.97	23.13	18.71	75	95.53	25.55	20.16	100	88.78	25.87	20.03

TABLA 4.3 Datos simulados para el ejemplo 4.70

	A^*	B^*	$A^* > 2B^*$
1	25.554	12.083	1
2	66.711	11.384	1
3	61.189	15.191	1
4	9.153	119.150	0
5	98.794	45.258	1
6	14.577	139.149	0
7	65.126	9.877	1
8	13.205	12.106	0
9	20.535	21.613	0
10	62.278	13.289	1
⋮	⋮	⋮	⋮
1 000	19.705	12.873	0

Entre los primeros diez pares (A_i^*, B_i^*) hay seis para los cuales $A_i^* > 2B_i^*$. Por consecuencia, si se basaran los resultados en los diez primeros valores, se estimaría que $P(A > 2B) = 6/10 = 0.06$. Obviamente, los diez pares simulados no se aproximan lo suficiente para calcular una estimación confiable. Entre los 1 000 pares simulados, hay 460 para los cuales $A_i^* > 2B_i^*$. Por tanto, se estima $P(A > 2B) = 0.460$. La ingeniera escoge el tipo B . Se observa que esta probabilidad se puede calcular sólo con una integral múltiple. La probabilidad exacta es $5/11 = 0.4545$. La aproximación de la simulación es muy buena.

Una muestra simulada correctamente a partir de una distribución de probabilidad dada es, de hecho, una muestra aleatoria simple de dicha distribución. Por consiguiente, se puede utilizar la media y la varianza muestrales simuladas para estimar la media y la varianza de la distribución, y se puede utilizar una gráfica de probabilidad para determinar si la distribución de probabilidad fue bien aproximada por una función de densidad estándar, como la curva normal. A continuación, se presentan algunos ejemplos.

Estimación de medias y varianzas

El ejemplo 4.71 muestra cómo se puede emplear los valores simulados para estimar la media y la desviación estándar de una población.

Ejemplo

4.71

Utilice los valores simulados R_i^* en la tabla 4.2 para estimar la media μ_R y la desviación estándar σ_R de la resistencia total R .

Solución

Se puede considerar a los valores $R_1^*, \dots, R_{100}^*$ como si fueran una muestra aleatoria de las resistencias reales. Por consiguiente, se estima μ_R con la media muestral $\bar{R}^*$ y σ_R con la des-

viación estándar muestral s_{R^*} . La media y la desviación estándar muestrales de $R_1^*, \dots, R_{100}^*$ son $\bar{R}^* = 19.856$ y $s_{R^*} = 1.6926$, respectivamente. Éstas son las estimaciones de μ_R y σ_R .

Comparación con la propagación de errores

En el ejemplo 4.71 se utilizó la simulación para aproximar la media y la desviación estándar de una función de variables aleatorias $R = XY/(X + Y)$. El método de propagación de errores, analizado en la sección 3.4, también se puede emplear con este propósito (véase el ejemplo 3.20). Por supuesto, la simulación puede hacer muchas cosas que la propagación de errores no, como estimar probabilidades y determinar si una función dada de variables aleatorias se distribuye normalmente. Pero si lo que se necesita es estimar la desviación estándar de una función de variables aleatorias, es natural preguntar si la simulación o la propagación de errores es la mejor técnica. La respuesta es que cada método tiene ventajas y desventajas.

Para que el análisis de esta cuestión sea concreto, sean $X_1, \dots, X_n$ variables aleatorias independientes y $U = U(X_1, \dots, X_n)$ una función. Se desea calcular σ_U . Lo primero que se necesita decir es que, en muchos casos, ambos métodos funcionan bien y dan resultados similares, por lo que sólo es una cuestión de conveniencia lo que se utiliza. La simulación tiene una ventaja, que no requiere que las desviaciones estándar de $X_1, \dots, X_n$ sean pequeñas, como lo hace la propagación de errores. Sin embargo, esta última tiene dos grandes ventajas. Primero, no es necesario conocer las distribuciones de $X_1, \dots, X_n$, como se necesita en simulación. Segundo, puede indicar cuál de las X contribuye más a la incertidumbre de U , lo que no se puede hacer tan fácil con la simulación.

Uso de la simulación para determinar si una población es aproximadamente normal

Una de las cuestiones que surge con mayor frecuencia en el análisis de datos es si una población se distribuye aproximadamente normal. Cuando se tiene a la disposición una muestra simulada de una población, es posible resolver esta cuestión.

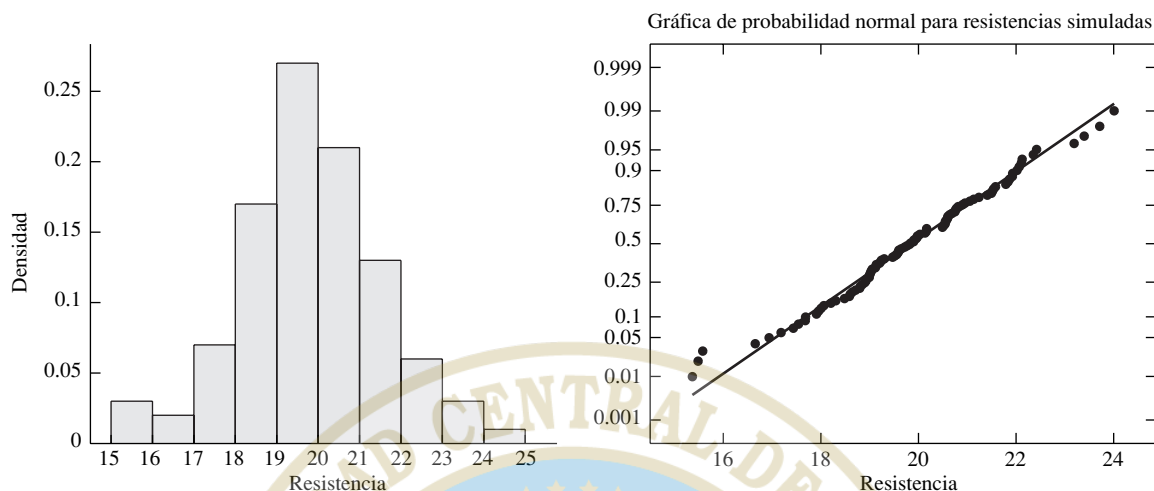
Ejemplo

4.72

Construya un histograma de los valores simulados de R^* presentados en la tabla 4.2. Construya una gráfica de probabilidad normal para determinar si la densidad de la resistencia R es aproximadamente normal.

Solución

La siguiente figura muestra el histograma y la gráfica de probabilidad. El primero es aproximadamente simétrico y tiene una moda. Esto es consistente con la normalidad. La segunda indica una ligera desviación de la normalidad, especialmente en las colas. Es muy razonable decir que la distribución parece ser aproximadamente normal. En la práctica, una muestra de tamaño 1 000 o más generaría un histograma más preciso. Una muestra de 100 es adecuada para la gráfica de probabilidad, aunque no haya más problemas para generar una muestra más grande.



El ejemplo 4.73 muestra cómo se puede usar la simulación para determinar si el tamaño de una muestra es lo suficientemente grande para que sea válido el teorema del límite central, si se conoce la distribución de la cual se extrajo la muestra.

Ejemplo 4.73

El artículo “Thermal Absorption from Pesticide Residues” (M. Reddy y A. Bunge, en *The Practical Applicability of Toxicokinetic Models in the Risk Assessment of Chemicals*, 2002:55-79) modela la cantidad de pesticida absorbida en el sistema como una variable aleatoria lognormal, cuya media es proporcional a la dosis. Suponga que para cierta dosis, la cantidad absorbida sigue una distribución lognormal con parámetros $\mu = 1$ y $\sigma = 0.5$. Se llevará a cabo un experimento en el que se aplica esta dosis en cada uno de los cinco experimentos independientes, y la cantidad absorbida se determinará cada vez. ¿La cantidad absorbida en promedio sigue una distribución aproximadamente normal?

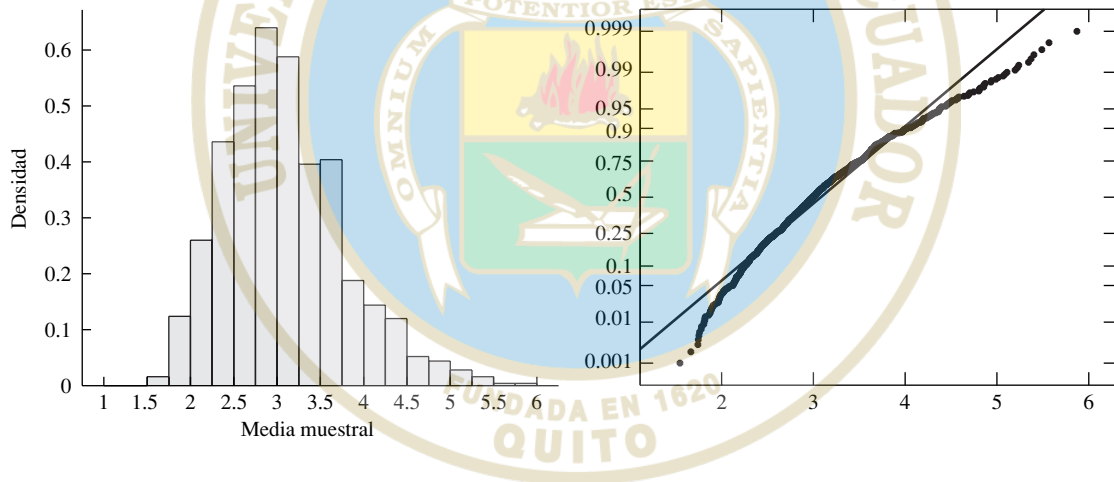
Solución

Sea $X_1, \dots, X_5$ una muestra aleatoria de una distribución lognormal con parámetros $\mu = 1$ y $\sigma = 0.5$. La pregunta es si la media muestral $\bar{X}$ tiene una distribución aproximadamente normal. Se responderá esta pregunta al generar 1 000 muestras aleatorias simuladas de tamaño cinco a partir de esta distribución lognormal, al calcular la media muestral de cada una de ellas, y después al construir una gráfica de probabilidad para las 1 000 medias muestrales. La tabla 4.4 presenta las primeras tres y las últimas tres de las muestras. Las primeras cinco columnas en cada renglón de la tabla 4.4 constituyen una muestra aleatoria simple $X_{1i}^*, \dots, X_{5i}^*$ proveniente de una distribución lognormal con parámetros $\mu = 1$ y $\sigma = 0.5$. La sexta columna es la media muestral $\bar{X}_i^*$. Por consecuencia, las 1 000 entradas en la sexta columna son una muestra aleatoria de medias muestrales. Al construir una gráfica de probabilidad normal de estos valores, se puede determinar si la media muestral está distribuida normalmente.

TABLA 4.4 Datos simulados para el ejemplo 4.73

	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*	$\bar{X}^*$
1	2.3220	1.5087	1.2144	2.5092	3.3408	2.1790
2	3.3379	2.8557	1.0023	3.8088	2.3320	2.6673
3	2.9338	3.0364	3.1488	2.0380	4.7030	3.1720
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
998	4.7993	3.7609	1.5751	3.6382	2.0254	3.1598
999	3.7929	2.9527	6.3663	1.8057	10.4450	5.0725
1 000	3.7680	4.5899	2.8609	2.1659	5.0658	3.6901

A continuación se muestra un histograma y una gráfica de probabilidad normal de los 1 000 valores de $\bar{X}^*$. El histograma muestra que la distribución está sesgada a la derecha. La gráfica de probabilidad confirma que la distribución está muy lejos de ser normal.



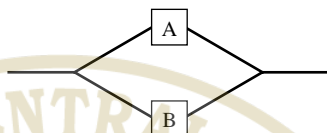
Uso de la simulación en los análisis de confiabilidad

Un sistema está formado por componentes, cada uno tiene un tiempo de vida aleatorio. Por tanto, el tiempo de vida de un sistema también es aleatorio. Con frecuencia los ingenieros especialistas en confiabilidad conocen las distribuciones de probabilidad de los tiempos de vida de los componentes y desean determinar la distribución de probabilidad del sistema. En la práctica puede ser muy difícil calcular directamente la distribución del tiempo de vida del sistema a partir de las distribuciones de los tiempos de vida de los componentes. No obstante, si los tiempos de vida de éstos son independientes, con frecuencia ello puede realizarse fácilmente con simulación. En seguida se muestra un ejemplo sencillo.

Ejemplo

4.74

Un sistema está formado por los componentes A y B conectados en paralelo, tal y como se muestra en la siguiente ilustración esquemática. El tiempo de vida en meses del componente A se distribuye $\text{Exp}(1)$ y el tiempo de vida en meses del componente B se distribuye $\text{Exp}(0.5)$. El sistema funcionará hasta que A y B fallen. Estime la media del tiempo de vida del sistema, la probabilidad de que el sistema funcione durante menos de un mes, y el 10o. percentil de los tiempos de vida del sistema.



Solución

Se genera una muestra $A_1^*, \dots, A_{1000}^*$ de los tiempos de vida simulados del componente A a partir de una distribución $\text{Exp}(1)$. Después se genera una muestra $B_1^*, \dots, B_{1000}^*$ de los tiempos de vida simulados del componente B a partir de una distribución $\text{Exp}(0.5)$. Observe que la media del tiempo de vida del componente A es de un mes y la media del tiempo de vida del componente B es $1/0.5 = 2$ meses. El tiempo de vida del i -ésimo sistema simulado es $L_i^* = \max(A_i^*, B_i^*)$. La tabla 4.5 muestra los resultados de las primeras diez muestras y de la última muestra.

TABLA 4.5 Datos simulados para el ejemplo 4.74

	A^*	B^*	L^*
1	0.0245	0.5747	0.5747
2	0.3623	0.3998	0.3998
3	0.8858	1.7028	1.7028
4	0.1106	14.2252	14.2252
5	0.1903	0.4665	0.4665
6	2.2259	1.4138	2.2259
7	0.8881	0.9120	0.9120
8	3.3471	3.2134	3.3471
9	2.5475	1.3240	2.5475
10	0.3614	0.8383	0.8383
⋮	⋮	⋮	⋮
1 000	0.3619	1.8799	1.8799

La media muestral de los primeros diez valores de L_i^* es 2.724. Cinco de ellos son menores a 1. El 10o. percentil de estos valores es $(0.3998 + 0.4665)/2 = 0.43315$. Por tanto, si las estimaciones se basaran en las primeras diez muestras, se estimaría que la media del tiempo de vida del sistema es de 2.724 meses, la probabilidad de que el sistema falle en un mes es 0.5, y el 10o. percentil de las duraciones del sistema es 0.43315. Obviamente, diez muestras

no es ni siquiera una estimación suficientemente confiable. Con base en las 1 000 muestras, la estimación de la media del tiempo de vida fue de 2.29 meses, la estimación de la probabilidad de error dentro de un mes fue de 0.278 y el 10o. percentil fue de 0.516 meses.

Uso de la simulación para estimar sesgos

La simulación se puede utilizar para estimar sesgos. El ejemplo 4.75 muestra cómo.

Ejemplo 4.75

Si $X_1, \dots, X_n$ es una muestra aleatoria, entonces se usa la desviación estándar muestral s para estimar la desviación estándar poblacional σ . Sin embargo, s es un estimador sesgado. Si $X_1, \dots, X_6$ es una muestra aleatoria de una distribución $N(0, 1)$, utilice la simulación para estimar el sesgo en s . Asimismo, estime la desviación estándar σ_s de s .

Solución

Se generarán 1 000 muestras aleatorias $X_{1i}^*, \dots, X_{6i}^*$ de tamaño seis a partir de $N(0, 1)$ y para cada una se calculará la desviación estándar muestral s_i^* . La tabla 4.6 muestra los resultados de las diez primeras y de la última muestra.

TABLA 4.6 Datos simulados para el ejemplo 4.75

	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*	X_6^*	s^*
1	-0.4326	0.7160	-0.6028	0.8304	-0.1342	-0.3560	0.6160
2	-1.6656	1.5986	-0.9934	-0.0938	0.2873	-1.8924	1.3206
3	0.1253	-2.0647	1.1889	-0.4598	0.3694	0.4906	1.1190
4	-1.7580	0.1575	-0.8496	0.3291	-1.5780	-1.1100	0.8733
5	1.6867	0.3784	0.3809	0.4870	0.9454	-0.4602	0.7111
6	1.3626	0.7469	-2.1102	2.6734	-0.5311	1.1611	1.6629
7	-2.2424	-0.5719	-1.9659	0.1269	-0.2642	0.3721	1.0955
8	1.3765	-0.4187	-0.5014	1.9869	-0.0532	-0.7086	1.1228
9	-1.8045	0.5361	-0.9121	1.4059	-1.2156	-0.9619	1.2085
10	0.3165	0.6007	-0.5363	-0.2300	0.2626	0.0523	0.4092
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1 000	0.3274	0.1787	0.2006	-1.1602	1.1020	0.3173	0.7328

Los valores $s_1^*, \dots, s_{1\,000}^*$ son una muestra aleatoria de la población de todos los posibles valores de s que se puede calcular a partir de una muestra normal de tamaño seis. Por consiguiente, la media muestral $\bar{s}^*$ es una estimación de la media poblacional μ_s . Ahora, la desviación estándar real de la distribución a partir de la cual se generó los datos simulados es $\sigma = 1$, por lo que el sesgo en s es $\mu_s - 1$. Se estima el sesgo con $\bar{s}^* - 1$.

La media muestral de los primeros diez valores de s_i^* es 1.0139. Por tanto, si los resultados se basaran en los primeros diez valores, se estimaría que el sesgo es $1.0139 - 1 = 0.0139$. Por supuesto, diez valores no son suficientes para construir una estimación confiable.

La media muestral de los 1 000 valores s_i^* es $\bar{s}^* = 0.9601$. Se estima que el sesgo es $0.9601 - 1 = -0.0399$.

La desviación estándar muestral de los 1 000 valores es 0.3156. Ésta es la estimación de σ_s .

La estimación Bootstrap

En los ejemplos analizados hasta ahora en esta sección, se ha especificado la distribución a partir de la cual se generan los datos simulados. En algunos casos, esta distribución se puede determinar de los datos. Los métodos de simulación en los cuales la distribución que será extraída se determina a partir de los datos se llaman de **estimación bootstrap**. Para ilustrar lo anterior, se presenta una variación del ejemplo 4.75, donde la distribución de la muestra se determina a partir de los datos.

Ejemplo 4.76

Una muestra de tamaño seis es tomada de una distribución normal cuya media y varianza son desconocidas. Los valores de la muestra son 5.23, 1.93, 5.66, 3.28, 5.93 y 6.21. La media muestral es $\bar{X} = 4.7067$ y la desviación estándar muestral es $s = 1.7137$. El valor de s se usará para estimar la desconocida desviación estándar σ de la población. Estime el sesgo en s .

Solución

Si se conociera la media poblacional μ y la desviación estándar σ de la distribución normal de la que proviene la muestra, se puede utilizar el método del ejemplo 4.75, para simular una distribución $N(\mu, \sigma)$. En virtud de que no se conocen estos valores, se estimarán con los valores muestrales $\bar{X} = 4.7067$ y $s = 1.7137$. Se procederá exactamente igual que en el ejemplo 4.75, excepto que se extraerá una muestra de una distribución $N(4.7067, 1.7137^2)$. Puesto que la distribución se determinó a partir de los datos, éste es un método de estimación bootstrap.

Se generarán 1 000 muestras aleatorias $X_{1i}^*, \dots, X_{6i}^*$ de tamaño seis de $N(4.7067, 1.7137^2)$ y para cada una se calculará la desviación estándar muestral s_i^* . La tabla 4.7 presenta los resultados de las primeras diez y de la última muestra.

TABLA 4.7 Datos simulados para el ejemplo 4.76

	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*	X_6^*	s^*
1	2.3995	4.8961	3.6221	6.9787	4.4311	4.5367	1.5157
2	2.6197	4.3102	3.2350	6.2619	4.4233	3.5903	1.2663
3	3.0114	5.2492	7.6990	6.0439	6.5965	3.7505	1.7652
4	3.9375	5.2217	1.9737	4.5434	3.0304	3.8632	1.1415
5	5.8829	5.3084	4.6003	2.6439	2.3589	2.3055	1.6054
6	7.8915	3.9731	5.1229	5.1749	3.5255	3.3330	1.6884
7	4.2737	5.5189	2.3314	5.1512	5.7752	4.0205	1.2705
8	5.8602	5.3280	5.5860	6.8256	7.5063	3.9393	1.2400
9	5.7813	4.9364	2.5893	3.7633	0.9065	3.8372	1.7260
10	3.3690	1.8618	2.7627	3.2837	3.9863	6.0382	1.4110
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1 000	2.0496	6.3385	6.2414	5.1580	3.7213	8.4576	2.2364

Los valores $s_1^*, \dots, s_{1000}^*$ son una muestra aleatoria de la población de todos los posibles valores de s que se puede calcular a partir de una muestra normal de tamaño seis. Por tanto, la media muestral $\bar{s}^*$ es una estimación de la media poblacional μ_s . Ahora, la desviación estándar poblacional a partir de la cual se generaron los datos simulados es $\sigma^* = 1.7137$. El sesgo se estima con $\bar{s}^* - 1.7137$.

La media muestral de los diez primeros valores de s_i^* es 1.4630. Por consiguiente, si los resultados se basaran en los diez primeros valores, se estimaría que el sesgo es $1.4630 - 1.7137 = -0.2507$. Por supuesto, diez valores no son suficientes para construir una estimación confiable. La media muestral de los 1 000 valores de s_i^* es 1.6188. Se estima que el sesgo es $1.6188 - 1.7137 = -0.0949$.

Algunas veces se pueden utilizar los resultados de la estimación bootstrap para ajustar estimaciones con el fin de que sean más exactas. El ejemplo 4.77 muestra cómo se puede realizar esto último con la desviación estándar muestral.

Ejemplo 4.77

En el ejemplo 4.76 se tomó una muestra de tamaño seis de una población $N(\mu, \sigma^2)$. La desviación estándar muestral $s = 1.7137$ es una estimación de la desviación estándar poblacional no conocida σ . Utilice los resultados de la estimación bootstrap en el ejemplo 4.76 para reducir el sesgo en esta estimación.

Solución

Se estima que el sesgo en s es -0.0949 . Esto significa que, en promedio, la desviación estándar muestral calculada a partir de esta población $N(\mu, \sigma^2)$ será menor que la desviación estándar real σ por aproximadamente -0.0949 . Por consecuencia, se corrige el sesgo sumando 0.0949 a la estimación. La estimación con sesgamiento corregido de la desviación estándar poblacional es $1.7137 + 0.0949 = 1.81$.

Estimación bootstrap paramétrica y no paramétrica

En el ejemplo 4.76 se sabía que la muestra provenía de una distribución normal, pero no se conocían la media ni la varianza. Por tanto, se empleó los datos para estimar los parámetros μ y σ . A este procedimiento se le llama **estimación bootstrap paramétrica**, ya que los datos son utilizados para estimar parámetros. ¿Qué pasaría si no se hubiera tenido conocimiento de que la distribución era normal? Entonces se hubiera utilizado la **estimación bootstrap no paramétrica**. En ésta se hace una simulación mediante el muestreo de los propios datos. Además, sirve para construir intervalos de confianza y realizar pruebas de hipótesis. Aquí se definirá la estimación bootstrap no paramétrica, y después se presentarán algunas aplicaciones en las secciones 5.8 y 6.15.

Si se tuviera una muestra $X_1, \dots, X_n$ de una distribución desconocida, se simularían muestras $X_{1i}^*, \dots, X_{ni}^*$ de la siguiente manera. Imagine que se colocan los valores $X_1, \dots, X_n$ en una caja, y que se saca un valor aleatoriamente. Después se reemplaza el valor y se saca otra vez. La segunda extracción es también una extracción muestral $X_1, \dots, X_n$. Se continúa hasta que se han realizado n extracciones. Ésta es la primera muestra simulada, denominada muestra de estimación bootstrap: $X_{11}^*, \dots, X_{n1}^*$. Observe que puesto que se realiza el mues-

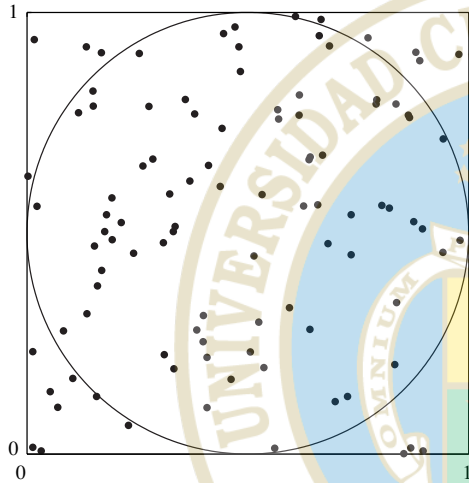
treo con reemplazo, la muestra de estimación bootstrap probablemente contendrá más de una vez a algunos de los elementos originales de la muestra y a otros que no aparezcan. Ahora se toma más muestras de estimación bootstrap: tantas como las que se tomarían en cualquier simulación, probablemente 1 000 o más. Después se procede como en cualquier otra simulación.

Para mayor información acerca de la estimación bootstrap y otros procesos de simulación, Efron y Tibshirani (1993) son una excelente fuente de información.

Ejercicios para la sección 4.11

- El vendedor A distribuye partes, donde cada una tiene una probabilidad de 0.03 de estar defectuosa. El vendedor B también distribuye partes y cada una tiene una probabilidad de 0.05 de estar defectuosa. Usted recibe un envío de 100 partes por parte de cada vendedor.
 - Sea X el número de partes defectuosas en el envío provenientes del vendedor A y Y el envío proveniente del vendedor B. ¿Cuáles son las distribuciones de X y Y ?
 - Genere muestras simuladas de tamaño 1 000 a partir de las distribuciones de X y Y .
 - Utilice las muestras para estimar la probabilidad de que el número total de partes defectuosas sea menor a diez.
 - Utilice las muestras para estimar la probabilidad de que el envío del vendedor A tiene más partes defectuosas que el envío del vendedor B.
 - Construya una gráfica de probabilidad normal para el número total de partes defectuosas. ¿La cantidad sigue una distribución aproximadamente normal?
- Dos diseños de cierto circuito de semiconductores están compitiendo entre sí. La duración del primero (en horas) se distribuye exponencialmente con $\lambda = 10^{-4}$ y la duración del segundo tiene una distribución lognormal con $\mu = 6$ y $\sigma^2 = 5.4$.
 - Utilice una muestra simulada de tamaño 1 000 para estimar la probabilidad de que un circuito con el primer diseño dure más que otro con el segundo diseño.
 - Estime la probabilidad de que un circuito con el primer diseño dure el doble de tiempo que otro con el segundo diseño.
- Se fabrican placas rectangulares cuyas longitudes se distribuyen como $N(2.0, 0.1^2)$ y cuyos anchos se distribuyen como $N(3.0, 0.2^2)$. Suponga que las longitudes y los anchos son independientes. El área de una placa está dada por $A = XY$.
 - Utilice una muestra simulada de tamaño 1 000 para estimar la media y varianza de A .
 - Estime la probabilidad de que $P(5.9 < A < 6.1)$.
 - Construya una gráfica de probabilidad normal para las áreas. ¿El área de una placa sigue una distribución aproximadamente normal?
- Un cable está compuesto por cuatro alambres. La fuerza de ruptura de cada alambre es una variable aleatoria distribuida normalmente con media de 10 kN y desviación estándar de 1 kN. Utilizando el método de cable quebradizo, se estima que la fuerza del cable es igual a la fuerza de alambre más frágil multiplicada por el número de alambres.
 - Utilice muestras simuladas de tamaño 1 000 para estimar la fuerza media de este tipo de cable.
 - Estime la mediana de la fuerza del cable.
 - Estime la desviación estándar de la fuerza del cable.
 - Para que sea aceptable en cierta aplicación, la probabilidad de que el cable se rompa con una carga de 28 kN debe ser menor a 0.01. ¿Parece ser que el cable es aceptable? Explique.
- La duración de un láser (en horas) tiene una distribución lognormal con $\mu = 8$ y $\sigma^2 = 2.4$. Dos de esos láser funcionan de forma independiente.
 - Utilice una muestra de tamaño 1 000 para estimar la probabilidad de que la suma de las dos duraciones sea mayor a 20 000 horas.
 - Estime la probabilidad de que ambos láser duren más de 3 000 horas.
 - Estime la probabilidad de que ambos láser fallen antes de las 10 000 horas.
- Estimación del valor de π .* La siguiente figura sugiere cómo estimar el valor de π con una simulación. En la figura, un

círculo con un área igual a $\pi/4$ está inscrito en un cuadrado cuya área es igual a 1. Se elige de forma aleatoria 100 puntos dentro del cuadrado. La probabilidad de que un punto esté dentro del círculo es igual a la fracción del área del cuadrado que abarca a éste, la cual es $\pi/4$. Por tanto, se puede estimar el valor de $\pi/4$ al contar el número de puntos dentro del círculo, que es 79, y al dividir entre el número total de puntos, que es 100, para obtener la estimación $\pi/4 \approx 0.79$. De esto último se concluye que $\pi \approx 4(0.79) = 3.16$. Este ejercicio presenta un experimento de simulación que fue diseñado para estimar el valor de π al generar 1 000 puntos en el cuadrado.



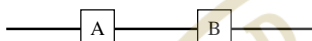
- Genere 1 000 coordenadas $x_1^*, \dots, x_{1\,000}^*$. Utilice la **distribución uniforme** con valor mínimo de 0 y valor máximo de 1. La distribución uniforme genera variables aleatorias que tienen la misma probabilidad de venir de cualquier parte del intervalo (0, 1).
- Genere 1 000 coordenadas $y_1^*, \dots, y_{1\,000}^*$, utilizando nuevamente la **distribución uniforme** con valor mínimo de 0 y valor máximo de 1.
- Cada punto (x_i^*, y_i^*) se encuentra dentro del círculo si su distancia desde el centro (0.5, 0.5) es menor a 0.5. Para cada par (x_i^*, y_i^*) determine si la distancia desde el centro es menor a 0.5. Esto último se puede realizar al calcular el valor $(x_i^* - 0.5)^2 + (y_i^* - 0.5)^2$, que es el cuadrado de la distancia, y al determinar si es menor que 0.25.
- ¿Cuántos de los puntos están dentro del círculo? ¿Cuál es su estimación de π ? (Nota: Con sólo 1 000 puntos, es probable que su estimación sea inferior por 0.05 o más. Una simulación con 10 000 y 100 000 puntos tiene mayores probabilidades de dar como resultado una estimación muy cercana al valor verdadero.

7. *Aplicación a redes de computadoras móviles.* Con frecuencia los expertos en computación modelan el movimiento de una computadora móvil como una trayectoria aleatoria dentro de un rectángulo. Esto es, se elige aleatoriamente dos puntos dentro del rectángulo y la computadora se mueve en una línea recta desde el primer punto al segundo. En el estudio de redes de computadoras móviles, es importante conocer la media de la longitud de una trayectoria (véase el artículo “Stationary Distributions for Random Waypoint Models”, W. Navidi y T. Camp, en *IEEE, Transactions on Mobile Computing*, 2004:99-108). Es muy difícil calcular directamente esta media, pero es sencillo estimarla con una simulación. Si los puntos finales de una trayectoria están representados por (X_1, Y_1) y (X_2, Y_2) , entonces la longitud de la trayectoria es $\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$. La media de la longitud se estima al generar puntos finales (X_1^*, Y_1^*) y (X_2^*, Y_2^*) para muchas trayectorias, calcular la longitud de cada una, y estimar su media. Este ejercicio presenta un experimento de simulación que fue diseñado para estimar la media de la distancia entre dos puntos seleccionados aleatoriamente dentro de un cuadrado de lado 1.

- Genere 1 000 conjuntos de puntos finales (X_{1i}^*, Y_{1i}^*) y (X_{2i}^*, Y_{2i}^*) . Utilice la distribución uniforme con valor mínimo de 0 y valor máximo de 1 para cada coordenada en cada punto. La distribución uniforme genera valores que tienen la misma probabilidad de provenir de cualquier parte del intervalo (0, 1).
 - Calcule las longitudes de las 1 000 trayectorias $L_i^* = \sqrt{(X_{2i}^* - X_{1i}^*)^2 + (Y_{2i}^* - Y_{1i}^*)^2}$.
 - Calcule la media muestral de la longitud de la trayectoria $\bar{L}^*$. La media verdadera, con seis dígitos significativos, es 0.521405. ¿Qué tan cercano está su resultado?
 - Estime la probabilidad de que una trayectoria tenga más de una unidad de largo.
8. Con referencia al ejemplo 4.74 (p. 289), con el fin de incrementar el tiempo de vida del sistema, los ingenieros deben decidir entre reemplazar el componente A por uno cuyo tiempo de vida se distribuya $\text{Exp}(1/2)$, o reemplazar el componente B con uno cuyo tiempo de vida se distribuya $\text{Exp}(1/3)$.
- Genere, mediante simulación, un gran número (al menos 1 000) de los tiempos de vida del sistema, bajo el supuesto de que se reemplaza el componente A.
 - Genere, con simulación, un gran número (al menos 1 000) de los tiempos de vida del sistema, bajo el supuesto de que se reemplaza el componente B.
 - Si el objetivo es maximizar la media del tiempo de vida del sistema, ¿cuál es la mejor opción? Explique.

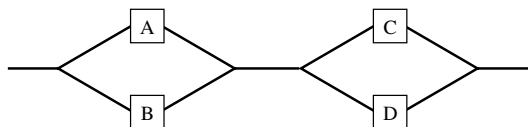
- d) Si el objetivo es minimizar la probabilidad de que el sistema falle en un mes, ¿cuál es la mejor opción? Explique.
- e) Si el objetivo es minimizar al 10o. percentil de los tiempos de vida del sistema, ¿cuál es la mejor opción? Explique.

9. Un sistema está compuesto por los componentes A y B conectados en serie, como lo muestra la siguiente ilustración esquemática. El tiempo de vida en meses del componente A sigue una distribución lognormal con $\mu = 1$ y $\sigma = 0.5$, y la duración en meses del componente B tiene una distribución lognormal con $\mu = 2$ y $\sigma = 1$. El sistema sólo funcionaría si A y B lo hacen.



- a) Genere, por simulación, un gran número (al menos 1 000) de los tiempos de vida del sistema.
- b) Estime la media del tiempo de vida del sistema.
- c) Estime la probabilidad de que el sistema falle en dos meses.
- d) Estime el 20o. percentil de los tiempos de vida del sistema.
- e) Construya una gráfica de probabilidad normal de los tiempos de vida del sistema. ¿El tiempo de vida del sistema tiene una distribución aproximadamente normal?
- f) Construya un histograma de los tiempos de vida del sistema. ¿Está sesgado a la izquierda, sesgado a la derecha, o es aproximadamente simétrico?

10. Un sistema está compuesto por dos subsistemas conectados en serie, como lo muestra la siguiente ilustración esquemática. Cada subsistema consiste en dos componentes conectados en paralelo. El subsistema AB falla cuando no funcionan A y B. El subsistema CD falla cuando lo hacen C y D. El sistema falla tan pronto como alguno de los dos subsistemas falla. Suponga que los tiempos de vida de los componentes, en meses, tiene las siguientes distribuciones: A: Exp(1), B: Exp(0.1), C: Exp(0.2), D: Exp(0.2).



- a) Genere, mediante simulación, un gran número (al menos 1 000) de los tiempos de vida del sistema.
- b) Estime la media del tiempo de vida del sistema.
- c) Estime la mediana del tiempo de vida del sistema.

- d) Estime la probabilidad de que el sistema funcione más de seis meses.
- e) Estime el 90o. percentil de los tiempos de vida del sistema.
- f) Estime la probabilidad de que el subsistema AB falle antes que lo haga el subsistema CD.

11. (Continuación del ejercicio 12 de la página 280.) La edad de una antigua pieza de materia orgánica se puede estimar a partir de la tasa a la que emite partículas beta como resultado del decaimiento del carbono-14. Por ejemplo, si X es el número de partículas emitidas durante diez minutos por un fragmento óseo con 10 000 años de antigüedad que contiene 1 g de carbono, entonces X tiene una distribución de Poisson con media $\lambda = 45.62$. Un arqueólogo descubrió un pequeño fragmento óseo que contiene exactamente 1 g de carbono. Si t es la edad desconocida del hueso, en años, el arqueólogo contará el número X de partículas emitidas en diez minutos y calculará una edad estimada $\hat{t}$ con la fórmula

$$\hat{t} = \frac{\ln 15.3 - \ln(X/10)}{0.0001210}$$

El arqueólogo no lo sabe, pero el hueso tiene exactamente 10 000 años de antigüedad, por lo que X tiene una distribución de Poisson con $\lambda = 45.62$.

- a) Genere una muestra simulada de 1 000 valores de X y sus valores correspondientes de $\hat{t}$.
- b) Estime la media de $\hat{t}$.
- c) Estime la desviación estándar de $\hat{t}$.
- d) Estime la probabilidad de que $\hat{t}$ esté a 1 000 años con una edad real de 10 000 años.
- e) Estime la probabilidad de que $\hat{t}$ esté a más de 2 000 años con una edad real de 10 000 años.
- f) Construya una gráfica de probabilidad normal para $\hat{t}$. ¿ $\hat{t}$ se distribuye aproximadamente en forma normal?
12. Se toma una muestra aleatoria de una distribución normal con el propósito de estimar la media poblacional μ . Puesto que μ es la mediana y la media, parece ser que tanto la mediana muestral m como la media muestral $\bar{X}$ son estimadores razonables. Este ejercicio está diseñado para determinar cuál de estos estimadores tiene la menor incertidumbre.

- a) Genere un gran número de (al menos 1 000) muestras de tamaño cinco de una distribución $N(0, 1)$.
- b) Calcule las medianas muestrales $m_1^*, \dots, m_{1000}^*$ de las 1 000 muestras.
- c) Calcule la media $\bar{m}^*$ y la desviación estándar s_{m^*} de $m_1^*, \dots, m_{1000}^*$.

- d) Calcule las medias muestrales $\bar{X}^*, \dots, \bar{X}_{1000}^*$ de las 1 000 muestras.
- e) Calcule la media y la desviación estándar $s_{\bar{X}}^*$ de $\bar{X}_1^*, \dots, \bar{X}_{1000}^*$.
- f) El valor verdadero de μ es 0. Estime el sesgo y la incertidumbre (σ_m) en m . (Nota: De hecho, la mediana no tiene sesgos, por lo que la estimación del sesgo debe aproximarse a 0.)
- g) Estime el sesgo y la incertidumbre ($\sigma_{\bar{X}}$) en $\bar{X}$. ¿Su estimación del sesgo se aproxima a 0? Explique por qué. ¿Su estimación de la incertidumbre se aproxima a $1/\sqrt{5}$? Explique por qué.
13. Se toma una muestra aleatoria de tamaño ocho de una distribución $\text{Exp}(\lambda)$, donde no se conoce λ . Los valores de la

muestra son 2.74, 6.41, 4.96, 1.65, 6.38, 0.19, 0.52 y 8.38. Este ejercicio muestra cómo se emplea la estimación bootstrap para estimar el sesgo y la incertidumbre ($\sigma_{\hat{\lambda}}$) en la estimación $\hat{\lambda} = 1/\bar{X}$.

- a) Calcule $\hat{\lambda} = 1/\bar{X}$ para la muestra específica.
- b) Genere 1 000 muestras de estimación bootstrap de tamaño ocho a partir de la distribución $\text{Exp}(\hat{\lambda})$.
- c) Calcule los valores $\hat{\lambda}_i^* = 1/\bar{X}_i^*$ para cada una de las 1 000 muestras bootstrap.
- d) Calcule la media muestral $\bar{\hat{\lambda}}^*$ y la desviación estándar muestral $s_{\hat{\lambda}}^*$ de $\hat{\lambda}_1^*, \dots, \hat{\lambda}_{1000}^*$.
- e) Estime el sesgo y la incertidumbre ($\sigma_{\hat{\lambda}}$) en $\hat{\lambda}$.

Ejercicios adicionales para el capítulo 4

- Un avión tiene 100 asientos de pasajeros. Suponga que la probabilidad de que una persona con boleto llegue al vuelo es de 0.90. Si la aerolínea vende 105 boletos, ¿cuál es la probabilidad de que todas las personas tengan asiento?
- El número de grietas grandes en un pavimento de determinada longitud a lo largo de una cierta calle tiene una distribución de Poisson con una media de una grieta por cada 100 m.
 - ¿Cuál es la probabilidad de que haya exactamente ocho grietas en una longitud de 500 m de pavimento?
 - ¿Cuál es la probabilidad de que no haya grietas en una longitud de 100 m de pavimento?
 - Sea T la distancia en metros entre dos grietas sucesivas, ¿cuál es la función de densidad de probabilidad de T ?
 - ¿Cuál es la probabilidad de que la distancia entre dos grietas sucesivas sea de más de 50 m?
- Unas plantas de frijoles contienen dos genes del color de la semilla, donde cada uno puede ser Y (para semillas amarillas) o G (para semillas verdes). A las plantas que contienen uno de cada tipo de gen se les llama *heterozygous*. De acuerdo con la teoría genética de Mendel, si se cruzan dos plantas *heterozygous*, cada uno de sus descendientes tendrá probabilidad de 0.75 de tener semillas amarillas y 0.25 de que sean verdes.
 - De los diez descendientes de las plantas *heterozygous*, ¿cuál es la probabilidad de que sólo tres tengan semillas verdes?
 - De los diez descendientes de las plantas *heterozygous*, ¿cuál es la probabilidad de que más de dos tengan semillas verdes?
 - De los 100 descendientes de las plantas *heterozygous*, ¿cuál es la probabilidad de que más de 30 tengan semillas verdes?
 - De los 100 descendientes de las plantas *heterozygous*, ¿cuál es la probabilidad de que entre, e incluyendo a, 30 y 35, tengan semillas verdes?
 - De los 100 descendientes de las plantas *heterozygous*, ¿cuál es la probabilidad de que menos de 80 tengan semillas amarillas?
- Se extrae una muestra aleatoria $X_1, \dots, X_n$ de una población, y las cantidades $\ln X_1, \dots, \ln X_n$ se incluyen en una gráfica de probabilidad normal. Los puntos siguen casi una línea recta. Verdadero o falso:
 - $X_1, \dots, X_n$ provienen de una población que es aproximadamente lognormal.
 - $X_1, \dots, X_n$ provienen de una población que es aproximadamente normal.
 - $\ln X_1, \dots, \ln X_n$ provienen de una población que es aproximadamente lognormal.
 - $\ln X_1, \dots, \ln X_n$ provienen de una población que es aproximadamente normal.
- La agencia de protección ambiental (EPA, por sus siglas en inglés) ha contratado a una compañía para que dé segui-

miento a la calidad del agua de diversos lagos en su distrito. Se emplean diez dispositivos. Suponga que cada uno tiene una probabilidad de 0.01 de fallar durante el transcurso del periodo de seguimiento.

- a) ¿Cuál es la probabilidad de que no falle ninguno de estos dispositivos?
 - b) ¿Cuál es la probabilidad de que fallen dos o más dispositivos?
 - c) Si la EPA requiere que la probabilidad de que ninguno de estos dispositivos falle sea de al menos 0.95, ¿cuál es la probabilidad de falla individual más grande que puede permitirse?
6. En el artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, en *Water Environment Research*, 2002:177-188), se midieron las concentraciones de amoníaco (en mg/l) de gran número de pozos en Iowa. La media de la concentración fue de 0.71, la mediana fue de 0.22 y la desviación estándar de 1.09. ¿Sería posible determinar si estas concentraciones tienen una distribución aproximadamente normal? Si es así, diga si éstas se distribuyen normalmente, y explique cómo lo sabe. Si no, describa la información adicional que necesitaría para determinar si se distribuyen normalmente.
7. El medicamento utilizado para tratar cierta condición se administra mediante una inyección. La dosis objetivo en una aplicación particular es μ . Debido a las variaciones de la jeringa, en la escala de lectura, y en la mezcla de la suspensión del fluido, la dosis realmente administrada se distribuye con media μ y varianza σ^2 .
- a) ¿Cuál es la probabilidad de que la dosis administrada difiera de la media μ por menos de σ ?
 - b) Si X representa la dosis administrada, encuentre el valor de z de tal forma que $P(X < \mu < z\sigma) = 0.90$.
 - c) Si la media de la dosis es de 10 mg, la varianza es de 2.6 mg^2 y una sobredosis clínica se define como una dosis mayor a 15 mg, ¿cuál es la probabilidad de que un paciente reciba una sobredosis?
8. Usted tiene una caja grande de resistores cuyas resistencias se distribuyen normalmente con media de 10Ω y desviación estándar de 1Ω .
- a) ¿Qué proporción de los resistores tienen resistencias entre 9.3 y 10.7Ω ?
 - b) Si usted extrae una muestra de 100 resistores, ¿cuál es la probabilidad de que 50 o más de ellos tengan resistencias entre 9.3 y 10.7Ω ?
 - c) ¿Cuántos resistores debe extraer de la muestra de tal forma que la probabilidad sea de 0.99 de que 50 o más de los resistores de la muestra tengan resistencias entre 9.3 y 10.7Ω ?
9. La holgura de las válvulas de entrada de unos motores nuevos de cierto tipo se distribuye normalmente con media $200 \mu\text{m}$ y desviación estándar $10 \mu\text{m}$.
- a) ¿Cuál es la probabilidad de que la holgura sea mayor a $215 \mu\text{m}$?
 - b) ¿Cuál es la probabilidad de que la holgura esté entre 180 y $205 \mu\text{m}$?
 - c) Un motor tiene seis válvulas de entrada. ¿Cuál es la probabilidad de que sólo dos de ellas tengan holguras mayores a $215 \mu\text{m}$?
10. La rigidez de cierto tipo de viga de acero utilizada en la construcción de edificios tiene media de 30 kN/mm y desviación estándar de 2 kN/mm .
- a) ¿Es posible calcular la probabilidad de que la rigidez de una viga seleccionada aleatoriamente sea mayor que 32 kN/mm ? Si es así, calcule la probabilidad. Si no, explique por qué.
 - b) En una muestra de 100 vigas, ¿sería posible calcular la probabilidad de que la rigidez media muestral de vigas sea mayor a 30.2 kN/mm ? Si es así, calcule la probabilidad. Si no, explique por qué.
11. En cierto proceso, la probabilidad de producir un montaje de mayor tamaño es de 0.05.
- a) En una muestra de 300 montajes elegidos aleatoriamente, ¿cuál es la probabilidad de que menos de 20 tengan un tamaño mayor?
 - b) En una muestra de diez montajes elegidos aleatoriamente, ¿cuál es la probabilidad de que uno o más de ellos tengan tamaño mayor?
 - c) ¿A qué valor debe reducirse la probabilidad de un montaje de mayor tamaño para que sólo 1% de la población de 300 montajes tengan 20 o más con mayor tamaño?
12. Un proceso de producir placas de vidrio deja en promedio tres burbujas pequeñas por cada 10 m^2 de vidrio. El número de burbujas en una lámina de vidrio sigue una distribución de Poisson.
- a) ¿Cuál es la probabilidad de que una pieza de vidrio de $3 \times 5 \text{ m}$ contendrá más de dos burbujas?
 - b) ¿Cuál es la probabilidad de que una pieza de vidrio de $4 \times 6 \text{ m}$ no tendrá ninguna burbuja?

- c) ¿Cuál es la probabilidad de que 50 piezas de vidrio, cada una de 3×6 m contendrá más de 300 burbujas?
13. Células de levadura se encuentran suspendidas en un medio líquido. Se toma de la suspensión una muestra de 2 mL. Se cuentan 56 células de levadura.
- a) Estime la concentración de células de levadura por mL de suspensión y determine la incertidumbre.
- b) ¿Qué volumen de la suspensión debe extraerse para reducir la incertidumbre a una célula por mL?
14. Una placa está sujeta a su asiento utilizando diez pernos. Se revisa cada perno antes de la instalación y la probabilidad de pasar la inspección es de 0.9. Sólo se instalan los pernos que pasan la inspección. Sea X el número de pernos que son revisados para sujetar una placa.
- a) Determine $P(X = 12)$.
- b) Determine μ_X .
- c) Determine σ_X .
15. Los espesores de cuñas se distribuyen normalmente con una media de 1.5 mm y una desviación estándar de 0.2 mm. Se apilan tres cuñas, una sobre otra.
- a) Determine la probabilidad de que una pila tenga un espesor de más de 5 mm.
- b) Determine el 80o. percentil del espesor de la pila.
- c) ¿Cuál es el número mínimo de cuñas que se debe apilar para que la probabilidad de que la pila tenga un espesor mayor a 5 mm sea de al menos 0.99?
16. El tiempo de vida de un microprocesador se distribuye exponencialmente con una media de 3 000 horas.
- a) ¿Qué proporción de microprocesadores fallará dentro de 3 000 horas?
- b) ¿Qué proporción de microprocesadores funcionará durante más de 6 000 horas?
- c) Se instala un microprocesador al lado de otro que ha funcionado durante 1 000 horas. Suponga que los dos microprocesadores funcionan de manera independiente. ¿Cuál es la probabilidad de que uno nuevo falle antes que el viejo?
17. La duración de un cojinete (en años) sigue una distribución de Weibull con parámetros $\alpha = 1.5$ y $\beta = 0.8$.
- a) ¿Cuál es la probabilidad de que un cojinete dure más de un año?
- b) ¿Cuál es la probabilidad de que un cojinete dure menos de dos años?
18. El tiempo para hacer un cambio de aceite en cierta tienda se distribuye normalmente con media de 29.5 minutos y desviación estándar de tres minutos. ¿Cuál es la probabilidad de que un mecánico pueda realizar 16 cambios de aceite en una jornada de ocho horas al día?
19. Un topógrafo estimará la altura de un acantilado al realizar n mediciones independientes y promediándolas. Cada medición no está sesgada y tiene una desviación estándar $\sigma = 1$ m. ¿Cuántas mediciones se debe hacer para que la probabilidad sea de 0.95 de que el promedio esté a ± 0.25 del valor verdadero?
20. Se revisa que las latas de aluminio recibidas por una compañía de bebidas satisfagan las especificaciones de resistencia. De un envío muy grande, se selecciona aleatoriamente y se prueba 100. El envío será rechazado si 12 latas o más no pasan la prueba. Suponga que 20% de las latas del envío no satisface la especificación. ¿Cuál es la probabilidad de que se rechace el envío?
21. Un productor de cereales afirma que el peso bruto (incluyendo el empaque) de una caja de cereal etiquetada con peso de 12 onzas tiene media de 12.2 onzas y desviación estándar de 0.1 onzas. Usted junta 75 cajas y las pesa todas juntas. Sea S el peso total de las 75 cajas de cereal.
- a) Si la afirmación es cierta, ¿a qué es igual $P(S \leq 914.8)$?
- b) Con base en la respuesta al inciso (a), si la afirmación es cierta, ¿914.8 onzas es un peso total inusualmente pequeño para una muestra de 75 cajas?
- c) Si el peso total de las cajas fuera de 914.8 onzas, ¿estaría usted convencido de que la afirmación es falsa? Explique.
- d) Si la afirmación es cierta, ¿a qué es igual $P(S \leq 910.3)$?
- e) Con base en la respuesta al inciso (d), si la afirmación es cierta, ¿910.3 onzas es un peso total inusualmente pequeño para una muestra de 75 cajas?
- f) Si el peso total de las cajas fuera de 910.3 onzas, ¿estaría usted convencido de que la afirmación es falsa? Explique.
22. Una persona afirma que el número de visitas a su sitio web tiene una distribución de Poisson con media de 20 por hora. Sea X el número de visitas en cinco horas.
- a) Si la afirmación es cierta, ¿a qué es igual $P(X \leq 95)$?

- b) Con base en la respuesta al inciso a), si la afirmación es cierta, ¿95 visitas durante un periodo de cinco horas es una cifra inusualmente pequeña?
- c) Si usted observó 95 visitas en un periodo de cinco horas, ¿esto sería una evidencia de que la afirmación es falsa? Explique.
- d) Si la afirmación es cierta, ¿a qué es igual $P(X \leq 65)$?
- e) Con base en la respuesta al inciso d), si la afirmación es cierta, ¿65 visitas durante un periodo de cinco horas es una cifra inusualmente pequeña?
- f) Si usted observó 65 visitas en un periodo de cinco horas, ¿esto sería una evidencia de que la afirmación es falsa? Explique.

23. $X \sim \text{Geom}(p)$. Sea $s \geq 0$ un entero.

- a) Demuestre que $P(X > s) = (1 - p)^s$. (*Sugerencia:* La probabilidad de que se necesite más de s experimentos para obtener el primer éxito es igual a la probabilidad de que todos los primeros s experimentos resulten en fracaso.)
- b) Sea $t \geq 0$ un entero. Demuestre que $P(X > s + t | X > s) = P(X > t)$. A esto último se le denomina *propiedad de falta de memoria*. [*Sugerencia:* $P(X > s + t \text{ y } X > s) = P(X > s + t)$.]
- c) Existe dos monedas, de uno y cinco centavos, respectivamente. Se lanza al aire tres veces la primera moneda y todas las veces sale “cruz”. Enseguida, ambas monedas se lanzan al aire dos veces, de tal forma que la moneda de un centavo se lance un total de cinco veces y la de cinco centavos dos. Utilice la propiedad de falta de

memoria para calcular la probabilidad condicional de que en los cinco lanzamientos de la moneda de un centavo salga “cruz”, en virtud de que en los primeros tres lanzamientos salió “cruz”. Después, calcule la probabilidad de que en los dos lanzamientos de la moneda de cinco centavos salga “cruz”. ¿Las dos probabilidades son iguales?

24. Sea $X \sim \text{Bin}(n, p)$.

- a) Demuestre que si x es un entero positivo entre, e incluyendo a, 1 y n , entonces

$$\frac{P(X = x)}{P(X = x - 1)} = \left(\frac{n - x + 1}{x} \right) \left(\frac{p}{1 - p} \right)$$

- b) Demuestre que si $X \sim \text{Bin}(n, p)$, el valor más probable de X es el entero más grande, menor o igual a $np + p$. [*Sugerencia:* Utilice el inciso a) para demostrar que $P(X = x) \geq P(X = x - 1)$ si y sólo si $x \leq np + p$.]

25. Sea $X \sim \text{Poisson}(\lambda)$.

- a) Demuestre que si x es un entero positivo, entonces

$$\frac{P(X = x)}{P(X = x - 1)} = \frac{\lambda}{x}$$

- b) Demuestre que si $X \sim \text{Poisson}(\lambda)$, el valor más probable de X es el entero más grande, menor o igual a λ . [*Sugerencia:* Utilice el inciso (a) para demostrar que $P(X = x) \geq P(X = x - 1)$ si y sólo si $x \leq \lambda$.]

Capítulo 5

Intervalos de confianza

Introducción

En el capítulo 4 se analizaron estimaciones de diferentes parámetros; por ejemplo, $\hat{p}$ representa la estimación de una probabilidad de éxito p , y $\bar{X}$ lo es de una media poblacional μ . Ambas son llamadas **estimaciones puntuales**, porque son sólo números, o puntos. Un hecho importante de las estimaciones puntuales es que casi nunca son exactamente iguales a los valores reales que están estimando.

Difieren casi siempre, a veces por poco y otras por mucho. Con la finalidad de que una estimación sea útil, se necesita describir qué tan alejada está del valor verdadero. Una manera de lograr esto último es reportar una estimación de la desviación estándar, o de la incertidumbre. En este capítulo se mostrará que cuando la estimación tiene una distribución normal, se puede obtener más información acerca de su precisión cuando se calcula un intervalo de confianza. El siguiente ejemplo presenta la idea básica.

Suponga que se hace gran número de mediciones independientes, todas mediante el mismo procedimiento, del diámetro de un pistón. La media muestral de las mediciones es 14.0 cm, y la incertidumbre en esta cantidad, que representa la desviación estándar de la media muestral, es 0.1 cm. Suponga que las mediciones no están sesgadas. El valor 14.0 proviene de una distribución normal, ya que es el promedio de un importante número de mediciones. Ahora el diámetro verdadero del pistón no será exactamente igual a la media muestral de 14.0 cm. Sin embargo, dado que ésta proviene de una distribución normal, se puede utilizar dicha desviación estándar para determinar qué tan cerca está probablemente del diámetro verdadero. Por ejemplo, es muy improbable que la media muestral sea diferente del diámetro verdadero en más de tres desviaciones estándares. Por tanto, se tiene una enorme confianza de que el diámetro verdadero esté en el intervalo (13.7, 14.3). Por otro lado, es muy probable que la media muestral difiera del valor verdadero en más de una desviación estándar. Por tanto, se tiene que existe poca certeza de que el diámetro verdadero se encuentre en dicha cercanía (13.9, 14.1).

Los intervalos (13.7, 14.3) y (13.9, 14.1) son **intervalos de confianza** para el diámetro verdadero del pistón. En este capítulo se verá cómo calcular una medida cuantitativa del ni-

vel de la confianza que se puede tener en estos intervalos, así como en otros que se pueda construir. Específicamente, los resultados de la sección 5.1 mostrarán que se puede tener una confianza de 99.7% de que el diámetro verdadero del pistón se encuentre en el intervalo (13.7, 14.3), pero sólo certeza de 68% de que ese valor verdadero lo esté en (13.9, 14.1).

5.1 Intervalos de confianza para la media poblacional con muestras grandes

Se comienza con un ejemplo. Un ingeniero que supervisa el control de calidad quiere calcular la media del peso de cajas que se han llenado con cereal por una máquina específica durante cierto día. Toma una muestra aleatoria de 100 cajas que se han llenado con esa máquina en ese día. Calcula que la media muestral del peso de llenado es de $\bar{X} = 12.05$ oz y la desviación estándar $s = 0.1$ oz.

Debido a que la media poblacional no será exactamente igual a la media muestral de 12.05, es mejor construir un **intervalo de confianza** alrededor de 12.05 que quizá contenga a aquélla. Después se puede cuantificar el nivel de confianza de que la media poblacional esté realmente contenida por el intervalo. Con el propósito de ver cómo construir un intervalo de confianza en este ejemplo, sea μ la media poblacional desconocida y σ^2 la varianza respectiva. Sean $X_1, \dots, X_{100}$ los 100 pesos del llenado de las cajas muestreadas. El valor observado de la media muestral es $\bar{X} = 12.05$. Ya que $\bar{X}$ es la media de una muestra grande, el teorema del límite central especifica que proviene de una distribución normal cuya media es μ y cuya desviación estándar es $\sigma_{\bar{X}} = \sigma/\sqrt{100}$.

La figura 5.1 presenta una curva normal, que representa la distribución de $\bar{X}$. Aquí se indica que 95% intermedio de la curva se extiende una distancia $1.96 \sigma_{\bar{X}}$ a cada lado de la media poblacional. El valor observado $\bar{X} = 12.05$ constituye una sola muestra de esta distribución. No se tiene manera alguna de saber de qué parte de la curva fue extraído este valor especial de $\bar{X}$. La figura 5.1 presenta una posibilidad: que la media muestral esté dentro del

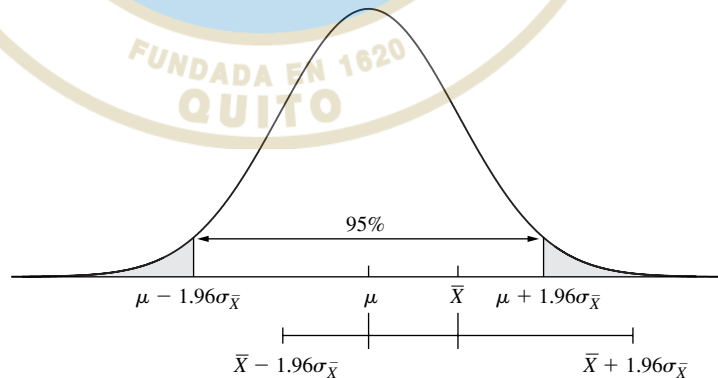


FIGURA 5.1 La media muestral $\bar{X}$ se extrae de una distribución normal con media μ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Para esta muestra en particular, $\bar{X}$ proviene de 95% intermedio de la distribución, por lo que el intervalo de confianza 95% $\bar{X} \pm 1.96\sigma_{\bar{X}}$ contiene con seguridad la media poblacional μ .

95% intermedio de la distribución. Noventa y cinco por ciento de todas las muestras que se podía haber tomado están dentro de esta categoría. La recta horizontal debajo de la curva en la figura 5.1 significa un intervalo alrededor de $\bar{X}$ que tiene exactamente la misma longitud de la parte central de 95% de la distribución; es decir, el intervalo $\bar{X} \pm 1.96 \sigma_{\bar{X}}$. Este **intervalo de confianza es de 95%** para la media poblacional μ . Es obvio que aquél contiene la media poblacional μ .

Por otra parte, la figura 5.2 representa una muestra cuya media $\bar{X}$ está fuera de 95% intermedio de la curva. Sólo 5% de todas las muestras se encuentra en dicha categoría. Para estas muestras más inusuales el intervalo de confianza de 95% $\bar{X} \pm 1.96 \sigma_{\bar{X}}$ no contiene la media poblacional μ .

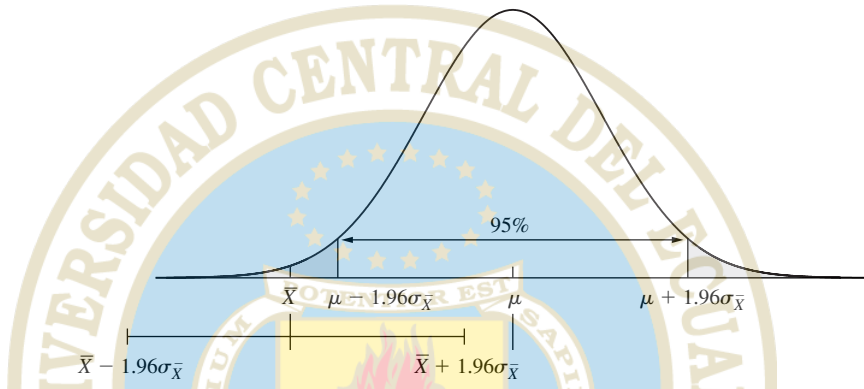


FIGURA 5.2 La media muestral $\bar{X}$ viene de una distribución normal con media μ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Para esta muestra en particular, $\bar{X}$ proviene de 5% exterior de la distribución, por lo que el intervalo de confianza de 95% $\bar{X} \pm 1.96 \sigma_{\bar{X}}$ no contiene la media poblacional μ .

Ahora se calculará un intervalo de confianza de 95% $\bar{X} \pm 1.96 \sigma_{\bar{X}}$ para la media del peso de llenado. El valor de $\bar{X}$ es 12.05. La desviación estándar poblacional σ y, por tanto, $\sigma_{\bar{X}} = \sigma/\sqrt{100}$ son desconocidos. Sin embargo, en este ejemplo, debido a que el tamaño muestral es grande, se podría aproximar σ con la desviación estándar muestral $s = 0.1$. Por tanto, se calcula al intervalo de confianza de 95% para la media del peso de llenado μ como $12.05 \pm (1.96)(0.01)$, o (12.0304, 12.0696). Se puede decir que hay **95% de confianza**, o **un nivel de confianza de 95%**; que la media del peso de llenado esté entre 12.0304 y 12.0696.

¿Este intervalo de confianza de 95% realmente contiene la media poblacional μ ? Esto último depende de si esta muestra en particular ocurrió en otra cuya media proviene de 95% intermedio de la distribución, o si era una muestra cuya media era inusualmente grande o pequeña, en el 5% exterior de la distribución. No hay ninguna manera de saber con seguridad en qué categoría está contenida esa muestra particular. Pero imagine que el ingeniero repitió este procedimiento todos los días, extrayendo una muestra grande y calculando el intervalo de confianza de 95% $\bar{X} \pm 1.96 \sigma_{\bar{X}}$. A la larga, 95% de las muestras que extrae tendrán medias en el 95% intermedio de la distribución, por lo que en tal porcentaje de los intervalos de confianza que el ingeniero calcula estará contenida la media poblacional. En otras palabras, un intervalo de confianza de 95% se calcula mediante un procedimiento que con seguridad contiene a la media poblacional 95% de las veces.

Es posible utilizar este mismo razonamiento para calcular intervalos de confianza con diferentes niveles. Por ejemplo, se puede construir un intervalo de confianza de 68%, como se muestra a continuación. Se sabe que 68% intermedio de la curva normal corresponde al intervalo que se extiende una distancia $1.0 \sigma_{\bar{X}}$ a cualquier lado de la media poblacional μ . Por consecuencia, un intervalo del mismo largo alrededor de $\bar{X}$, específicamente $\bar{X} \pm \sigma_{\bar{X}}$, contendrá la media poblacional en 68% de las muestras que se pudiera extraer. Por tanto, un intervalo de confianza de 68% para la media del peso de llenado de las cajas es $12.05 \pm (1.0)(0.01)$, o $(12.04, 12.06)$.

Observe que el intervalo de confianza de 95% es más ancho que el de 68%. Esto es intuitivamente creíble. Con el fin de aumentar la confianza que se tiene de contener a la verdadera media poblacional, se debe hacer más ancho al intervalo, que proporcionar un margen más ancho de error. Tomando dos casos extremos, se tiene una confianza de 100% de que la verdadera media poblacional está en un intervalo infinitamente ancho $(-\infty, \infty)$, y una confianza de 0% de que la verdadera media poblacional esté en el intervalo de ancho cero $[12.05, 12.05]$ que contenga a la media muestral y no a otro punto.

A continuación se mostrará cómo encontrar un intervalo de confianza con cualquier nivel de confianza deseado. Sea α un número entre 0 y 1, y $100(1 - \alpha)\%$ el nivel de confianza requerido. La figura 5.3 muestra la curva normal que representa la distribución de $\bar{X}$. Se define a $z_{\alpha/2}$ como el puntaje z que corta un área de $\alpha/2$ en la cola del lado derecho. Por ejemplo, la tabla z (tabla A.2) indica que $z_{0.025} = 1.96$, ya que 2.5% del área bajo la curva normal estandarizada está a la derecha de 1.96. De manera similar, la cantidad $-z_{\alpha/2}$ corta un área de $\alpha/2$ en la cola del lado izquierdo. El área $1 - \alpha$ intermedia bajo la curva corresponde al intervalo $\mu \pm z_{\alpha/2}\sigma_{\bar{X}}$. Como consecuencia del razonamiento que se muestra en las figuras 5.1 y 5.2, el intervalo $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$ contendrá a la media poblacional μ para una proporción $1 - \alpha$ de todas las muestras que se pudieran extraer. Por tanto, un **intervalo de confianza de nivel $100(1 - \alpha)\%$** para μ es $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$, o $\bar{X} \pm z_{\alpha/2}\sigma/\sqrt{n}$.

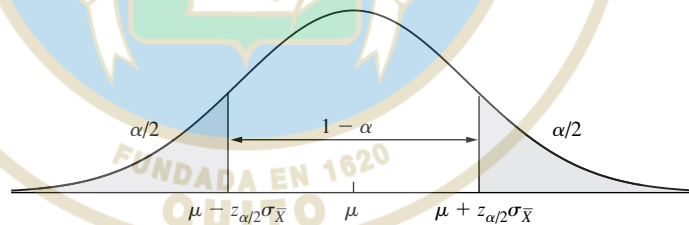


FIGURA 5.3 La media muestral $\bar{X}$ se extrae de una distribución normal con media μ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. La cantidad $z_{\alpha/2}$ constituye el puntaje z que corta un área de $\alpha/2$ en la cola del lado derecho. Asimismo, $-z_{\alpha/2}$ representa el que corta un área de $\alpha/2$ la cola del lado izquierdo. El intervalo $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$ contendrá la media poblacional μ para una proporción $1 - \alpha$ de todas las muestras que se pudiera extraer. Por tanto, $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$ significa un intervalo de confianza de nivel $100(1 - \alpha)\%$ para μ .

Nótese que aun para muestras grandes, la distribución de $\bar{X}$ es sólo *aproximadamente* normal, y no exactamente normal. Por tanto, los niveles establecidos para los intervalos de confianza son aproximados. Cuando el tamaño muestral es lo suficientemente grande para que se utilice el teorema del límite central, la distinción entre los niveles aproximados y exactos se ignora en la práctica.

Resumen

Sea $X_1, \dots, X_n$ una muestra aleatoria *grande* ($n > 30$) de una población con media μ y desviación estándar σ , por lo que $\bar{X}$ es aproximadamente normal. Entonces un intervalo de confianza $100(1 - \alpha)\%$ para μ es

$$\bar{X} \pm z_{\alpha/2} \sigma_{\bar{X}} \quad (5.1)$$

donde $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Cuando el valor de σ es desconocido, se puede sustituir por la desviación estándar muestral s .

En particular,

- $\bar{X} \pm \frac{s}{\sqrt{n}}$ es intervalo de confianza de 68% para μ .
- $\bar{X} \pm 1.645 \frac{s}{\sqrt{n}}$ es intervalo de confianza de 90% para μ .
- $\bar{X} \pm 1.96 \frac{s}{\sqrt{n}}$ es intervalo de confianza de 95% para μ .
- $\bar{X} \pm 2.58 \frac{s}{\sqrt{n}}$ es intervalo de confianza de 99% para μ .
- $\bar{X} \pm 3 \frac{s}{\sqrt{n}}$ es intervalo de confianza de 99.7% para μ .

Ejemplo

5.1

La media y desviación estándar muestrales para todos los pesos de llenado de las 100 cajas son $\bar{X} = 12.05$ y $s = 0.1$. Encuentre un intervalo de confianza de 85% para la media de los pesos de llenado de las cajas.

Solución

Con el propósito de determinar un intervalo de confianza de 85%, haga $1 - \alpha = 0.85$ para obtener $\alpha = 0.15$ y $\alpha/2 = 0.075$. Cuando se busca en la tabla a $z_{0.075}$, el puntaje z que corta 7.5% del área en la cola del lado derecho. Se encuentra $z_{0.075} = 1.44$. Se aproxima $\sigma_{\bar{X}} \approx s/\sqrt{n} = 0.01$. Por lo que el intervalo de confianza de 85% es $12.05 \pm (1.44)(0.01)$. Esto último se puede escribir como 12.05 ± 0.0144 , o como (12.0356, 12.0644).

Ejemplo

5.2

El artículo “Study on the Life Distribution of Microdrills” (Z. Yang, Y. Chen y Y. Yang, en *Journal of Engineering Manufacture*, 2002:301-305) notifica que en una muestra de 50 microperforadoras, éstas perforan una aleación de acero con bajo contenido de carbono, el tiempo de vida promedio (expresado como el número de huecos perforados antes de que falle) era de 12.68 con desviación estándar de 6.83. Determine un intervalo de confianza de 95% para la media del tiempo de vida de las microperforadoras bajo estas condiciones.

Solución

Primero se traslada el problema al lenguaje estadístico. Se tiene una muestra aleatoria simple $X_1, \dots, X_{50}$ de los tiempos de vida. La media y desviación estándar muestrales son $\bar{X} = 12.68$ y $s = 6.83$. La media poblacional es desconocida y se denota por μ .

El intervalo de confianza tiene la forma $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$, como se especifica en la expresión (5.1). Dado que se quiere un intervalo de confianza de 95%, el nivel respectivo $1 - \alpha$ es igual a 0.95. Por lo que $\alpha = 0.05$ y $z_{\alpha/2} = z_{0.025} = 1.96$. Se aproxima σ con $s = 6.83$ y se obtiene $\sigma_{\bar{X}} \approx 6.83/\sqrt{50} = 0.9659$. Por lo que el intervalo de confianza de 95% es $12.68 \pm (1.96)(0.9659)$. Lo anterior se puede escribir como 12.68 ± 1.89 , o como (10.79, 14.57).

El siguiente resultado de computadora (de MINITAB) presenta el intervalo de confianza de 95% calculado en el ejemplo 5.2.

```
One-Sample Z
The assumed standard deviation = 6.830000
N      Mean    SE Mean      95% CI
50    12.680000 0.965908  (10.786821, 14.573179)
```

La mayor parte del resultado se explica solo. La cantidad etiquetada con “SE Mean” representa la desviación estándar de la media muestral $\sigma_{\bar{X}}$, aproximada por $s/\sqrt{n}$. (“SE Mean” establece el error estándar de la media, que es otro término para la desviación estándar de la media muestral.)

Ejemplo**5.3**

En el ejemplo 5.2 determine un intervalo de confianza de 80 por ciento.

Solución

Para determinar un intervalo de confianza de 80%, haga $1 - \alpha = 0.80$ para obtener $\alpha = 0.20$. Después busque en la tabla para $z_{.10}$, el puntaje z que corta 10% del área en la cola del lado derecho. El valor es $z_{.10} = 1.28$. Por lo que el intervalo de confianza de 80% es $12.68 \pm (1.28)(0.9659)$. Éste se puede escribir como 12.68 ± 1.24 , o bien (11.44, 13.92).

Se ha visto cómo calcular un intervalo de confianza con un nivel de confianza específico. Es posible calcular el nivel de un intervalo de confianza dado. El ejemplo 5.4 ilustra el método apropiado.

Ejemplo**5.4**

Con base en los datos del tiempo de vida de las microperforadoras que se presentó en el ejemplo 5.2, un ingeniero notifica un intervalo de confianza de (11.09, 14.27), pero olvidó especificar el nivel. ¿Cuál es el nivel de confianza de este intervalo de confianza?

Solución

El intervalo de confianza tiene la forma $\bar{X} \pm z_{\alpha/2}s/\sqrt{n}$. Se despeja a $z_{\alpha/2}$ y después se consulta la tabla z para determinar el valor de α . Ahora $\bar{X} = 12.68$, $s = 6.83$ y $n = 50$. El límite supe-

rior de confianza de 14.27 satisface la ecuación $14.27 = 12.68 + z_{\alpha/2}(6.83/\sqrt{50})$. Por tanto, $z_{\alpha/2} = 1.646$. De la tabla z se determina que $\alpha/2$, el área a la derecha de 1.646, es aproximadamente 0.05. El nivel es $100(1 - \alpha)\%$, o 90 por ciento.

Más acerca de niveles de confianza

El nivel de confianza de un intervalo mide la confiabilidad del método utilizado para calcular el intervalo. Un intervalo de confianza de un nivel $100(1 - \alpha)\%$ se calcula mediante un método que a la larga dará como resultado que la media poblacional se sitúe en una proporción $1 - \alpha$ todas las veces que se utilice. En la práctica, cuando se calcula un intervalo de confianza, se debe decidir qué nivel de confianza se utilizará. Esta decisión implica un intercambio, porque los intervalos con niveles de confianza más grandes son menos precisos. Por ejemplo, un intervalo de confianza de 68% especifica la media poblacional dentro de $\pm 1.0\sigma_{\bar{X}}$, mientras que un intervalo de confianza de 95% especifica a éste solamente dentro de $\pm 1.96\sigma_{\bar{X}}$; por tanto, tiene sólo casi la mitad de la precisión del intervalo de confianza de 68%. La figura 5.4 ilustra el intercambio entre confianza y precisión. Se extrajeron cien muestras de una población con media μ . La figura 5.4b presenta cien intervalos de confianza de 95%, cada uno basado en una de estas muestras. Los intervalos de confianza son todos diferentes, porque cada muestra tiene una media $\bar{X}$ diferente. (También tienen valores diferentes de s con los que se aproxima a σ , pero esto último tiene un efecto mucho muy pequeño.) Cerca de 95% de estos intervalos contiene la media poblacional μ . La figura 5.4a presenta intervalos de confianza de 68% basados en las mismas muestras. Estos intervalos son más precisos (más angostos), pero muchos de ellos no contienen la media poblacional. La figura 5.4c presenta intervalos de confianza de 99.7%. Estos intervalos son muy confiables. A la larga, solamente tres de los 1 000 intervalos no contendrán la media poblacional. Sin embargo, son menos precisos (más anchos); por tanto, no transmiten mucha información.

El nivel de confianza más utilizado en la práctica es de 95%. Para muchas aplicaciones, este nivel proporciona un buen compromiso entre precisión y confiabilidad. Los niveles de confianza inferiores a 90% rara vez se utilizan. Para algunas aplicaciones de aseguramiento de calidad, donde la confiabilidad de producto es importante, se utilizan intervalos con niveles de confianza muy altos, de 99.7 por ciento.

Probabilidad contra confianza

En el ejemplo del peso de llenado analizado al inicio de esta sección, se calculó un intervalo de confianza de 95% para la media poblacional μ de (12.304, 12.696). Es arriesgado decir que la probabilidad es de 95% y que μ está entre 12.304 y 12.696. Sin embargo, esto último no es correcto. El término *probabilidad* se refiere a los eventos aleatorios que pueden resultar diferentes cuando se repiten los experimentos. Los números 12.304 y 12.696 son fijos, no aleatorios. La media poblacional es también fija. La media del peso de llenado está ya sea en el intervalo de 12.304 a 12.696, o no lo está. No hay aleatoriedad implicada. Por tanto, se dice que se tiene *confianza* de 95% (no una probabilidad) que la media poblacional esté en tal intervalo.

Por otra parte, se dice que se está analizando un *método* utilizado para calcular un intervalo de confianza de 95%. El método dará como resultado que la media poblacional esté 95% de las veces, y no el otro 5%. En este caso, si la media poblacional está contenida o no es un evento aleatorio, porque puede variar entre experimentos. Por tanto, es correcto decir

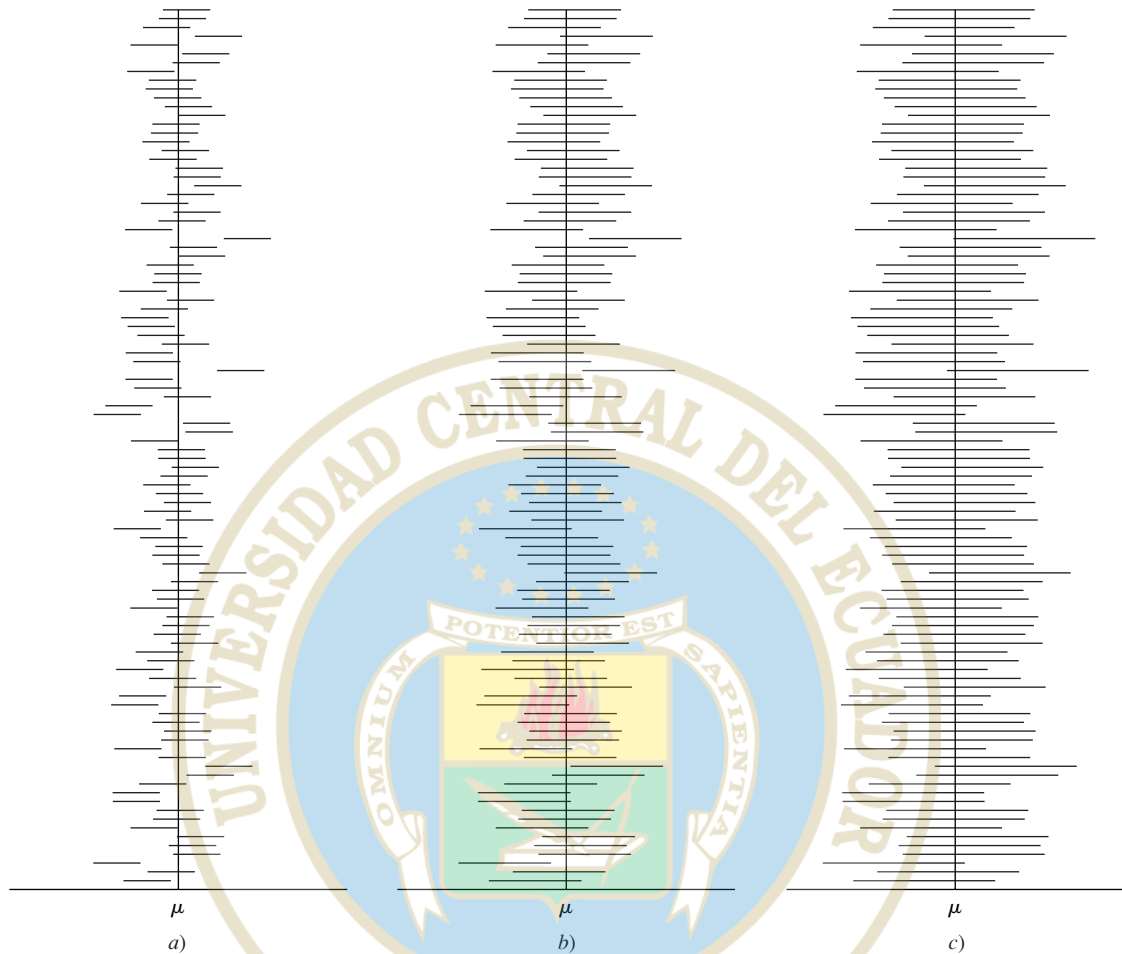


FIGURA 5.4 *a)* Cien intervalos de confianza de 68% para una media poblacional, cada uno calculado con una muestra diferente. Aunque precisos, no contienen a la media poblacional 32% de las veces. Esta alta tasa de fallas hace que el intervalo de confianza de 68% sea inaceptable para propósitos prácticos. *b)* Cien intervalos de confianza de 95% calculado de estas muestras. Éste presenta un buen compromiso entre precisión y confiabilidad para muchos propósitos. *c)* Cien intervalos de confianza de 99.7% calculado de estas muestras. Estos intervalos no contienen a la media poblacional solamente tres veces en 1 000. Son sumamente confiables, pero poco precisos.

que un *método* para calcular un intervalo de confianza de 95% tiene esa probabilidad de contener a la media poblacional.

Ejemplo



Un intervalo de confianza de 90% para la media del diámetro (en cm) de varillas de acero fabricadas en cierta máquina de extrusión se calcula de (14.73, 14.91). Verdadero o falso: La probabilidad de que la media del diámetro de las varillas fabricadas por este proceso esté entre 14.73 y 14.91 es de 90 por ciento.

Solución

Falso. Un intervalo específico de confianza está dado. La media está o no en el intervalo. Hay confianza de 90% de que la media poblacional esté entre 14.73 y 14.91. El término *probabilidad* no es adecuado.

Ejemplo

Una ingeniera planea calcular un intervalo de confianza de 90% para la media del diámetro de varillas de acero. Medirá los diámetros de una muestra grande de varillas, calculará $\bar{X}$ y s , y después el intervalo $\bar{X} \pm 1.645s/\sqrt{n}$. Verdadero o falso: La probabilidad de que la media del diámetro estará en este intervalo es de 90 por ciento.

Solución

Verdadero. Lo que se describe aquí es un método para calcular un intervalo de confianza, más que un valor numérico específico. Es correcto decir que un método para calcular un intervalo de confianza de 90% tiene probabilidad de 90% de contener la media poblacional.

Ejemplo**5.7**

Un equipo de geólogos planea medir los pesos de 250 rocas. Después de pesar cada roca muchas veces, calculará un intervalo de confianza de 95% para su peso. Suponga que no hay sesgo en el procedimiento de pesado. ¿Cuál es la probabilidad de que más de 240 de los intervalos de confianza contengan los pesos verdaderos de las rocas?

Solución

Aquí se han analizado 250 implementaciones planeadas de un método de cálculo de intervalos de confianza, no 250 intervalos específicos que ya han sido calculados. Por tanto, es adecuado calcular la probabilidad de que un número específico de estos intervalos contendrá los pesos verdaderos de sus rocas respectivas. Debido a que el procedimiento de pesado no tiene sesgos, el peso verdadero de una roca es igual a la media poblacional de sus mediciones. Se puede pensar en cada uno de los 250 intervalos de confianza como un ensayo de Bernoulli, con el éxito ocurriendo si el intervalo de confianza contiene la media poblacional. Como consecuencia de que un intervalo de confianza de 95% se calcula con un proceso que contiene la media poblacional 95% de las veces, la probabilidad de éxito para cada ensayo es de 0.95. Sea Y el número de intervalos de confianza que contiene al peso verdadero. Entonces $Y \sim \text{Bin}(250, 0.95) \approx N(237.5, 11.875)$. La desviación estándar de Y es $\sigma = \sqrt{11.875} = 3.45$. Con el uso de la curva normal, la probabilidad de que $Y > 240$ es 0.1922. Véase la figura 5.5. Observe que se ha utilizado la corrección de continuidad (véase la sección 4.10).

Determinación del tamaño muestral necesario para un intervalo de confianza de ancho específico

En el ejemplo 5.2, un intervalo de confianza de 95% fue dado por 12.68 ± 1.89 , o (10.79, 14.57). Éste especifica que la media está dentro de ± 1.89 . Ahora suponga que es demasiado ancho para ser útil. Suponga que es deseable obtener un intervalo de confianza de 95% que especifique que la media esté dentro de ± 0.50 . Con este propósito se debe aumentar el tama-

ño muestral. Se ilustra cómo calcular el tamaño necesario para obtener un intervalo de confianza de cualquier ancho específico.

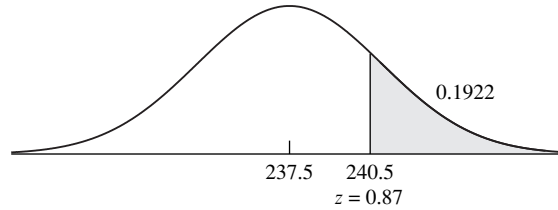


FIGURA 5.5 Solución al ejemplo 5.7.

Se sigue de la expresión (5.1) que el ancho de un intervalo de confianza para una media poblacional basado en una muestra de tamaño n extraída de una población con desviación estándar σ es $\pm z_{\alpha/2}\sigma/\sqrt{n}$. Si se especifica el nivel de confianza $100(1 - \alpha)\%$ se puede buscar el valor $z_{\alpha/2}$. Si la desviación estándar σ poblacional también se especifica, se calculará el valor de n necesario para producir un ancho específico. En el ejemplo 5.2, el nivel de confianza es de 95% y la desviación estándar se calcula de 6.83. Se busca $z_{\alpha/2} = z_{0.025} = 1.96$. El tamaño muestral necesario para obtener un intervalo de confianza de 95% con ancho ± 0.50 se encuentra al despejar n de la ecuación $(1.96)(6.83)/\sqrt{n} = 0.50$. Se obtiene $n = 716.83$, que se redondea hacia arriba a $n = 717$.

Ejemplo 5.8

En el ejemplo del peso de llenado que se analizó en esta sección, la desviación estándar muestral de pesos de las 100 cajas era de $s = 0.1$ oz. ¿Cuántas cajas se probarán para obtener un intervalo de confianza de 99% de ancho ± 0.012 oz?

Solución

El nivel es 99%, por lo que $1 - \alpha = 0.99$. Por tanto, $\alpha = 0.01$ y $z_{\alpha/2} = 2.58$. Se calcula el valor de σ con $s = 0.1$. El tamaño muestral necesario se encuentra con $(2.58)(0.1)/\sqrt{n} = 0.012$. Se obtiene $n \approx 463$.

Intervalos de confianza de un lado

Los intervalos de confianza que se han analizado son de dos lados, ya que especifican tanto un límite inferior como otro superior. Ocasionalmente se tiene interés sólo en uno de estos límites. En estos casos son adecuados los intervalos de confianza de un lado. Por ejemplo, suponga que un ingeniero que supervisa la confiabilidad quiere calcular la media de la fuerza de compresión de cierto tipo de bloque de concreto, con el propósito de determinar los tipos de aplicaciones para los que será adecuado. El ingeniero estará interesado solamente en un límite inferior para la fuerza, ya que las especificaciones para diferentes aplicaciones en general especificarán sólo una fuerza mínima.

Suponga que una muestra grande tiene una media muestral $\bar{X}$ y desviación estándar $\sigma_{\bar{X}}$. La figura 5.6 muestra cómo se puede adaptar la idea detrás del intervalo de confianza de dos lados para obtener un intervalo de confianza de un lado para la media poblacional μ . La cur-

va normal representa la distribución de $\bar{X}$. Para 95% de todas las muestras que se pueden extraer, $\bar{X} < \mu + 1.645\sigma_{\bar{X}}$ y, por tanto, el intervalo $(\bar{X} - 1.645\sigma_{\bar{X}}, \infty)$ contiene a μ . Este último no contendrá a μ sólo si la media muestral está en el 5% superior de su distribución. El intervalo $(\bar{X} - 1.645\sigma_{\bar{X}}, \infty)$ es un intervalo de confianza de un lado de 95% desigual para μ , y la cantidad $\bar{X} - 1.645\sigma_{\bar{X}}$ es un límite inferior de confianza del 95% para μ .

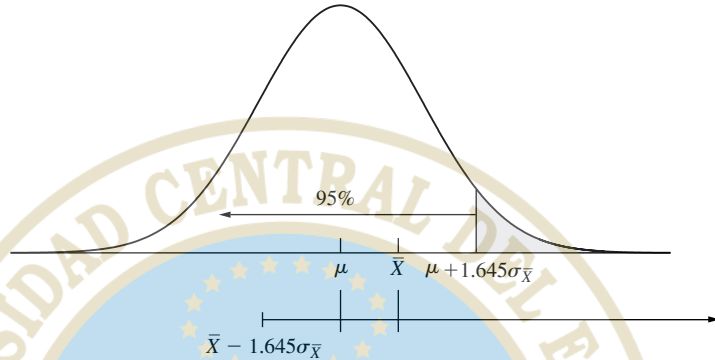


FIGURA 5.6 La media muestral $\bar{X}$ es extraída de una distribución normal con media μ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Para esta muestra en particular, $\bar{X}$ proviene de 95% más bajo de la distribución, por lo que el intervalo de confianza de un lado de 95% $(\bar{X} - 1.645\sigma_{\bar{X}}, \infty)$ seguramente contiene a la media poblacional μ .

Al construir una figura, como la 5.6, con 5% de la cola inferior sombreada, se puede ver que la cantidad $\bar{X} + 1.645\sigma_{\bar{X}}$ es el límite superior de confianza de 95% superior para μ . Ahora se generalizará el método para obtener intervalos de confianza de un lado a cualquier nivel deseado. Se define z_{α} como el puntaje z que corta un área α en la cola de la derecha de la curva normal. Por ejemplo, $z_{0.05} = 1.645$. Mediante el razonamiento que se utilizó para obtener un intervalo de confianza de 95%, se pueden apreciar un nivel $100(1 - \alpha)\%$ con límite inferior de confianza para μ dado por $\bar{X} - z_{\alpha}\sigma_{\bar{X}}$ y un nivel $1 - \alpha$ con límite superior de confianza para μ dado por $\bar{X} + z_{\alpha}\sigma_{\bar{X}}$.

Resumen

Sea $X_1, \dots, X_n$ una muestra aleatoria grande ($n > 30$) de una población con media μ y desviación estándar σ , se tiene que $\bar{X}$ es aproximadamente normal. Entonces el nivel de confianza $100(1 - \alpha)\%$ con límite inferior de confianza para μ es

$$\bar{X} - z_{\alpha}\sigma_{\bar{X}} \quad (5.2)$$

y un nivel $100(1 - \alpha)\%$ con límite de confianza superior para μ es

$$\bar{X} + z_{\alpha}\sigma_{\bar{X}} \quad (5.3)$$

donde $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Cuando el valor de σ es desconocido, se puede sustituir por la desviación estándar muestral s .

En particular,

- $\bar{X} + 1.28 \frac{s}{\sqrt{n}}$ representa un límite superior de confianza de 90% para μ .
- $\bar{X} + 1.645 \frac{s}{\sqrt{n}}$ constituye un límite superior de confianza de 95% para μ .
- $\bar{X} + 2.33 \frac{s}{\sqrt{n}}$ significa un límite superior de confianza de 99% para μ .

Los correspondientes límites inferiores se encuentran al reemplazar el “+” con el “-”.

Ejemplo

5.9

Con referencia al ejemplo 5.2, encuentre tanto un límite inferior de confianza de 95% como uno superior de 99% para la media del tiempo de vida de las microperforadoras.

Solución

La media muestral y la desviación estándar son $\bar{X} = 12.68$ y $s = 6.83$, respectivamente. El tamaño muestral es $n = 50$. Se calcula $\sigma_{\bar{X}} \approx s/\sqrt{n} = 0.9659$. El límite inferior de confianza de 95% es $\bar{X} - 1.645\sigma_{\bar{X}} = 11.09$ y el de 99% es $\bar{X} + 2.33\sigma_{\bar{X}} = 14.93$.

En el ejemplo 5.2, el intervalo de confianza de 95% de dos lados se calculó de (10.79, 14.57). El límite inferior de confianza de 95% de 11.09, calculado en el ejemplo 5.9, es más grande que el límite inferior del intervalo de confianza de dos lados. La razón de esto último consiste en que el intervalo de dos lados puede fallar en dos maneras: el valor de μ podría ser demasiado alto o demasiado bajo. El intervalo de confianza de 95% de dos lados está diseñado para fallar 2.5% de las veces en el lado superior y 2.5% en el inferior. En contraparte, el límite inferior de confianza de 95% nunca falla sobre el lado superior. Éste está, por tanto, diseñado para fallar 5% de las veces en el lado inferior, por lo que su límite inferior es mayor que para el intervalo de dos lados.

Intervalos de confianza que deben estar basados en muestras aleatorias

Los métodos descritos en esta sección requieren que los datos sean una muestra aleatoria de una población. Cuando se utiliza para otras muestras, los resultados podrían ser no significativos. Los siguientes son dos ejemplos en los que se incumple la suposición de muestreo aleatorio.

Ejemplo

5.10

Un ingeniero químico desea calcular la media de la producción de un nuevo proceso. El proceso está operando 100 veces durante un periodo de varios días. La figura 5.7 presenta las 100 producciones graficadas en función del tiempo. ¿Sería adecuado calcular un intervalo de confianza para la media de la producción mediante el cálculo de $\bar{X}$ y s para las producciones y después utilizar la expresión (5.1)?

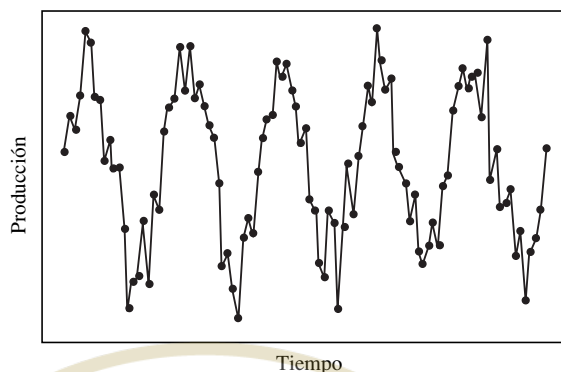


FIGURA 5.7 Producción de 100 operaciones de un proceso químico, graficadas en función del tiempo. Hay un patrón claro que indica que los datos no forman una muestra aleatoria.

Solución

No. La expresión (5.1) es válida sólo cuando los datos son una muestra aleatoria de una población. La figura 5.7 muestra un patrón cíclico. Éste podría indicar que la producción de cada operación está influida por la producción de la operación previa, lo que violaría la suposición de independencia. Otra posibilidad es que la producción esté influida por condiciones ambientales que fluctúan en forma regular. En cualesquiera de los dos casos, los datos no satisfacen las condiciones de una muestra aleatoria y no se debe utilizar la expresión (5.1).

Ejemplo

5.11

El ingeniero de quien se habló en el ejemplo 5.10 está investigando la producción de otro proceso. La figura 5.8 presenta las producciones de 100 operaciones de éste, graficadas en función del tiempo. ¿Se debe utilizar la expresión (5.1) para calcular un intervalo de confianza para la media de la producción de este proceso?

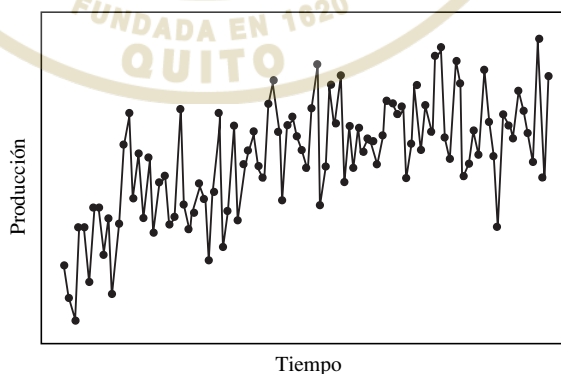


FIGURA 5.8 Producción de 100 operaciones de un proceso químico, graficadas en función del tiempo. Hay tendencia creciente con el tiempo, al menos en la parte inicial de la gráfica, lo que indica que los datos no forman una muestra aleatoria.

Solución

No. Como en el ejemplo 5.10, hay un patrón en el tiempo. En este caso, las producciones tienden a aumentar con el tiempo, al menos en la parte inicial de la gráfica. Esto último podría indicar un “efecto de aprendizaje”; conforme un operador se hace más experimentado respecto de un proceso, los resultados mejoran. Un análisis más minucioso de los datos podría indicar un momento donde el aumento parece parar y, en tal caso, la parte que tiene éxito se utiliza para formar un intervalo de confianza.

Ejercicios para la sección 5.1

- Determine el valor de $z_{\alpha/2}$ para utilizar la expresión (5.1) con el fin de construir un intervalo de confianza con nivel
 - 90%
 - 83%
 - 99.5%
 - 75%
- Determine los niveles de los intervalos de confianza que tienen los siguientes valores de $z_{\alpha/2}$:
 - $z_{\alpha/2} = 1.96$
 - $z_{\alpha/2} = 2.17$
 - $z_{\alpha/2} = 1.28$
 - $z_{\alpha/2} = 3.28$
- Conforme se eleva el nivel de confianza, la confiabilidad _____ y la precisión _____. Opciones: aumenta, disminuye.
- Los métodos de interpolación se utilizan para calcular alturas superiores al nivel del mar para ubicaciones donde las mediciones directas no están disponibles. En el artículo “Transformation of Ellipsoid Heights to Local Leveling Heights” (M. Yanalak y O. Baykal, en *Journal of Surveying Engineering*, 2001:90-103), se evalúa un método de interpolación para un polinomio de segundo orden que tiene como objetivo calcular las alturas de mediciones GPS (sistema de posicionamiento global). En una muestra de 74 ubicaciones, los errores del método tienen promedio de 3.8 cm, con desviación estándar de 4.8 cm.
 - Determine un intervalo de confianza de 95% para la media del error de este método.
 - Determine un intervalo de confianza de 98% para la media del error de dicho método.
 - Un topógrafo afirma que el error de media está entre 3.2 y 4.4 cm. ¿Con qué nivel de confianza se puede hacer esta afirmación?
 - ¿Aproximadamente cuántas ubicaciones se deben muestrear con el propósito de que un intervalo de confianza de 95% especificará la media dentro de ± 0.7 cm?
 - ¿Aproximadamente cuántas ubicaciones se debe muestrear con el propósito de que un intervalo de confianza de 98% especificará la media dentro de ± 0.7 cm?
- En una muestra aleatoria de 100 baterías producidas por cierto método, el promedio del tiempo de vida fue de 150 horas y la desviación estándar de 25 horas.
 - Determine un intervalo de confianza de 95% para la media del tiempo de vida de las baterías producidas por este método.
 - Determine un intervalo de confianza de 99% para la media del tiempo de vida de baterías producidas por dicho método.
 - Un ingeniero afirma que la media del tiempo de vida está entre 147 y 153 horas. ¿Con qué nivel de confianza se puede hacer esta afirmación?
 - ¿Aproximadamente cuántas baterías se deben muestrear con el propósito de que un intervalo de confianza de 95% especificará la media dentro de ± 2 horas?
 - ¿Aproximadamente cuántas baterías se deben muestrear con el fin de que un intervalo de confianza de 99% especificará la media dentro de ± 2 horas?
- En una muestra aleatoria de 53 especímenes de concreto, la media de la porosidad (en %) fue de 21.6 y la desviación estándar de 3.2.
 - Determine un intervalo de confianza de 90% para la media de la porosidad de los especímenes de este tipo de concreto.
 - Determine un intervalo de confianza de 95% para la media de la porosidad de los especímenes de este tipo de concreto.
 - ¿Cuál es el nivel de confianza del intervalo (21.0, 22.2)?

- d) ¿Cuántos especímenes se deben muestrear con el propósito de que un intervalo de confianza de 90% especifique la media dentro ± 0.3 ?
- e) ¿Cuántos especímenes se deben muestrear con el propósito de que un intervalo de confianza de 95% especifique la media dentro ± 0.3 ?
7. En una muestra de 80 clavos con costo de diez centavos, el peso promedio era 1.56 g y la desviación estándar era de 0.1 gramos.
- a) Determine un intervalo de confianza de 95% para la media del peso de este tipo de clavo.
- b) Determine un intervalo de confianza de 98% para la media del peso de este tipo de clavo.
- c) ¿Cuál es el nivel de confianza del intervalo (1.54, 1.58)?
- d) ¿Cuántos clavos se deben muestrear con el propósito de que un intervalo de confianza de 95% especifique la media dentro de ± 0.01 g?
- e) ¿Aproximadamente cuántos clavos se deben muestrear con el fin de que un intervalo de confianza de 98% especifique la media dentro de ± 0.01 g?
8. Una etapa en la fabricación de cierta abrazadera de metal implica perforar cuatro huecos. En una muestra de 150 abrazaderas, el promedio del tiempo necesario para completar dicha etapa era de 72 segundos y la desviación estándar de 10 segundos.
- a) Determine un intervalo de confianza de 95% para la media del tiempo necesario para completar tal etapa.
- b) Determine un intervalo de confianza de 99.5% para la media del tiempo necesario para completar esta etapa.
- c) ¿Cuál es el nivel de confianza del intervalo (71, 73)?
- d) ¿Cuántas abrazaderas se deben muestrear con el propósito de que un intervalo de confianza de 95% especifique la media dentro de ± 1.5 segundos?
- e) ¿Cuántas abrazaderas se deben muestrear con el objetivo de que un intervalo de confianza 99.5% especifique la media dentro de ± 1.5 segundos?
9. Un proveedor vende fibras sintéticas a una compañía de manufactura. Se selecciona una muestra aleatoria simple de 81 fibras de un envío. El promedio de la fuerza de ruptura de éstas es de 29 lb y la desviación estándar de 9 lb.
- a) Determine un intervalo de confianza de 95% para la media de la fuerza de ruptura de todas las fibras del envío.
- b) Determine un intervalo de confianza de 99% para la media de la fuerza de ruptura de todas las fibras del envío.
- c) ¿Cuál es el nivel de confianza del intervalo (27.5, 30.5)?
- d) ¿Cuántas fibras se deben muestrear con el propósito de que un intervalo de confianza de 95% especifique la media dentro de ± 1 lb?
- e) ¿Cuántas fibras se deben muestrear con el propósito de que un intervalo de confianza de 99% especifique la media dentro de ± 1 lb?
10. Con referencia al ejercicio 5.
- a) Determine un límite inferior de confianza de 95% para la media del tiempo de vida de tal tipo de batería.
- b) Un ingeniero afirma que la media del tiempo de vida es mayor de 148 horas. ¿Con qué nivel de confianza se puede hacer esta afirmación?
11. Con referencia al ejercicio 6.
- a) Determine un límite superior de confianza de 99% para la media de la porosidad.
- b) Se hizo una afirmación de que la media de la porosidad es menor que 22.7%. ¿Con qué nivel de confianza se pudo haber hecho tal afirmación?
12. Con referencia al ejercicio 7.
- a) Encuentre un límite superior de confianza de 90% para la media del peso.
- b) Alguien dice que la media del peso es menor que 1.585 g. ¿Con qué nivel de confianza se pudo haber hecho dicha afirmación?
13. Con referencia al ejercicio 8.
- a) Determine un límite inferior de confianza de 98% para completar la etapa.
- b) Un especialista en eficiencia dice que la media del tiempo es mayor de 70 segundos. ¿Con qué nivel de confianza se pudo haber hecho esta afirmación?
14. Con referencia al ejercicio 9.
- a) Determine un límite superior de confianza de 95% para la media de la fuerza de ruptura.
- b) El proveedor afirma que la media de la fuerza de ruptura es mayor que 28 lb. ¿Con qué nivel de confianza se pudo haber hecho tal afirmación?
15. Una investigadora calcula un intervalo de confianza de 95% para una media poblacional con base en una muestra de tamaño 70. ¿Si desea calcular un intervalo de confianza de 95% que sea la mitad de ancho, ¿qué tamaño muestral necesita?

16. Un intervalo de confianza de 95% para una media poblacional se calcula de una muestra de tamaño 50. Se calculará otro intervalo de confianza de 95% para una muestra de tamaño 200, extraída de la misma población. Elija la mejor respuesta que complete el espacio en blanco: El intervalo de una muestra de tamaño 50 será aproximadamente _____ del intervalo de la muestra de tamaño 200.
- Un octavo de ancho.
 - Un cuarto de ancho.
 - La mitad de ancho.
 - El mismo ancho.
 - Dos veces de ancho.
 - Cuatro veces de ancho.
 - Ocho veces de ancho.
17. Con base en pruebas de comportamiento de una gran muestra de uniones soldadas, se calculó un intervalo de confianza de 90% para la media de la dureza Rockwell B de cierto tipo de soldadura de (83.2, 84.1). Determine un intervalo de confianza de 95% para la media de la dureza Rockwell B de este tipo de soldadura.
18. Se hicieron 64 mediciones independientes de la velocidad de la luz. Con un promedio de 299 795 km/s y tenían una desviación estándar de 8 km/s. Verdadero o falso:
- Un intervalo de confianza de 95% para la velocidad de la luz es $299\,795 \pm 1.96$ km/s.
 - La probabilidad es de 95% de que la velocidad de la luz esté en el intervalo $299\,795 \pm 1.96$.
 - Si se hace la medición 65, la probabilidad es de 95% de que estuviera en el intervalo $299\,795 \pm 1.96$.
19. Una caja grande contiene 10 000 cojinetes de bola. Se elige una muestra aleatoria de 120. La media muestral del diámetro es 10 mm y la desviación estándar es 0.24 mm. Verdadero o falso:
- Un intervalo de confianza de 95% para la media del diámetro de los 120 cojinetes en la muestra es $0 \pm (1.96)(0.24)/\sqrt{120}$.
 - Un intervalo de confianza de 95% para la media del diámetro de los 10 000 cojinetes en la caja es $10 \pm (1.96)(0.24)/\sqrt{120}$.
 - Un intervalo de confianza de 95% para la media del diámetro de los 10 000 cojinetes en la caja diez es $10 \pm (1.96)(0.24)/\sqrt{10\,000}$.
20. Todos los días un ingeniero de control de calidad selecciona una muestra aleatoria de 100 pernos de la producción del día, mide sus longitudes y calcula un intervalo de confianza de 95% para la media de la longitud de todos los pernos fabricados ese día. ¿Cuál es la probabilidad de que más de 15 de los intervalos de confianza construidos en los siguientes 250 días no contendrán la media verdadera?
21. Con base en una muestra de registros de reparación, un ingeniero calcula el intervalo de confianza de 95% para la media del costo de reparar un componente de fibra óptica de (\$140, \$160). Un supervisor resume este resultado en un informe, diciendo: “Se tiene una confianza de 95% de que la media del costo de las reparaciones es menor que \$160”. ¿El supervisor está subestimando la confianza, sobreestimándola u obteniéndola de manera correcta? Explique.
22. Un meteorólogo mide la temperatura en el centro de la ciudad de Denver a mediodía todos los días durante un año. Las 365 lecturas tienen un promedio de 57°F y una desviación estándar de 20°F. El meteorólogo calcula un intervalo de confianza de 95% para la media de la temperatura a mediodía de $57^\circ \pm (1.96)(20)/\sqrt{365}$. ¿Es esto correcto? ¿Por qué si o por qué no?

5.2 Intervalos de confianza para proporciones

Los métodos de la sección 5.1, en particular la expresión (5.1), se pueden utilizar con el fin de determinar los intervalos de confianza para la media de cualquier población de la cual se ha extraído una muestra grande. Cuando la población tiene una distribución de Bernoulli, esta expresión toma una forma especial. Se muestra esto último con un ejemplo.

En el ejemplo 5.2 (de la sección 5.1), se construyó un intervalo de confianza para la media del tiempo de vida de una microperforadora cuando perforaba una aleación de acero con bajo contenido de carbono. Ahora suponga que se ha establecido una especificación de que una perforadora debe tener un tiempo de vida mínimo de diez huecos perforados antes de fa-

llar. Se prueba una muestra de 144 microperforadoras, y 120, 83.3%, satisfacen la especificación. Sea p la proporción de microperforadoras en la población que satisface la especificación. Se desea encontrar un intervalo de confianza de 95% para p .

Se empieza construyendo un estimador de p . Sea X el número de perforadoras en la muestra que satisface la especificación. Entonces $X \sim \text{Bin}(n, p)$, donde $n = 144$ es el tamaño muestral. El estimador de p es $\hat{p} = X/n$. En este ejemplo, $X = 120$, por lo que $\hat{p} = 120/144 = 0.833$. La incertidumbre, o desviación estándar de $\hat{p}$, es $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$. Puesto que el tamaño muestral es grande, se tiene por el teorema del límite central (ecuación 4.52 de la sección 4.10) que

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

El razonamiento que se ilustra en las figuras 5.1 y 5.2 (de la sección 5.1) muestra que en 95% de todas las muestras posibles, la proporción poblacional p satisface la siguiente desigualdad:

$$\hat{p} - 1.96\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + 1.96\sqrt{\frac{p(1-p)}{n}} \quad (5.4)$$

A primera vista, la expresión (5.4) parece un intervalo de confianza de 95% para p . Sin embargo, los límites $\hat{p} \pm 1.96\sqrt{p(1-p)/n}$ contienen una p desconocida, y por eso no se puede calcular. El punto de vista tradicional es sustituir p con $\hat{p}$, obtener el intervalo de confianza $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$. Investigaciones recientes muestran que ese intervalo se puede mejorar modificando ligeramente tanto a n como a $\hat{p}$. En específico, se debe sumar 4 al número de los ensayos y 2 al de los éxitos. Así que en lugar de n se utiliza $\tilde{n} = n + 4$, y en lugar de $\hat{p}$ se usa $\tilde{p} = (X + 2)/\tilde{n}$. Un intervalo de confianza de 95% para p es así dado por $\tilde{p} \pm 1.96\sqrt{\tilde{p}(1-\tilde{p})/\tilde{n}}$. En este ejemplo, $\tilde{n} = 148$ y $\tilde{p} = 122/148 = 0.8243$, por eso el intervalo de confianza de 95% es 0.8243 ± 0.0613 , o $(0.763, 0.886)$.

Se justifica lo anterior con base en el teorema del límite central, que requiere que n sea grande. Sin embargo, este método de cálculo de intervalos de confianza es adecuado para cualquier tamaño n de muestra. Cuando se utiliza con muestras pequeñas, podría ocurrir que el límite inferior sea menor a 0 o que el superior a 1. Dado que $0 < p < 1$, un límite inferior menor que 0 se debe sustituir con 0, y un límite superior mayor que 1 se debe sustituir con 1.

Resumen

Sea X el número de éxitos en n ensayos de Bernoulli independientes con probabilidad de éxito p , por lo que $X \sim \text{Bin}(n, p)$.

Se define $\tilde{n} = n + 4$ y $\tilde{p} = \frac{X + 2}{\tilde{n}}$. Entonces un nivel $100(1 - \alpha)\%$ de un intervalo de confianza para p es

$$\tilde{p} \pm z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \quad (5.5)$$

Si el límite inferior es menor que 0, se reemplaza éste con 0. Si el superior es mayor que 1, se reemplaza éste con 1.

El intervalo de confianza dado por la expresión (5.5) algunas veces es llamado intervalo de *Agresti-Coull*, en honor a Alan Agresti y Brent Coull, quienes lo desarrollaron. Para mayor información consulte “Approximate Is Better Than ‘Exact’ for Interval Estimation of Binomial Proportions” (A. Agresti y B. Coull, en *The American Statistician*, 1998:119-126).

Ejemplo

5.12

Los métodos de interpolación se usan para calcular las alturas sobre el nivel del mar para ubicaciones donde las mediciones directas no están disponibles. En el artículo “Transformation of Ellipsoid Heights to Local Leveling Heights” (M. Yanalak y O. Baykal, en *Journal of Surveying Engineering*, 2001:90-103), se evalúa un método de promedio ponderado de interpolación para calcular las alturas de mediciones GPS. El método se estableció para interpretar errores “grandes” (errores cuya magnitud están por encima de umbral comúnmente aceptado) en 26 de 74 ubicaciones de prueba. Determine un intervalo de confianza de 90% para la proporción de ubicaciones en las que este método tendrá errores grandes.

Solución

El número de éxitos es $X = 26$ y el de ensayos $n = 74$. Por tanto, se calcula $\tilde{n} = 74 + 4 = 78$, $\tilde{p} = (26 + 2)/78 = 0.3590$, y $\sqrt{\tilde{p}(1 - \tilde{p})/\tilde{n}} = \sqrt{(0.3590)(0.6410)/78} = 0.0543$. Para un intervalo de confianza de 90%, el valor de $\alpha/2$ es 0.05, por lo que $z_{\alpha/2} = 1.645$. El intervalo de confianza de 90% es, por tanto, $0.3590 \pm (1.645)(0.0543)$, o (0.270, 0.448).

Los intervalos de confianza de un lado se pueden calcular también para proporciones. Son análogos a los intervalos de un lado para una media poblacional (ecuaciones 5.2 y 5.3 de la sección 5.1). Los niveles para los intervalos de confianza de un lado son sólo aproximaciones burdas para muestras pequeñas.

Resumen

Sea X el número de éxitos en n ensayos de Bernoulli independientes con probabilidad p de éxito, por lo que $X \sim \text{Bin}(n, p)$.

Se define $\tilde{n} = n + 4$ y $\tilde{p} = \frac{X + 2}{\tilde{n}}$. Entonces un nivel $100(1 - \alpha)\%$ de un límite inferior de confianza para p es

$$\tilde{p} - z_{\alpha} \sqrt{\frac{\tilde{p}(1 - \tilde{p})}{\tilde{n}}} \quad (5.6)$$

y nivel $100(1 - \alpha)\%$ de un límite superior de confianza para p es

$$\tilde{p} + z_{\alpha} \sqrt{\frac{\tilde{p}(1 - \tilde{p})}{\tilde{n}}} \quad (5.7)$$

Si el límite inferior es menor que 0, se reemplaza con 0. Si el superior es mayor que 1, se reemplaza con 1.

El ejemplo 5.13 muestra cómo calcular el tamaño muestral necesario para un intervalo de confianza que tenga un ancho específico cuando se conoce un valor preliminar de $\tilde{p}$.

Ejemplo

5.13

En el ejemplo 5.12, ¿qué tamaño muestral se necesita para obtener un intervalo de confianza de 95% con ancho ± 0.08 ?

Solución

Un intervalo de confianza de 95% tiene un ancho $\pm 1.96\sqrt{\tilde{p}(1-\tilde{p})/\tilde{n}}$, donde $\tilde{n} = n + 4$. Por tanto, se determina el tamaño muestral n con la ecuación $1.96\sqrt{\tilde{p}(1-\tilde{p})/(n+4)} = 0.08$. De los datos del ejemplo 5.12, $\tilde{p} = 0.3590$. Al sustituir este valor para $\tilde{p}$ y despejando a n , se obtiene $n \approx 135$.

A veces se puede desear calcular un tamaño muestral sin tener disponible un estimador $\tilde{p}$ confiable. La cantidad $\tilde{p}(1-\tilde{p})$, que determina el ancho del intervalo de confianza, se maximiza por $\tilde{p} = 0.5$. Debido a que el ancho es el más grande cuando $\tilde{p}(1-\tilde{p})$ es mayor, se puede calcular un estimador de tamaño muestral conservador con $\tilde{p} = 0.5$ y prosiguiendo como en el ejemplo 5.13.

Ejemplo

5.14

En el ejemplo 5.12, ¿qué tamaño muestral es necesario para garantizar que el ancho del intervalo de confianza de 95% no será mayor que ± 0.08 , si no se ha tomado alguna muestra preliminar?

Solución

Un intervalo de confianza de 95% tiene un ancho $\pm 1.96\sqrt{\tilde{p}(1-\tilde{p})/(n+4)}$. El intervalo de confianza más ancho posible, para una muestra de tamaño n , es $\pm 1.96\sqrt{(0.5)(1-0.5)/(n+4)}$, o $\pm 0.98/\sqrt{n+4}$. Al despejar a n de la ecuación $0.98/\sqrt{n+4} = 0.08$, se obtiene $n \approx 147$. Observe que este cálculo es un poco más grande que el que se obtuvo en el ejemplo 5.13.

El método tradicional

El método que se ha descrito se ha desarrollado recientemente (aunque se creó para simplificar un método mucho más antiguo). Muchas personas todavía usan un método más tradicional. Éste utiliza el tamaño muestral n real en lugar de $\tilde{n}$ y la proporción real $\hat{p}$ en lugar de $\tilde{p}$. Aunque este método todavía es usado, falla para lograr la probabilidad de cobertura establecida, incluso para algunos valores bastante grandes de n . Esto significa que intervalos de confianza $100(1-\alpha)\%$ que se calculan con los métodos tradicionales contendrán la proporción verdadera menos del $100(1-\alpha)\%$ de las veces. El método tradicional no puede ser usado para todas las muestras pequeñas; una regla práctica respecto del tamaño muestral es que tanto $n\hat{p}$ (el número de éxitos) como $n(1-\hat{p})$ (el número de fracasos) deben ser mayores que 10.

Debido a que el método tradicional todavía es muy usado, lo resumimos en el siguiente cuadro. Para tamaños muestrales muy grandes, los resultados del método tradicional son casi idénticos a los obtenidos con el método moderno. Para tamaños muestrales pequeños o medianamente grandes, el punto de vista moderno es mejor.

Resumen

El método tradicional para calcular intervalos de confianza para una proporción (ampliamente usado pero no recomendado)

Sea $\hat{p}$ la proporción de éxitos en un *gran* número n de ensayos de Bernoulli independientes con probabilidad de éxito p . Entonces el intervalo de confianza tradicional de nivel $100(1 - \alpha)\%$ para p es

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \quad (5.8)$$

El método no se puede utilizar a menos que la muestra contenga un mínimo de diez éxitos y diez fracasos.

Ejercicios para la sección 5.2

- Las concentraciones de contaminantes atmosféricos, como monóxido de carbono (CO), se pueden medir con un espectrómetro. En una prueba de calibración, se hicieron 50 mediciones de una muestra de gas del laboratorio que se sabía tenía una concentración de CO de 70 partes por millón (ppm). Se considera que una medición es satisfactoria si está dentro de 5 ppm de la concentración verdadera. De las 50 mediciones, 37 fueron satisfactorias.
 - ¿Qué proporción de mediciones de la muestra fue satisfactoria?
 - Determine un intervalo de confianza de 95% para la proporción de mediciones hechas por este instrumento que serán satisfactorias.
 - ¿Cuántas mediciones se debe tomar para especificar la proporción de mediciones satisfactorias dentro de ± 0.10 con una confianza de 95%?
 - Determine un intervalo de confianza de 99% para la proporción de mediciones hechas por este instrumento que será satisfactorio.
 - ¿Cuántas mediciones se debe tomar para especificar la proporción de mediciones satisfactorias dentro de ± 0.10 con una confianza de 99%?
- En cierto día, se fabricó gran número de fusibles, cada uno tasado a 15 A. Al extraer una muestra de 75 de la producción del día, se encontró que 17 de ellos tenían amperajes de quemado mayores de 15 A.
 - Determine un intervalo de confianza de 95% para la proporción de fusibles fabricada ese día, cuyo amperaje de quemado es mayor que 15 A.
 - Determine un intervalo de confianza de 98% para la proporción de fusibles fabricados ese día, cuyo amperaje de quemado es mayor que 15 A.
 - Determine el tamaño muestral necesario para que un intervalo de confianza de 95% especifique la proporción dentro de ± 0.05 .
 - Determine el tamaño muestral necesario para que un intervalo de confianza de 98% especifique la proporción dentro de ± 0.05 .
 - Si se calcula un intervalo de confianza de 95% a diario durante 200 días, ¿cuál es la probabilidad de que más de 192 intervalos de confianza contengan las proporciones verdaderas?
- Un fabricante de refresco compra latas de aluminio de un distribuidor externo. Se selecciona una muestra aleatoria de 70 latas de un envío grande, se prueba la resistencia de cada una aplicando una carga creciente en los lados de la lata hasta que se perfora. De las 70 latas, 52 satisfacen la especificación para la resistencia de perforación.
 - Determine un intervalo de confianza de 95% para la proporción de latas que satisface la especificación en el envío.
 - Determine un intervalo de confianza de 90% para la proporción de latas que satisface la especificación en el envío.
 - Determine el tamaño muestral necesario para que un intervalo de confianza de 95% especifique la proporción dentro de ± 0.05 .
 - Determine el tamaño de la muestra necesario para que un intervalo de confianza de 90% especifique la proporción dentro de ± 0.05 .

- e) Si un intervalo de confianza de 90% se calcula a diario durante 300 días, ¿cuál es la probabilidad de que más de 280 intervalos de confianza contengan las proporciones verdaderas?
4. Con referencia al ejercicio 1, encuentre un límite inferior de confianza de 95% para la proporción de medidas satisfactorias.
5. Con referencia al ejercicio 2, encuentre un límite superior de confianza de 98% para la proporción de fusibles con amperajes de quemado mayores de 15 A.
6. Con referencia al ejercicio 3, encuentre un límite inferior de confianza de 99% para la proporción de latas que satisface la especificación.
7. Se prueba una muestra aleatoria de 400 componentes electrónicos fabricados por cierto proceso y se encuentra que 30 están defectuosos.
- Sea p la proporción de componentes fabricados con este proceso que están defectuosos. Determine un intervalo de confianza de 95% para p .
 - ¿Cuántos componentes se deben muestrear con el propósito de que el intervalo de confianza de 95% especifique la proporción defectuosa dentro de ± 0.02 ?
 - (Difícil) La compañía envía los componentes en lotes de 200. Los lotes que contienen más de 20 componentes defectuosos pueden ser regresados. Determine un intervalo de confianza de 95% para la proporción de lotes que serán regresados.
8. Con referencia al ejercicio 7, se fabricará un dispositivo en el cual se conectarán en serie dos de los componentes del ejercicio 7. Los componentes funcionan de manera independiente, el dispositivo funcionará sólo si ambos componentes funcionan. Sea q la probabilidad de que un dispositivo funcione. Determine un intervalo de confianza de 95% para q . (Sugerencia: exprese q en función de p , y después utilice el resultado del ejercicio 7a.)
9. El artículo “Leachate from Land Disposed Residential Construction Waste” (W. Weber, Y. Jang y cols., en *Journal of Environmental Engineering*, 2002:237-245) presenta un estudio de la contaminación en basureros que contienen desechos de construcción y desperdicio de demolición. De un sitio de prueba se tomaron muestras de lixiviado. De cada 42 muestras, 26 contienen niveles detectables de plomo, 41 de arsénico y 32 de cromo.
- Encuentre un intervalo de confianza de 90% para la probabilidad de que una muestra contendrá un nivel detectable de plomo.
 - Determine un intervalo de confianza de 95% para la probabilidad de que una muestra contenga un nivel detectable de arsénico.
 - Determine un intervalo de confianza de 99% para la probabilidad de que una muestra contenga un nivel detectable de cromo.
10. Los aceros inoxidables pueden ser susceptibles al agrietamiento de corrosión por tensión bajo ciertas condiciones. Un ingeniero especializado en materiales está interesado en determinar la proporción de fallas de aleaciones de acero que son atribuibles al agrietamiento de corrosión por tensión.
- En ausencia de datos preliminares, ¿de qué tamaño debe ser una muestra para asegurar que el intervalo de confianza de 95% especificará la proporción dentro de ± 0.05 ?
 - En una muestra de 100 fallas, 20 eran ocasionadas por el agrietamiento de corrosión por tensión. Encuentre un intervalo de confianza de 95% para la proporción de fallas ocasionadas por el agrietamiento de corrosión por tensión.
 - Con base en los datos del inciso (b), calcule el tamaño muestral necesario con el propósito de que el intervalo de confianza de 95% especificará la proporción dentro de ± 0.05 .
11. Para que los proyectos de remediación ecológica muy importantes sean exitosos, deben tener apoyo público. El artículo “Modelling the Non-Market Environmental Costs and Benefits of Biodiversity Using Contingent Value Data” (D. Macmillan, E. Duff y D. Elston, en *Environmental and Resource Economics*, 2001:391-410) notifica los resultados de una encuesta en que a votantes escoceses se les preguntó si estarían dispuestos a pagar impuestos adicionales con la finalidad de restaurar el bosque Affric. De los 189 que respondieron, 61 decían que sí apoyarían esa medida.
- Suponiendo que los 189 votantes que respondieron, constituyen una muestra aleatoria, determine un intervalo de confianza de 90% para la proporción de votantes que estarían dispuestos a pagar para restaurar el bosque Affric.
 - ¿Cuántos votantes se deben muestrear para especificar la proporción dentro de ± 0.03 con una confianza de 90%?
 - Se planea realizar otra encuesta en la cual se les preguntará a los votantes si estarían dispuestos a pagar con la

finalidad de restaurar el bosque Strathspey. En este momento, no hay alguna estimación disponible. Determine un cálculo conservador del tamaño muestral necesario con el propósito de que la proporción estará especificada dentro de ± 0.03 con una confianza de 90 por ciento.

12. Un analista del mercado de valores observa que, en cierto año, el precio de cada acción de la IBM aumentó en 131 de los 252 días bursátiles. ¿Con estos datos se puede encontrar un intervalo de confianza de 95% para la proporción de días en que la acción de IBM se incrementa? Explique.

5.3 Intervalos de confianza para la media poblacional con muestras pequeñas

Los métodos descritos en la sección 5.1 con el fin de calcular intervalos de confianza para la media de una población requieren que el tamaño muestral sea grande. Cuando éste es pequeño, no hay ningún buen método general para encontrar intervalos de confianza. Sin embargo, cuando la población es aproximadamente normal, se puede utilizar una distribución de probabilidad denominada *t* de Student para calcular los intervalos de confianza para una media poblacional. En esta sección se describe dicha distribución y se muestra cómo utilizarla.

Distribución *t* de Student

Si $\bar{X}$ es la media de una muestra grande de tamaño n de una población con media μ y varianza σ^2 , entonces el teorema del límite central especifica que $\bar{X} \sim N(\mu, \sigma^2/n)$. La cantidad $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ tiene una distribución normal con media 0 y varianza 1. Además, la desviación estándar muestral s estará cerca de la desviación estándar σ poblacional. Por esta razón la cantidad $(\bar{X} - \mu)/(s/\sqrt{n})$ es aproximadamente normal con media 0 y varianza 1, por lo que se pueden buscar las probabilidades relacionadas con esta cantidad en la tabla normal estándar (tabla *z*). Esto último permite que se calcule intervalos de confianza para diferentes niveles para la media poblacional μ .

¿Qué se puede hacer si $\bar{X}$ es la media de una muestra *pequeña*? Si éste es pequeño, s podría no estar cercano a σ , y $\bar{X}$ puede no ser aproximadamente normal. Si no se sabe nada acerca de la población de la que la muestra pequeña fue extraída, entonces no hay ningún método fácil para calcular intervalos de confianza. Sin embargo, si la población es aproximadamente normal, $\bar{X}$ lo será incluso cuando el tamaño muestral sea pequeño. Lo anterior propicia que aún se puede utilizar la cantidad $(\bar{X} - \mu)/(s/\sqrt{n})$, pero debido a que s no está necesariamente cercana a σ , esta cantidad no tendrá una distribución normal. En su lugar, tiene la **distribución *t* de Student** con $n - 1$ grados de libertad, que se denota por t_{n-1} . El número de grados de libertad para la distribución *t* es uno menos que el tamaño muestral.

La distribución *t* de Student fue descubierta en 1908 por William Sealy Gossett, un estadístico que trabajó en la cervecera Guinness, en Dublín, Irlanda. La dirección de Guinness consideró que el descubrimiento era información privada y prohibió a Gossett que lo publicara. Aun así, él lo publicó, usando el seudónimo “Estudiante”. Gossett había hecho ya esto antes; véase la sección 4.3.

Resumen

Sea $X_1, \dots, X_n$ una muestra *pequeña* (por ejemplo $n < 30$) de una población *normal* con media μ . Entonces la cantidad

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

tiene una distribución *t* de Student con $n - 1$ grados de libertad, denotada por t_{n-1} .

Cuando n es grande, la distribución de la cantidad $(\bar{X} - \mu)/(s/\sqrt{n})$ es muy cercana a la curva normal, de esta forma la curva normal puede usarse en lugar de la *t* de Student.

La función de densidad de probabilidad de la distribución *t* de Student es diferente para distintos grados de libertad. La figura 5.9 presenta gráficas de la función de densidad de probabilidad para diferentes elecciones de grados de libertad. Las curvas tienen una forma similar a la curva normal, o *z*, es una curva con media 0 y desviación estándar 1. Sin embargo, las curvas *t* son más extendidas. Por ejemplo, la curva *t* con un grado de libertad corresponde a un tamaño muestral de 2. Cuando se extraen muestras de tamaño 2, ocurrirá con frecuencia que la desviación estándar muestral *s* sea mucho más pequeña que σ , lo que llevará a que el valor de $(\bar{X} - \mu)/(s/\sqrt{n})$ sea muy grande (ya sea positivo o negativo). Por esta razón, la curva *t* con un grado de libertad tiene mucho más área en las colas. Para tamaños muestrales más grandes, el valor de *s* es menos probable que esté lejos de σ y la curva *t* es más cercana a la curva normal. Con diez grados de libertad (correspondiendo a un tamaño muestral de 11), la diferencia entre la curva *t* y la curva normal no es grande. Si una curva *t* con 30 grados de libertad estuviera dibujada en la figura 5.9, sería indistinguible de la curva normal.

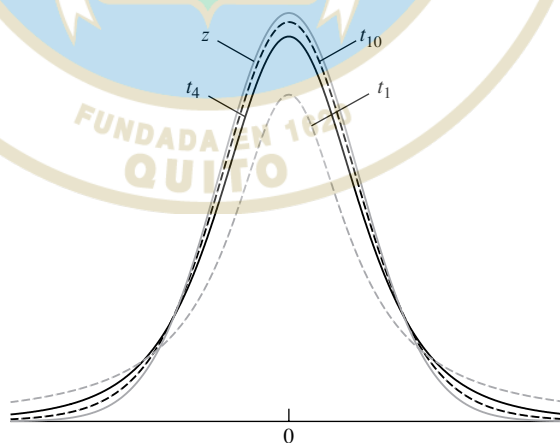


FIGURA 5.9 Gráficas de la función de densidad de probabilidad de la curva *t* de Student para diferentes grados de libertad. La curva normal con media 0 y varianza 1 (curva *z*) es graficada para comparar. Las curvas *t* están más extendidas que la normal, pero la cantidad de extensión adicional disminuye conforme se aumenta el número de grados de libertad.

La tabla A.3 (en el Apéndice A), denominada **tabla t** , proporciona probabilidades relacionadas con la distribución t de Student. Se presentan algunos ejemplos para mostrar el uso de la tabla.

Ejemplo 5.15

Se extrae una muestra aleatoria de tamaño 10 de una distribución normal con media 4. La estadística t de Student $t = (\bar{X} - 4)/(s/\sqrt{10})$ es calculada. ¿Cuál es la probabilidad de que $t > 1.833$?

Solución

Esta estadística t tiene $10 - 1 = 9$ grados de libertad. De la tabla t , $P(t > 1.833) = 0.05$. Véase la figura 5.10.

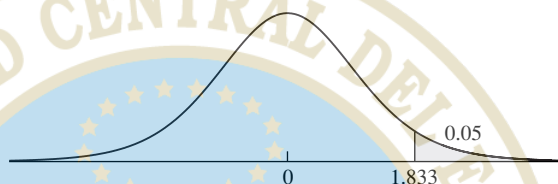


FIGURA 5.10 Solución al ejemplo 5.15.

Ejemplo 5.16

Con referencia al ejemplo 5.15, determine $P(t > 1.5)$.

Solución

Buscando a través del renglón correspondiente a 9 grados de libertad, se ve que la tabla t no lista el valor 1.5. Se encuentra que $P(t > 1.383) = 0.10$ y que $P(t > 1.833) = 0.05$. Se concluye entonces que $0.05 < P(t > 1.5) < 0.10$. Véase la figura 5.11. Un resultado más preciso que esta desigualdad se puede obtener mediante interpolación lineal

$$P(t > 1.5) \approx 0.10 - \frac{1.5 - 1.383}{1.833 - 1.383}(0.10 - 0.05) = 0.0870$$

Un software proporciona la respuesta correcta con tres dígitos significativos como 0.0839.

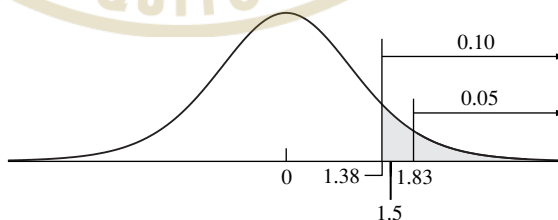


FIGURA 5.11 Solución al ejemplo 5.16.

Ejemplo 5.17

Determine el valor para la distribución t_{12} cuya cola superior de probabilidad es 0.025.

Solución

Al buscar hacia abajo la columna encabezada con “0.025” en el renglón correspondiente a 12 grados de libertad. El valor para t_{12} es 2.179.

Ejemplo

5.18

Determine el valor para la distribución de t_{14} cuya cola de la probabilidad inferior es 0.01.

Solución

Busque hacia abajo de la columna encabezada con “0.01” en el renglón que corresponde a 14 grados de libertad. El valor para t_{14} es 2.624. Este valor corta un área, o probabilidad, de 1% en la cola superior. El valor cuya cola inferior de probabilidad es 1% es -2.624 .

No use la estadística t de Student si la muestra contiene datos atípicos

Para que la estadística t de Student sea válida, la muestra debe provenir de una población que es aproximadamente normal. Tales muestras rara vez contienen datos atípicos. Por tanto, los métodos que implican la estadística t de Student *no* se deben utilizar en muestras que contienen datos atípicos.

Intervalos de confianza al usar la distribución t de Student

Cuando el tamaño muestral es pequeño y la población es aproximadamente normal, se puede utilizar la distribución t de Student para calcular intervalos de confianza. Se muestra esto último con un ejemplo.

Un metalúrgico estudia un nuevo proceso de soldadura. Fabrica cinco uniones soldadas y mide la resistencia producida por cada uno. Los cinco valores (en ksi) son 56.3, 65.4, 58.7, 70.1 y 63.9. Suponga que estos valores son una muestra aleatoria de una población aproximadamente normal. La tarea es determinar un intervalo de confianza para la media de la resistencia de las soldaduras hechas por este proceso.

Cuando el tamaño muestral es grande, no necesita preocuparse mucho acerca de la naturaleza de la población, porque el teorema del límite central garantiza que la cantidad $\bar{X}$ tendrá una distribución aproximadamente normal. Sin embargo, cuando la muestra es pequeña, la distribución de la población debe ser aproximadamente normal.

En este caso el intervalo de confianza se construye de la misma manera que en la sección 5.1, exceptuando que el puntaje z se reemplaza con un valor de la distribución t de Student. La cantidad

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

tiene una distribución t de Student con $n - 1$ grados de libertad. La figura 5.12 muestra la distribución t_4 . De la tabla t de Student se encuentra que 95% del área bajo la curva está contenida entre los valores $t = -2.776$ y $t = 2.776$. Por consecuencia, para 95% de todas las muestras que se pudo haber elegido,

$$-2.776 < \frac{\bar{X} - \mu}{s/\sqrt{n}} < 2.776$$

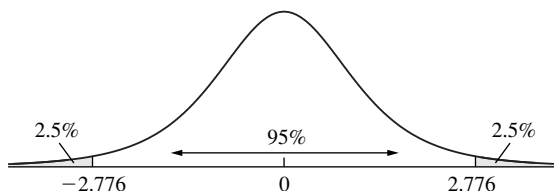


FIGURA 5.12 La distribución t de Student con cuatro grados de libertad. Un 95% del área cae entre $t = -2.776$ y $t = 2.776$.

Expresado de otra manera, para 95% de todas las muestras que se pudo haber elegido, se tiene que

$$-2.776 \frac{s}{\sqrt{n}} < \bar{X} - \mu < 2.776 \frac{s}{\sqrt{n}}$$

Multiplicando por -1 y al sumar $\bar{X}$ en todos los lados de la desigualdad, se obtiene un intervalo de confianza 95% para μ :

$$\bar{X} - 2.776 \frac{s}{\sqrt{n}} < \mu < \bar{X} + 2.776 \frac{s}{\sqrt{n}}$$

En este ejemplo, la media muestral es $\bar{X} = 62.88$ y la desviación estándar muestral es $s = 5.4838$. El tamaño muestral es $n = 5$. Al sustituir valores para $\bar{X}$, s y n , se encuentra que un intervalo de confianza de 95% para μ es $62.88 - 6.81 < \mu < 62.88 + 6.81$, o $(56.07, 69.69)$.

En general, para producir un intervalo de confianza de nivel $100(1 - \alpha)\%$, sea $t_{n-1, \alpha/2}$ el $1 - \alpha/2$ cuantil de la distribución t de Student con $n - 1$ grados de libertad, es el valor que corta un área de $\alpha/2$ en la cola de la derecha. Por ejemplo, antes se encontró que $t_{4, 0.025} = 2.776$. Entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ para la media μ poblacional es $\bar{X} - t_{n-1, \alpha/2}(s/\sqrt{n}) < \mu < \bar{X} + t_{n-1, \alpha/2}(s/\sqrt{n})$, o $\bar{X} \pm t_{n-1, \alpha/2}(s/\sqrt{n})$.

Resumen

Sea $X_1, \dots, X_n$ una muestra aleatoria *pequeña* de una población *normal* con media μ . Entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ para μ es

$$\bar{X} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \quad (5.9)$$

¿Cómo se determina si la distribución t de Student es adecuada?

La distribución t de Student es adecuada siempre que la muestra provenga de una población que es más o menos normal. A veces se sabe por experiencia si un proceso genera datos con una distribución aproximada. Sin embargo, en muchos casos, se debe decidir si una población es aproximadamente normal examinando la muestra. Por desgracia, cuando el tamaño muestral es pequeño, desviaciones a la normalidad pueden ser difíciles de detectar. Una manera ra-

zonable de proceder es construir un diagrama de caja o de puntos de la muestra. Si estos diagramas no revelan una asimetría fuerte o algún dato atípico, entonces en la mayor parte de los casos la distribución t de Student será confiable. En principio, también se puede determinar si una población es aproximadamente normal al construir una gráfica de probabilidad. Sin embargo, con muestras pequeñas, los diagramas de caja y de puntos son más fáciles de dibujar, especialmente a mano.

Ejemplo

5.19

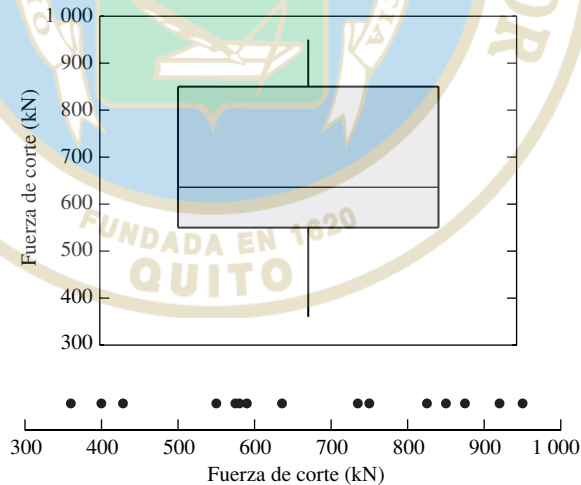
El artículo “Direct Strut-and-Tie Model for Prestressed Deep Beams” (K. Tan, K. Tong y C. Tang, en *Journal of Structural Engineering*, 2001:1076-1084) presenta mediciones de la fuerza nominal de corte (en kN) para una muestra de 15 vigas de concreto. Los resultados son

580	400	428	825	850	875	920	550
575	750	636	360	590	735	950	

¿Es adecuado utilizar la estadística t de Student para construir un intervalo de confianza de 99% para la media de la fuerza de corte? Si es así, construya el intervalo de confianza. Si no, explique por qué.

Solución

Para determinar si la estadística t de Student es adecuada, se hace un diagrama de caja y de puntos de la muestra. Éstos se muestran en la figura siguiente.



No hay evidencia de una desviación muy importante a la normalidad; en particular las gráficas no son fuertemente asimétricas, y no hay algún dato atípico. El método t de Student es adecuado. Por tanto, se calcula $\bar{X} = 668.27$ y $s = 192.089$. Se utiliza la expresión (5.9) con $n = 15$ y $\alpha/2 = 0.005$. De la tabla t con 14 grados de libertad, se encuentra $t_{14, 0.005} = 2.977$. El intervalo de confianza de 99% es $668.27 \pm (2.977)(192.089)/\sqrt{15}$, o $(520.62, 815.92)$.

El siguiente resultado de computadora (de MINITAB) presenta el intervalo de confianza calculado en el ejemplo 5.19.

One-Sample T: Strength

Test of $\mu = 0$ vs not $= 0$

Variable	N	Mean	StDev	SE Mean	99% CI
Strength	15	668.2667	192.0891	49.59718	(520.6159, 815.9175)

El resultado es muy claro. La cantidad etiquetada “SE Mean” representa la desviación estándar de la media muestral, $s/\sqrt{n}$.

Ejemplo

5.20

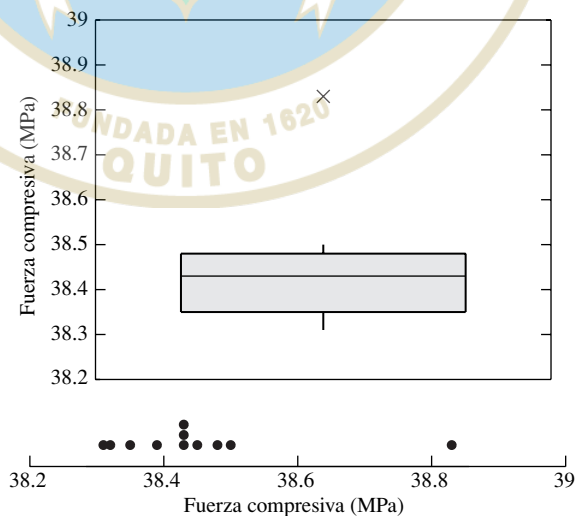
En el artículo referido en el ejemplo 5.19, la fuerza compresiva cilíndrica (en MPa) fue medida para 11 vigas. Los resultados fueron

38.43 38.43 38.39 38.83 38.45 38.35 38.43 38.31 38.32 38.48 38.50

¿Es adecuado utilizar la estadística t de Student para construir un intervalo de confianza de 95% para la media de la fuerza compresiva cilíndrica? Si es así, construya el intervalo de confianza. Si no, explique por qué.

Solución

Como en el ejemplo 5.19, se realizará un diagrama de caja y un diagrama de puntos de la muestra. Éstos se muestran en la figura siguiente.



Hay un dato atípico en esta muestra. La estadística t de Student no se debe utilizar.

Ejemplo**5.21**

Un ingeniero lee un informe que dice que una muestra de 11 vigas de concreto tenía una fuerza compresiva promedio de 38.45 MPa con desviación estándar de 0.14 MPa. ¿Se debe utilizar la curva t para encontrar un intervalo de confianza para la media de la fuerza compresiva?

Solución

No. El problema es que no hay ninguna manera de saber si las mediciones provienen de una población normal. Por ejemplo, si las mediciones contienen un dato atípico (como en el ejemplo 5.20), el intervalo de confianza sería inválido.

La distribución t de Student se puede utilizar para calcular intervalos de confianza de un lado. Las fórmulas son análogas a las que se utilizan con muestras grandes.

Sea $X_1, \dots, X_n$ una muestra aleatoria *pequeña* de una población *normal* con media μ . Entonces un límite superior de confianza de nivel $100(1 - \alpha)\%$ para μ es

$$\bar{X} + t_{n-1, \alpha} \frac{s}{\sqrt{n}} \quad (5.10)$$

y un límite inferior de confianza de nivel $100(1 - \alpha)\%$ para μ es

$$\bar{X} - t_{n-1, \alpha} \frac{s}{\sqrt{n}} \quad (5.11)$$

Utilice z , no t , si se conoce a σ

En ocasiones se puede tomar una pequeña muestra de una población normal cuya desviación estándar σ se conoce. En estos casos, no se utiliza la curva t de Student, porque no se está aproximando a σ con s . En su lugar, se utiliza la tabla z . El ejemplo 5.22 ilustra el método.

Ejemplo**5.22**

Con referencia al ejemplo 5.19. Suponga, con base en un número muy grande de mediciones previas de otras vigas, que la población de las fuerzas de corte es aproximadamente normal, con desviación estándar $\sigma = 180.0$ kN. Encuentre un intervalo de confianza de 99% para la media de la fuerza de corte.

Solución

Se calcula $\bar{X} = 668.27$. No se necesita calcular s , porque se conoce la desviación estándar poblacional σ . Dado que se quiere un intervalo de confianza de 99%, $\alpha/2 = 0.005$. Ya que se conoce σ , se utiliza $z_{\alpha/2} = z_{0.005}$, en lugar de un valor de t de Student, para calcular el intervalo de confianza. De la tabla z se obtiene $z_{0.005} = 2.58$. El intervalo de confianza es $668.27 \pm (2.58)(180.0)/\sqrt{15}$, o (548.36, 788.18).

Es importante recordar que cuando el tamaño muestral es pequeño, la población debe ser aproximadamente normal, se conozca o no la desviación estándar.

Resumen

Sea $X_1, \dots, X_n$ una muestra aleatoria (de cualquier tamaño) de una población *normal* con media μ . Si se conoce la desviación estándar σ , entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ es

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (5.12)$$

En ocasiones se tiene un solo valor que se muestra de una población normal con desviación estándar conocida. En estos casos se puede obtener un intervalo de confianza para μ y deducir como un caso especial de la expresión (5.12) al hacer $n = 1$.

Resumen

Sea X un solo valor que se muestra de una población *normal* con media μ . Si se conoce la desviación estándar σ , entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ para μ es

$$X \pm z_{\alpha/2} \sigma \quad (5.13)$$

Ejercicios para la sección 5.3

- Determine el valor de $t_{n-1, \alpha/2}$ necesario para construir un intervalo de confianza de dos lados de un nivel específico con los siguientes tamaños muestrales:
 - Nivel 90%, tamaño muestral 9.
 - Nivel 95%, tamaño muestral 5.
 - Nivel 99%, tamaño muestral 29.
 - Nivel 95%, tamaño muestral 2.
 - $t = 1.729$, tamaño muestral 20.
 - $t = 3.707$, tamaño muestral 7.
 - $t = 3.707$, tamaño muestral 27.
- Determine el valor de $t_{n-1, \alpha}$ necesario para construir un límite superior o inferior de confianza en cada uno de los casos del ejercicio 1.
- Determine el nivel de confianza para un intervalo de dos lados que está basado en el valor dado de $t_{n-1, \alpha/2}$ y el tamaño muestral específico.
 - $t = 2.179$, tamaño muestral 13.
 - $t = 3.365$, tamaño muestral 6.
- Verdadero o falso: La distribución t de Student se puede utilizar para construir un intervalo de confianza para la media de cualquier población, en tanto que el tamaño muestral sea pequeño.
- El artículo "Ozone for Removal of Acute Toxicity from Logyard Run-off" (M. Zenaitis y S. Duff, en *Ozone Science and Engineering*, 2002:83-90) presenta análisis químicos del agua que escurre de aserraderos en la Columbia Británica. Incluye seis mediciones de pH para seis muestras de agua: 5.9, 5.0, 6.5, 5.6, 5.9, 6.5. Suponiendo que éstas sean una muestra aleatoria de las muestras de agua de una población aproximadamente normal, encuentre un intervalo de confianza de 95% para la media del pH.

6. Los siguientes son resúmenes estadísticos para un conjunto de datos. ¿Sería adecuado utilizar la distribución t de Student para construir un intervalo de confianza de estos datos? Explique.

N	Mean	Median	TrMean	StDev	SE Mean
10	8.905	6.105	6.077	9.690	3.064
Minimum	Maximum	Q1	Q3		
0.512	39.920	1.967	8.103		

7. El artículo “An Automatic Visual System for Marble Tile Classification” (L. Carrino, W. Polini, y S. Turchetta, en *Journal of Engineering Manufacture*, 2002:1095-1108) describe una medida para la sombra del azulejo de mármol en el cual la cantidad de luz reflejada por éste se mide en una escala de 0-255. Un azulejo perfectamente negro no refleja luz alguna y mide 0, y un azulejo perfectamente blanco mediría 255. Se midió una muestra de nueve azulejos Mezza Perla, con los siguientes resultados:
- | | | | | |
|---------|---------|---------|---------|---------|
| 204.999 | 206.149 | 202.102 | 207.048 | 203.496 |
| 206.343 | 203.496 | 206.676 | 205.831 | |
- ¿Es adecuado utilizar la estadística t de Student para construir un intervalo de confianza de 95% para la media de la sombra del azulejo Mezza Perla? Si es así, hágalo. Si no, explique por qué.
8. Una química hizo ocho mediciones independientes del punto de fusión del tungsteno. Obtuvo una media muestral de 3 410.14°C y una desviación estándar muestral de 1.018°C.
- Determine un intervalo de confianza de 95% para el punto de fusión del tungsteno.
 - Determine un intervalo de confianza de 98% para el punto de fusión del tungsteno.
 - Si las ocho mediciones hubieran sido 3 409.76, 3 409.80, 3 412.66, 3 409.79, 3 409.76, 3 409.77, 3 409.80 y 3 409.78 ¿serían válidos los intervalos de confianza que se encuentran en los incisos a) y b)? Explique.
9. Se hacen ocho mediciones independientes del diámetro de un pistón. Las mediciones (en pulgadas) son 3.236, 3.223, 3.242, 3.244, 3.228, 3.253, 3.253 y 3.230.
- Realice un diagrama de puntos de los ocho valores.
 - ¿Se debe utilizar la curva t para encontrar un intervalo de confianza de 99% para el diámetro de este pistón? Si es así, encuentre el intervalo de confianza. Si no, explique por qué.
 - Se toman ocho mediciones independientes del diámetro de otro pistón. Las mediciones en este momento son 3.295, 3.232, 3.261, 3.248, 3.289, 3.245, 3.576 y 3.201. Realice un diagrama de puntos de estos valores.
- d) ¿Se debe utilizar la curva t para encontrar un intervalo de confianza de 95% para el diámetro de este pistón? Si es así, encuentre el intervalo de confianza. Si no, explique por qué.
10. Se toman cinco mediciones de la clasificación de octano para un tipo especial de gasolina. Los resultados (en %) son 87.0, 86.0, 86.5, 88.0, 85.3. Encuentre un intervalo de confianza de 99% para la media de la clasificación de octano de media para este tipo de gasolina.
11. Un modelo de transferencia de calor de un cilindro sumergido en un líquido predice que el coeficiente de transferencia de calor para el cilindro es constante en razones muy bajas de circulación del fluido. Se toma una muestra de diez mediciones. Los resultados, en W/m^2K , son
- | | | | | |
|------|------|------|------|------|
| 13.7 | 12.0 | 13.1 | 14.1 | 13.1 |
| 14.1 | 14.4 | 12.2 | 11.9 | 11.8 |
- Determine un intervalo de confianza de 95% para el coeficiente de transferencia de calor.
12. Los tensioactivos son agentes químicos, como detergentes, que bajan la tensión superficial de un líquido. Son importantes en la limpieza de suelos contaminados. En un experimento para determinar la eficacia de cierto método para retirar tolueno de arena, esta última fue lavada con un agente tensioactivo, y luego enjuagada con agua desionizada. Es importante la cantidad de tolueno que sale en el enjuague. En cinco de estos experimentos, las cantidades de tolueno eliminado en el ciclo de enjuague, expresado como porcentaje de la cantidad total originalmente presente, fueron de 5.0, 4.8, 9.0, 10.0 y 7.3. Determine el intervalo de confianza de 95% para el porcentaje de tolueno eliminado en el enjuague. (Este ejercicio está basado en el artículo “Laboratory Evaluation of the Use of Surfactants for Ground Water Remediation and the Potential for Recycling Them” D. Lee, R. Cody, y B. Hoyle, en *Ground Water Monitoring and Remediation*, 2001:49-57.)

13. En un experimento para medir la razón de absorción de pesticidas a través de la piel, 500 μg de uniconazol se aplicó a la piel de cuatro ratas. Después de diez horas, las cantidades absorbidas (en μg) fueron 0.5, 2.0, 1.4 y 1.1. Encuentre un intervalo de confianza de 90% para la media de la cantidad absorbida.
14. El siguiente resultado de MINITAB presenta un intervalo de confianza para una media poblacional.

One-Sample T: X

Variable	N	Mean	StDev	SE Mean	95% CI
X	10	6.59635	0.11213	0.03546	(6.51613, 6.67656)

- a) ¿Cuántos grados de libertad tiene la distribución t de Student?
- b) Utilice la información en el resultado, junto con la tabla t , para calcular un intervalo de confianza de 99 por ciento.
15. El siguiente resultado de MINITAB presenta un intervalo de confianza para una media poblacional, pero algunos de los números están borrosos y son ahora ilegibles. Complete los números faltantes para (a), (b) y (c).

One-Sample T: X

Variable	N	Mean	StDev	SE Mean	99% CI
X	20	2.39374	(a)	0.52640	((b), (c))

16. La concentración de monóxido de carbono (CO) en una muestra de gas se mide con un espectrómetro y se encuentra que es de 85 ppm. A partir de la gran experiencia con este instrumento, se cree que sus mediciones no tienen sesgos y se distribuyen normalmente, con incertidumbre (desviación estándar) de 8 ppm. Determine un intervalo de confianza de 95% para la concentración de CO en esta muestra.

5.4 Intervalos de confianza para la diferencia entre dos medias

Ahora se tratan ejemplos en los que se desea calcular la diferencia entre las medias de dos poblaciones. Los datos constarán de dos muestras, una para cada población. La idea básica es simple. Se calculará la diferencia de las medias muestrales y la desviación estándar de esa diferencia. Entonces una modificación simple de la expresión (5.1) (de la sección 5.1) proporcionará el intervalo de confianza. El método que se describe está basado en los resultados que se relacionan con la suma y la diferencia de dos variables aleatorias normales independientes que se presentaron en la sección 4.5. Aquí se repasan estos resultados:

Sean X y Y independientes, con $X \sim N(\mu_X, \sigma_X^2)$ y $Y \sim N(\mu_Y, \sigma_Y^2)$. Entonces

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2) \quad (5.14)$$

$$X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2) \quad (5.15)$$

Ahora se verá cómo construir un intervalo de confianza para la diferencia entre dos medias poblacionales. Como ejemplo, suponga que se ha estado desarrollando un diseño nuevo de foco que se piensa durará más que el diseño viejo. Una muestra aleatoria simple de 144 focos nuevos tiene un tiempo de vida promedio de 578 horas y una desviación estándar de 22 horas. Una muestra aleatoria simple de 64 focos viejos tiene tiempo de vida promedio de 551 horas y desviación estándar de 33 horas. Las muestras son independientes, de tal manera que los tiempos de vida para una muestra no influyen sobre los tiempos de vida de la otra. Se quiere encontrar un intervalo de confianza de 95% para la diferencia entre la media de los tiempos de vida de los focos de los dos diseños.

Se inicia por traducir el problema en el lenguaje estadístico. Se tiene una muestra aleatoria simple $X_1, \dots, X_{144}$ de los tiempos de vida de los focos nuevos. La media muestral es $\bar{X} = 578$ y la desviación estándar muestral es $s_X = 22$. Se tiene otra muestra aleatoria simple $Y_1, \dots, Y_{64}$ de los tiempos de vida de los focos viejos. Esta muestra tiene una media $\bar{Y} = 551$ y desviación estándar $s_Y = 33$. Las medias poblacionales y las desviaciones estándares no se conocen. Se denota a la media poblacional de los tiempos de vida de los nuevos focos por μ_X y la media poblacional de los focos viejos por μ_Y . Se denota las correspondientes desviaciones estándares por σ_X y σ_Y . Se tiene interés en la diferencia $\mu_X - \mu_Y$.

Se puede construir el intervalo de confianza para $\mu_X - \mu_Y$ determinando la distribución $\bar{X} - \bar{Y}$. Mediante el teorema del límite central, $\bar{X}$ proviene de una distribución normal con media μ_X y desviación estándar $\sigma_X/\sqrt{144}$, y $\bar{Y}$ proviene de una distribución normal con media μ_Y y desviación estándar $\sigma_Y/\sqrt{64}$. Dado que las muestras son independientes, se tiene por medio de la expresión (5.15) que la diferencia $\bar{X} - \bar{Y}$ proviene de una distribución normal con media $\mu_X - \mu_Y$ y varianza $\sigma_{\bar{X} - \bar{Y}}^2 = \sigma_X^2/144 + \sigma_Y^2/64$. La figura 5.13 muestra la distribución de $\bar{X} - \bar{Y}$ e indica que 95% intermedio de la curva tiene un ancho $\pm 1.96\sigma_{\bar{X} - \bar{Y}}$.

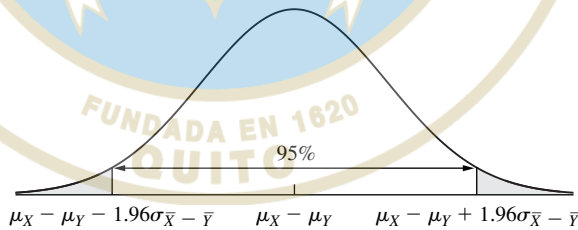


FIGURA 5.13 La diferencia observada $\bar{X} - \bar{Y} = 27$ se extrae de una distribución normal con media $\mu_X - \mu_Y$ y desviación estándar $\sigma_{\bar{X} - \bar{Y}} = \sqrt{\sigma_X^2/144 + \sigma_Y^2/64}$.

Al estimar las desviaciones estándares poblacionales σ_X y σ_Y con las desviaciones estándares muestrales $s_X = 22$ y $s_Y = 33$, respectivamente, se estima $\sigma_{\bar{X} - \bar{Y}} \approx \sqrt{22^2/144 + 33^2/64} = 4.514$. Por tanto, el intervalo de confianza de 95% para $\mu_X - \mu_Y$ es $578 - 551 \pm 1.96(4.514)$, o 27 ± 8.85 .

Resumen

Sea $X_1, \dots, X_{n_X}$ una muestra aleatoria *grande* de tamaño n_X de una población con media μ_X y desviación estándar σ_X , y sea $Y_1, \dots, Y_{n_Y}$ una muestra aleatoria simple *grande* de tamaño n_Y de una población con media μ_Y y desviación estándar σ_Y . Si las dos muestras son independientes, con un intervalo de confianza de nivel $100(1 - \alpha)\%$ para $\mu_X - \mu_Y$ es

$$\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} \quad (5.16)$$

Cuando los valores de σ_X y σ_Y son desconocidos, se pueden sustituir con las desviaciones estándares muestrales s_X y s_Y .

Ejemplo 5.23

La composición química del suelo varía con la profundidad. El artículo “Sampling Soil Water in Sandy Soils: Comparative Analysis of Some Common Methods” (M. Ahmed, M. Sharma y colaboradores, en *Communications in Soil Science and Plant Analysis*, 2001:1677-1686) describe análisis químicos del suelo tomado de una granja en Australia occidental. Se toman 50 muestras a profundidades de 50 y 250 cm. A una profundidad de 50 cm, la concentración promedio de NO_3 (en mg/l) era de 88.5 con una desviación estándar de 49.4. A una profundidad de 250 cm, la concentración promedio era de 110.6 con una desviación estándar de 51.5. Determine un intervalo de confianza de 95% para la diferencia entre las concentraciones de NO_3 a las dos profundidades.

Solución

Sean $X_1, \dots, X_{50}$ las concentraciones de 50 muestras tomadas a 50 cm y sean $Y_1, \dots, Y_{50}$ las concentraciones de 50 muestras tomadas a 250 cm. Entonces $\bar{X} = 88.5$, $\bar{Y} = 110.6$, $s_X = 49.4$ y $s_Y = 51.5$. Los tamaños muestrales son $n_X = n_Y = 50$. Ambas muestras son grandes, por lo que se puede utilizar la expresión (5.16). Como consecuencia de que se quiere un intervalo de confianza de 95%, $z_{\alpha/2} = 1.96$. El intervalo de confianza de 95% para la diferencia $\mu_Y - \mu_X$ es $110.6 - 88.5 \pm 1.96 \sqrt{49.4^2/50 + 51.5^2/50}$, o 22.1 ± 19.8 .

Ejercicios para la sección 5.4

1. Se están comparando los puntos de fusión de dos aleaciones. Se fusionaron 35 ejemplares de la aleación 1. La temperatura promedio de fusión fue de 517.0°F y la desviación estándar fue de 2.4°F . Se fusionaron 47 ejemplares de la aleación 2. La temperatura promedio fue de 510.1°F y la desviación estándar fue de 2.1°F . Determine un intervalo de confianza de 99% para la diferencia entre los puntos de fusión.
2. En un experimento para determinar el efecto de la temperatura en la tasa de deposición del tungsteno sobre placas de silicio, se procesaron 64 placas a 400°C y se procesaron 88 placas a 425°C . La tasa de deposición promedio para las placas procesadas a 400°C fue de $1\,840 \text{ \AA}/\text{min}$, con una desviación estándar de $244 \text{ \AA}/\text{min}$. Las placas procesadas a 425°C promedian $2\,475 \text{ \AA}/\text{min}$, con desviación estándar de

- 760 Å/min. Determine un intervalo de confianza de 95% para la diferencia entre las medias de las tasas de deposición.
3. El artículo “Vehicle-Arrival Characteristics at Urban Uncontrolled Intersections” (V. Rengeraju y V. Rao, en *Journal of Transportation Engineering*, 1995:317-323) presenta datos de las características del tránsito en diez cruces en Madras, India. En un cruce específico, la rapidez promedio para una muestra de 39 automóviles era de 26.50 km/h, con desviación estándar de 2.37 km/h. La rapidez promedio para una muestra de 142 motocicletas era de 37.14 km/h, con desviación estándar de 3.66 km/h. Determine un intervalo de confianza de 95% para la diferencia entre las medias de la rapidez de las motocicletas y de los automóviles.
 4. Se realizó un análisis de tensión en muestras aleatorias de uniones pegadas con resina epoxídica de dos clases de madera. Una muestra aleatoria de 120 uniones de la clase A tuvo una media de tensión de corte de 1 250 psi y una desviación estándar de 350 psi, y una muestra aleatoria de 90 uniones de la clase B tuvo una media de tensión de corte de 1 400 psi y una desviación estándar de 250 psi. Encuentre un intervalo de confianza de 98% para la diferencia en las medias de la tensión de corte de las dos clases.
 5. En un estudio para comparar dos inhibidores de corrosión diferentes, se sumergieron especímenes de acero inoxidable durante cuatro horas en una disolución que contenía ácido sulfúrico y un inhibidor de corrosión. Cuarenta y siete especímenes en presencia del inhibidor A tenían media de la pérdida de peso de 242 mg y desviación estándar de 20 mg, y 42 especímenes en presencia del inhibidor B tenían media de pérdida de peso de 220 mg y desviación estándar de 31 mg. Determine un intervalo de confianza de 95% para la diferencia en la media de pérdida de peso de los dos inhibidores.
 6. Un ingeniero eléctrico desea comparar las medias de los tiempos de vida de dos tipos de transistores en una aplicación que implica un desarrollo a alta temperatura. Se probó una muestra de 60 transistores del tipo A y se encontró que tenía media de tiempos de vida de 1 827 horas y desviación estándar de 168 horas. Se probó una muestra de 180 transistores del tipo B y se encontró que tenía media de tiempos de vida de 1 658 horas y desviación estándar de 225 horas. Determine un intervalo de confianza de 95% para la diferencia entre las medias de los tiempos de vida de los dos tipos de transistores.
 7. En un estudio del efecto de tasa de enfriamiento en la dureza de uniones soldadas, se enfriaron 50 soldaduras a tasa de 10°C/s que tenían un promedio de dureza de Rockwell (B) de 91.1 y desviación estándar de 6.23, y se enfriaron 40 soldaduras a tasa de 30°C/s que tenían una media de 90.7 y desviación estándar de 4.34.
 - a) Determine un intervalo de confianza de 95% para la diferencia en la dureza entre las soldaduras enfriadas a las tasas diferentes.
 - b) Alguien dice que la tasa de enfriamiento no tiene ningún efecto sobre la dureza. ¿Estos datos contradicen dicha afirmación? Explique.
 8. Con referencia al ejercicio 7 se hacen diez soldaduras más para incrementar la precisión del intervalo de confianza. ¿En cuánto se incrementaría la precisión, enfriando las diez soldaduras a una tasa de 10°C/s, enfriando las diez soldaduras a una tasa de 30°C/s, o enfriando cinco soldaduras a 10°C/s y cinco a 30°C/s? Explique.
 9. El artículo “The Prevalence of Daytime Napping and Its Relationship to Nighttime Sleep” (J. Pilcher, K. Michalkowski, y R. Canigan), en *Behavioral Medicine*, 2001:71-76) presenta resultados de un estudio de hábitos de sueño en gran número de personas. En una muestra de 87 adultos jóvenes, el promedio del tiempo por día que pasan en cama (ya sea despiertos o dormidos) fue de 7.70 horas, con desviación estándar de 1.02 horas, y el promedio del tiempo pasado en cama durmiendo fue de 7.06 horas, con desviación estándar de 1.11 horas. La media del tiempo pasado en cama despierto se calculó de $7.70 - 7.06 = 0.64$ horas. ¿Es posible calcular un intervalo de confianza de 95% por la media del tiempo pasado en la cama despierto? Si es así, construya el intervalo de confianza. Si no es posible, explique por qué.
 10. El artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, en *Water Environment Research*, 2002:177-186) describe las mediciones de las concentraciones de amonio (en mg/l) para gran número de pozos en Iowa. Éstos incluían 349 pozos aluviales y 143 pozos cuaternarios. Las concentraciones en los pozos aluviales promediaban 0.27 con desviación estándar de 0.40 y los pozos cuaternarios promediaban 1.62 con desviación estándar de 1.70. Determine un intervalo de confianza de 95% para la diferencia en las medias de las concentraciones de los pozos aluviales y los cuaternarios.

5.5 Intervalos de confianza para la diferencia entre dos proporciones

En una población de Bernoulli, la media es igual a la probabilidad p de éxito, que es la proporción de éxitos en la población. Cuando se realizan ensayos independientes de cada una de las dos poblaciones de Bernoulli, es posible utilizar métodos similares a los que se presentó en la sección 5.4 con el fin de encontrar un intervalo de confianza para la diferencia entre las dos probabilidades de éxito. Se presenta un ejemplo para explicarlo.

Dieciocho de 60 camiones ligeros producidos en la línea de montaje A tenían un defecto en el mecanismo de dirección, que se tenían que reparar antes de enviarlos. Solamente 16 camiones de los 90 producidos en la línea de montaje B tenían este defecto. Suponga que estos camiones se pueden considerar como dos muestras aleatorias simples independientes de los camiones fabricados en las dos líneas de montaje. Se desea determinar un intervalo de confianza de 95% para la diferencia entre las proporciones de camiones con este defecto en las dos líneas de montaje.

Ésta es una situación en la que se tendría que tener cuidado de asegurarse de que es razonable considerar que los datos sean muestras aleatorias simples. Elegir camiones secuencialmente de la línea no podría ser una buena idea, por ejemplo, si hay fluctuaciones sistemáticas en la calidad en el tiempo. Se supondrá que el muestreo se ha hecho con un procedimiento bien planeado y adecuado.

Para construir el intervalo de confianza se procede en forma similar a la de la sección 5.4, al sustituir a las medias con las proporciones. Sea p_X la proporción de camiones en la población de la línea A que tenían el defecto, y sea p_Y la proporción correspondiente de la línea B. Los valores de p_X y p_Y son desconocidos. Se desea encontrar un intervalo de confianza de 95% para $p_X - p_Y$.

Sea X el número de camiones en la muestra de la línea A que tenía defectos y sea Y el número correspondiente de la línea B. Entonces X es una variable aleatoria binomial con $n_X = 60$ ensayos y probabilidad de éxito p_X , y Y es una variable aleatoria binomial con $n_Y = 90$ ensayos y probabilidad de éxito p_Y . Las proporciones muestrales son $\hat{p}_X$ y $\hat{p}_Y$. En este ejemplo los valores observados son $X = 18$, $Y = 16$, $\hat{p}_X = 18/60$ y $\hat{p}_Y = 16/90$. En virtud de que los tamaños muestrales son grandes, se tiene por el teorema del límite central que $\hat{p}_X$ y $\hat{p}_Y$ están ambas distribuidas aproximadamente en forma normal con medias p_X y p_Y y desviaciones estándares $\sigma_{\hat{p}_X} = \sqrt{p_X(1-p_X)/n_X}$ y $\sigma_{\hat{p}_Y} = \sqrt{p_Y(1-p_Y)/n_Y}$. Por consecuencia, la diferencia $\hat{p}_X - \hat{p}_Y$ tiene una distribución normal con media $p_X - p_Y$ y desviación estándar $\sqrt{p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y}$. Se concluye que para 95% de todas las muestras posibles, la diferencia $p_X - p_Y$ satisface la siguiente desigualdad:

$$\begin{aligned} \hat{p}_X - \hat{p}_Y - 1.96\sqrt{\frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}} \\ < p_X - p_Y < \\ \hat{p}_X - \hat{p}_Y + 1.96\sqrt{\frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}} \end{aligned} \quad (5.17)$$

La expresión 5.17 no es un intervalo de confianza, ya que las cantidades expresadas como: $\sqrt{p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y}$ dependen de los valores verdaderos desconocidos p_X y

p_Y . El punto de vista tradicional es sustituir p_X y p_Y con $\hat{p}_X$ y $\hat{p}_Y$, produciendo el intervalo de confianza $\hat{p}_X - \hat{p}_Y \pm z_{\alpha/2} \sqrt{\hat{p}_X(1 - \hat{p}_X)/n_X + \hat{p}_Y(1 - \hat{p}_Y)/n_Y}$. Investigaciones recientes han mostrado que este intervalo se puede mejorar modificando un poco a n_X , n_Y , $\hat{p}_X$ y $\hat{p}_Y$. Simplemente se suma 1 a cada uno de los números de éxitos de X y Y , y se suma 2 a cada uno de los números de ensayos n_X y n_Y . Por lo que, se definen $\tilde{n}_X = n_X + 2$, $\tilde{n}_Y = n_Y + 2$, $\tilde{p}_X = (X + 1)/\tilde{n}_X$, y $\tilde{p}_Y = (Y + 1)/\tilde{n}_Y$. El intervalo de confianza de 95% es $\tilde{p}_X - \tilde{p}_Y \pm z_{\alpha/2} \sqrt{\tilde{p}_X(1 - \tilde{p}_X)/\tilde{n}_X + \tilde{p}_Y(1 - \tilde{p}_Y)/\tilde{n}_Y}$. En este ejemplo, $\tilde{n}_X = 62$, $\tilde{n}_Y = 92$, $\tilde{p}_X = 19/62 = 0.3065$, y $\tilde{p}_Y = 17/92 = 0.1848$. Por tanto, se obtiene $1.3065 - 0.848 \pm 0.1395$, o $(-0.0178, 0.2612)$.

Para obtener un intervalo de confianza de nivel $100(1 - \alpha)$ el intervalo de confianza, sustituya 1.96 con $z_{\alpha/2}$. Aunque se justifica este intervalo de confianza usando el teorema del límite central, que supone que n_X y n_Y son grandes, se ha encontrado que este método da buenos resultados para casi todos los tamaños muestrales.

Resumen

Sea X el número de éxitos en n_X ensayos de Bernoulli independientes con probabilidad de éxito p_X , y sea Y el número de éxitos en n_Y ensayos de Bernoulli independientes con probabilidad de éxito p_Y , tal que $X \sim \text{Bin}(n_X, p_X)$ y $Y \sim \text{Bin}(n_Y, p_Y)$. Se define $\tilde{n}_X = n_X + 2$, $\tilde{n}_Y = n_Y + 2$, $\tilde{p}_X = (X + 1)/\tilde{n}_X$ y $\tilde{p}_Y = (Y + 1)/\tilde{n}_Y$.

Entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ para la diferencia $p_X - p_Y$ es

$$\tilde{p}_X - \tilde{p}_Y \pm z_{\alpha/2} \sqrt{\frac{\tilde{p}_X(1 - \tilde{p}_X)}{\tilde{n}_X} + \frac{\tilde{p}_Y(1 - \tilde{p}_Y)}{\tilde{n}_Y}} \quad (5.18)$$

Si el límite inferior del intervalo de confianza es menor que -1 , sustituya éste con -1 .

Si el límite superior del intervalo de confianza es mayor que 1 , sustituya éste con 1 .

El ajuste que aquí se ha descrito para el intervalo de confianza de dos muestras es similar al que se describió en la sección 5.2 para el intervalo de confianza de una muestra. En ambos casos, se ha agregado un total de dos éxitos y cuatro ensayos. Para el caso de dos muestras, éstos se han dividido entre las muestras, así que se ha agregado un éxito y dos ensayos a cada muestra. Para el caso de una muestra, se han agregado dos éxitos y cuatro ensayos a una muestra. El intervalo de confianza dado por la expresión (5.18) puede ser llamado intervalo de *Agresti-Caffo*, después de que lo desarrollaran Alan Agresti y Brian Caffo. Para mayor información acerca de este intervalo de confianza consulte el artículo “Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures” (A. Agresti y B. Caffo, en *The American Statistician*, 2000:280-288).

Ejemplo

5.24

Los métodos para calcular la fuerza y requerimientos de la dureza deben ser conservadores, de tal forma que deben sobreestimar en vez de subestimar. La tasa de éxitos de este método

se puede medir con la probabilidad de una sobreestimación. El artículo “Discrete Bracing Analysis for Light-Frame Wood-Truss Compression Webs” (M. Waltz, T. McLain y cols., en *Journal of Structural Engineering*, 2000:1086-1093) presenta los resultados de un experimento que se evaluó con un método usual (el método de Plaut) para calcular la fuerza de soporte para una red de soportes de compresión. En una muestra de 380 columnas cortas de prueba (cuatro a seis pies en longitud), el método sobreestimó la fuerza para 304 de ellas, y en una muestra de 394 columnas largas de prueba (ocho a diez pies en longitud), el método sobreestimó la fuerza para 360 de ellas. Determine un intervalo de confianza de 95% para la diferencia entre las tasas de éxito para columnas largas y columnas cortas.

Solución

El número de éxitos en la muestra de columnas cortas es $X = 304$ y el número de éxitos en la muestra de columnas largas es $Y = 360$. Los números de ensayos son $n_X = 380$ y $n_Y = 394$. Se calcula $\tilde{n}_X = 382$, $\tilde{n}_Y = 396$, $\tilde{p}_X = (304 + 1)/382 = 0.7984$, y $\tilde{p}_Y = (360 + 1)/396 = 0.9116$. El valor de $z_{\alpha/2}$ es 1.96. El intervalo de confianza de 95% es $0.9116 - 0.7984 \pm 1.96 \sqrt{(0.7984)(0.2016)/382 + (0.9116)(0.0884)/396}$, o 0.1132 ± 0.0490 .

El método tradicional

Muchas personas usan el método tradicional para calcular intervalos de confianza para la diferencia entre proporciones. Este método utiliza las proporciones muestrales $\hat{p}_X$ y $\hat{p}_Y$ y los tamaños muestrales verdaderos n_X y n_Y . El método tradicional da los resultados muy similares a los del método moderno que se acaba de describir para tamaños muestrales grandes o moderadamente grandes. Para tamaños muestrales pequeños, sin embargo, el intervalo de confianza tradicional falla para lograr que contenga la probabilidad; en otras palabras, el intervalo de confianza de nivel $100(1 - \alpha)\%$ calculado por el método tradicional contiene el valor verdadero menos del $100(1 - \alpha)\%$ de las veces.

Resumen

El método tradicional para calcular los intervalos de confianza para la diferencia entre proporciones (ampliamente usado pero no recomendado)

Sea $\hat{p}_X$ la proporción de éxitos en un número *grande* n_X de ensayos de Bernoulli independientes con probabilidad de éxito p_X , y sea $\hat{p}_Y$ la proporción de éxitos en un número *grande* n_Y de ensayos de Bernoulli independientes con probabilidad de éxito p_Y . Entonces el intervalo de confianza de nivel $100(1 - \alpha)\%$ tradicional para $p_X - p_Y$ es

$$\hat{p}_X - \hat{p}_Y \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_X(1 - \hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1 - \hat{p}_Y)}{n_Y}} \quad (5.19)$$

Este método no se puede utilizar a menos que ambas muestras contengan al menos diez éxitos y diez fracasos.

Ejercicio para la sección 5.5

- En una prueba del efecto de la humedad en conexiones eléctricas, se probaron 100 conexiones eléctricas bajo condiciones húmedas y 150 en condiciones secas. Veinte de las primeras fallaron y sólo diez de las segundas no pasaron la prueba. Determine un intervalo de confianza de 90% para la diferencia entre las proporciones de las conexiones que fallaron, húmedas y secas.
- La especificación para la resistencia de tensión de un alambre que conecta un circuito integrado a su marco es de 10 g o más. En una muestra de 85 unidades hechas con alambres de oro, 68 satisfacen la especificación, y en una muestra de 120 unidades elaboradas con alambres de aluminio, 105 cumplen el requerimiento. Determine un intervalo de confianza de 95% para la diferencia en las proporciones de unidades que satisfacen las especificaciones entre unidades con alambre de oro y aquellas con alambre de aluminio.
- En una muestra aleatoria de 340 automóviles conducidos en altitudes bajas, 46 produjeron más de 10 g de contaminación de masa particulada por galón de combustible consumido. En una muestra aleatoria de 85 automóviles conducidos en altitudes altas, 21 contaminaron en las mismas circunstancias. Determine un intervalo de confianza de 98% para la diferencia entre las proporciones para vehículos a alta y baja altitudes.
- De 1 200 piezas de grava de una planta, 110 piezas se clasificaron como “grandes”. De 900 piezas de otra planta, 95 también lo fueron. Determine un intervalo de confianza de 99% para la diferencia entre las proporciones de piezas de grava grandes producidas en las dos plantas.
- Se comparan dos procesos para fabricar cierto microchip. Se seleccionó una muestra de 400 chips de un proceso menos costoso, donde 62 estaban defectuosos. Se seleccionó una muestra de 100 chips de un proceso más costoso, pero 12 tenían defecto.
 - Determine un intervalo de confianza de 95% para la diferencia entre las proporciones de los chips defectuosos producidos por los dos procesos.
 - Con la finalidad de aumentar la precisión del intervalo de confianza, se muestrearon más chips. Se están considerando tres planes de muestreo de igual costo. En el primer plan se probarán 100 chips más del proceso menos costoso. En el segundo se observaron 50 más del proceso más costoso. En el tercer plan se probaron 50 del menos costoso y 25 del más costoso. ¿Qué plan es más probable que proporcione el aumento más grande en la precisión del intervalo de confianza? Explique.
- El artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, en *Water Environment Research*, 2002:177-186) describe mediciones de concentraciones de amonio (en mg/l) en gran número de pozos en Iowa. Éstos incluían 349 pozos aluviales y 143 cuaternarios. De los primeros, 182 tenían concentraciones arriba de 0.1, y 112 de los segundos tenían concentraciones superiores a 0.1. Determine un intervalo de confianza de 95% para la diferencia entre las proporciones de dos tipos de pozos con concentraciones arriba de 0.1.
- El artículo referido en el ejercicio 9 en la sección 5.2 describe un experimento en el que 42 muestras de lixiviado fueron probadas para la presencia de algunos contaminantes, donde 26 contenían niveles detectables de plomo y 32 de cromo. ¿Es posible utilizar los métodos de esta sección para encontrar un intervalo de confianza de 95% para la diferencia entre la probabilidad de que una muestra contendrá alguna cantidad detectable de plomo y cromo? Si es así, encuentre el intervalo de confianza. Si no, explique.
- El artículo “Case Study Based Instruction of DOE and SPC” (J. Brady y T. Allen, en *The American Statistician*, 2002:312-315) describe un esfuerzo de un equipo de ingeniería para reducir la tasa de defectos en la fabricación de un tablero específico de circuitos impresos. El equipo determinó reconfigurar el pozo de calor del transistor. Se fabricaron 1 500 tableros a la semana antes de que se implementara la reconfiguración, y 345 de éstos estaban defectuosos. Asimismo, se hicieron 1 500 tableros una semana después de la reconfiguración, y 195 de éstos tenían defectos. Determine un intervalo de confianza de 95% para el decrecimiento de la tasa de defectos después de la reconfiguración.
- Compras repetidas es una buena medida de la satisfacción del cliente. Al final de año, un distribuidor de suministros de computadora extrajo una muestra de 120 cuentas y encontró que 56 habían ordenado un pedido más de una vez. Después lo hizo con 80 cuentas del año anterior, de éstas 30 habían ordenado un pedido más de una vez. Encuentre un intervalo de confianza de 95% para la diferencia entre las dos proporciones de clientes que ordenaron más de una vez.
- El artículo “Accidents on Suburban Highways-Tennessee’s Experience” (R. Margiotta y A. Chatterjee, en *Journal of Transportation Engineering*, 1995:255-261) compara tasas de accidentes de tránsito en intersecciones con medianas

elevadas con tasas en intersecciones con doble sentido y vuelta a la izquierda. De 4 644 accidentes en el primer tipo de intersecciones, 2 280 eran en la parte trasera, y de 4 584 accidentes del segundo tipo, 1 982 presentaban esta última condición. Suponiendo que éstas sean muestras aleatorias de accidentes de dos tipos de intersecciones, determine un intervalo de confianza de 90% para la diferencia entre las proporciones de accidentes que son del tipo de parte trasera en los dos tipos de intersecciones.

$1)/(365 + 2) = 0.22$, $(63 + 1)/(365 + 2) = 0.17$, y propo-
ne calcular un intervalo de confianza de 95% para la dife-
rencia entre las proporciones de días que nieva en las dos
ciudades; así:

$$0.22 - 0.17 \pm 1.96 \sqrt{\frac{(0.22)(0.78)}{367} + \frac{(0.17)(0.83)}{367}}$$

¿Es éste un intervalo de confianza válido? Explique.

11. En cierto año, había 80 días con nevadas medibles en Den-
ver, y 63 días en Chicago. Un meteorólogo calcula $(80 +$

5.6 Intervalos de confianza para la diferencia entre dos medias con pequeñas muestras

La distribución t de Student se puede utilizar en algunos casos donde las muestras son peque-
ñas y, por tanto, donde no es aplicable el teorema del límite central. Se presenta un ejemplo.

Una muestra de seis soldaduras de un tipo tenía promedio de prueba final de resisten-
cia (en ksi) de 83.2 y desviación estándar de 5.2, y una muestra de diez soldaduras de otro ti-
po tenía resistencia promedio de 71.3 y desviación estándar de 3.1. Suponga que ambos
conjuntos de soldaduras son muestras aleatorias de poblaciones normales. Se desea encontrar
un intervalo de confianza de 95% para la diferencia entre las medias de las resistencias de los
dos tipos de soldaduras.

Ambos tamaños muestrales son pequeños, por lo que no es aplicable el teorema del lí-
mite central. Si ambas poblaciones son normales, la distribución t de Student se puede utilizar
para calcular un intervalo de confianza para la diferencia entre las dos medias poblacionales.
El método es similar al que se presenta en la sección 5.4 para el caso donde las muestras son
grandes, excepto que el puntaje z se sustituye con un valor de la distribución t de Student.

Si $X_1, \dots, X_{n_X}$ es una muestra de tamaño n_X de una población *normal* con media μ_X y
 $Y_1, \dots, Y_{n_Y}$ es una muestra de tamaño n_Y de una población *normal* con media μ_Y , entonces la
cantidad

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

tiene una aproximada distribución t de Student.

El número de grados de libertad a usar para esta distribución está dado por

$$v = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{(s_X^2/n_X)^2}{n_X - 1} + \frac{(s_Y^2/n_Y)^2}{n_Y - 1}} \quad \text{redondeado hacia abajo al entero más cercano.} \quad (5.20)$$

En este ejemplo, sea $X_1, \dots, X_6$ de las seis soldaduras del primer tipo y sea $Y_1, \dots, Y_{10}$ las diez soldaduras del segundo tipo. Sustituyendo $s_X = 5.2$, $s_Y = 3.1$, $n_X = 6$, $n_Y = 10$ en la ecuación (5.20) se obtiene

$$v = \frac{\left(\frac{5.2^2}{6} + \frac{3.1^2}{10}\right)^2}{\frac{(5.2^2/6)^2}{5} + \frac{(3.1^2/10)^2}{9}} = 7.18 \approx 7$$

Si ambas poblaciones son normales, entonces la cantidad

$$[(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)] / \sqrt{s_X^2/6 + s_Y^2/10}$$

tiene una distribución aproximada t de Student con siete grados de libertad. La figura 5.14 presenta esta distribución. El 95% del área bajo la curva está contenida entre los valores $t = -2.365$ y $t = 2.365$. En consecuencia 95% de todas las muestras que se puede haber elegido,

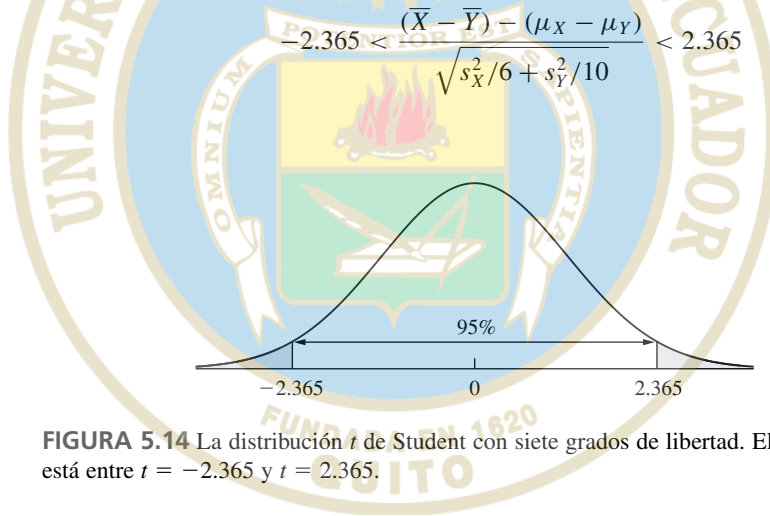


FIGURA 5.14 La distribución t de Student con siete grados de libertad. El 95% del área está entre $t = -2.365$ y $t = 2.365$.

Mediante el razonamiento utilizado en la sección 5.3, un intervalo de confianza de 95% para la diferencia $\mu_X - \mu_Y$ es $\bar{X} - \bar{Y} \pm 2.365 \sqrt{s_X^2/6 + s_Y^2/10}$. Al sustituir $\bar{X} = 83.2$, $\bar{Y} = 71.3$, $s_X = 5.2$ y $s_Y = 3.1$, se encuentra que un intervalo de confianza de 95% para $\mu_X - \mu_Y$ es 11.9 ± 5.53 , o $(6.37, 17.43)$.

En general, para producir un intervalo de confianza de $100(1 - \alpha)\%$, sea $t_{v, \alpha/2}$ el cuantil $1 - \alpha/2$ de la distribución t de Student con v grados de libertad, éste es el valor que corta un área de $\alpha/2$ en la cola del lado derecho. Por ejemplo, antes se encontró que $t_{7, 0.025} = 2.365$. Después un intervalo de confianza de nivel $100(1 - \alpha)\%$ para la diferencia entre las medias poblacionales $\mu_X - \mu_Y$, cuando los tamaños muestrales son n_X y n_Y , respectivamente, $\bar{X} - \bar{Y} \pm t_{v, \alpha/2} \sqrt{s_X^2/n_X + s_Y^2/n_Y}$.

Resumen

Sea $X_1, \dots, X_{n_X}$ una muestra aleatoria de tamaño n_X de una población *normal* con media μ_X , y $Y_1, \dots, Y_{n_Y}$ una muestra aleatoria de tamaño n_Y de una población *normal* con media μ_Y . Suponga que las dos muestras son independientes.

Si las poblaciones no necesariamente tienen la misma varianza, un intervalo de confianza de nivel $100(1 - \alpha)\%$ para $\mu_X - \mu_Y$ es

$$\bar{X} - \bar{Y} \pm t_{v, \alpha/2} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}} \quad (5.21)$$

El número de grados de libertad, v , está dado por

$$v = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{(s_X^2/n_X)^2}{n_X - 1} + \frac{(s_Y^2/n_Y)^2}{n_Y - 1}} \text{ redondeado hacia abajo al entero más cercano.}$$

Ejemplo 5.25

Los compuestos con base de resina se usan en odontología restauradora. El artículo “Reduction of Polymerization Shrinkage Stress and Marginal Leakage Using Soft-Start Polymerization” (C. Ernst, N. Brand y colaboradores, en *Journal of Esthetic and Restorative Dentistry*, 2003:93-104) presenta una comparación de la dureza de la superficie de las muestras curadas durante 40 segundos con fuerza constante con otra de 40 segundos con la fuerza aumentando en forma exponencial. Quince muestras fueron curadas con cada método. Las curadas con fuerza constante tuvieron promedio de dureza de superficie (en N/mm²) de 400.9 con desviación estándar de 10.6. Las curadas con fuerza que aumenta exponencialmente tenían una dureza de superficie promedio de 367.2 con desviación estándar de 6.1. Determine un intervalo de confianza de 98% para la diferencia en las medias de la dureza entre las muestras curadas por los dos métodos.

Solución

Se tiene $\bar{X} = 400.9$, $s_X = 10.6$, $n_X = 15$, $\bar{Y} = 367.2$, $s_Y = 6.1$ y $n_Y = 15$. El número de grados de libertad está dado por la ecuación (5.20) como

$$v = \frac{\left(\frac{10.6^2}{15} + \frac{6.1^2}{15}\right)^2}{\frac{(10.6^2/15)^2}{15 - 1} + \frac{(6.1^2/15)^2}{15 - 1}} = 22.36 \approx 22$$

De la tabla t (tabla A.3 en el Apéndice A) se encuentra que $t_{22, 0.01} = 2.508$. Se utiliza la expresión (5.21) para encontrar que el intervalo de confianza de 98% es

$$400.9 - 367.2 \pm 2.508 \sqrt{10.6^2/15 + 6.1^2/15}, \text{ o } 33.7 \pm 7.9.$$

Cuando las poblaciones tienen varianzas iguales

Ocasionalmente se presenta una situación en la que se sabe que dos poblaciones tienen varianzas casi iguales. En estos casos un método que explota este hecho se puede utilizar para producir intervalos de confianza que son más angostos que los producidos por el método más general descrito antes. Al igual que con el método más general, ambas poblaciones deben ser aproximadamente normales. Se presenta un ejemplo.

Dos pesas usuales, cada una etiquetada con 100 g, son pesadas varias veces en la misma balanza. La primera se pesa durante ocho veces, y la media de la lectura de la balanza es 18.2 μg arriba de 100 g, con desviación estándar de 2.0 μg . La segunda se pesa en 18 ocasiones y la media de las lecturas es 16.4 g arriba de 100 g, con desviación estándar de 1.8 μg . Suponga que cada conjunto de lecturas es una muestra de una población aproximadamente normal. Debido a que se usa la misma balanza para todas las mediciones, y dado que los pesos verdaderos son aproximadamente iguales, es razonable suponer que las desviaciones estándares poblacionales de las lecturas son las mismas para ambas pesas. Se supone que las mediciones no están sesgadas (es, en realidad, suficiente suponer que el sesgamiento es el mismo para ambas pesas). Se desea encontrar un intervalo de confianza de 95% para la diferencia entre los pesos verdaderos.

Sean $X_1, \dots, X_8$ las lecturas para la primera pesa, y $Y_1, \dots, Y_{18}$ las lecturas para la segunda. Sean μ_X y μ_Y los pesos verdaderos, que son las medias de las poblaciones de las cuales se extrajeron estas muestras. Se supone que ambas poblaciones siguen distribuciones normales con la misma varianza σ^2 . Por tanto, $\bar{X}$ tiene una distribución normal con media μ_X y varianza $\sigma^2/8$, y $\bar{Y}$ tiene una distribución normal con media μ_Y y varianza $\sigma^2/18$. Por tanto, la diferencia $\bar{X} - \bar{Y}$ tiene una distribución normal con media $\mu_X - \mu_Y$ y la varianza $\sigma^2(1/8 + 1/18)$. Dado que σ^2 no se conoce, se debe estimar. Se podría estimar esta cantidad utilizando cualesquiera de las varianzas muestrales $s_X^2 = \sum_{i=1}^8 (X_i - \bar{X})^2 / (8 - 1)$ o $s_Y^2 = \sum_{i=1}^{18} (Y_i - \bar{Y})^2 / (18 - 1)$. Pero la mejor estimación se obtiene al combinar la información en ambas muestras. La mejor estimación es la **varianza combinada** $s_p^2 = (7s_X^2 + 17s_Y^2) / (7 + 17)$. La varianza combinada s_p^2 es un promedio ponderado de dos varianzas muestrales. Los pesos son iguales a los tamaños muestrales menos uno. Es lógico utilizar un promedio ponderado con el propósito de que la varianza muestral basada en la muestra más grande cuente más. Con los valores dados para s_X y s_Y , el valor de la varianza combinada es $s_p^2 = [7(2.0^2) + 17(1.8^2)] / (7 + 17) = 3.4617$, así $s_p = 1.8606$.

La cantidad $[(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)] / (s_p \sqrt{1/8 + 1/18})$ tiene la distribución t de Student con $8 + 18 - 2 = 24$ grados de libertad. De la tabla t , se encuentra que $t_{24, 0.025} = 2.064$. Por lo que para 95% de todas las muestras que se puede haber elegido,

$$-2.064 < \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{s_p \sqrt{1/8 + 1/18}} < 2.064$$

Por el razonamiento utilizado en la sección 5.3, un intervalo de confianza de 95% para $\mu_X - \mu_Y$ es $\bar{X} - \bar{Y} \pm 2.064 s_p \sqrt{1/8 + 1/18}$. Al sustituir $\bar{X} = 18.2$, $\bar{Y} = 16.4$ y $s_p = 1.8606$ encuentra un intervalo de confianza de 95% para $\mu_X - \mu_Y$ es 1.8 ± 1.6318 , o $(0.1682, 3.4318)$.

Resumen

Sea $X_1, \dots, X_{n_X}$ una muestra aleatoria de tamaño n_X de una población *normal* con media μ_X y $Y_1, \dots, Y_{n_Y}$ una muestra aleatoria de tamaño n_Y de una población *normal* con media μ_Y . Se supone que las dos muestras son independientes.

Si se sabe que las poblaciones tienen casi la misma varianza, un intervalo de confianza de nivel $100(1 - \alpha)\%$ para $\mu_X - \mu_Y$ es

$$\bar{X} - \bar{Y} \pm t_{n_X + n_Y - 2, \alpha/2} \cdot s_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}} \quad (5.22)$$

La cantidad s_p^2 es la varianza combinada, dada por

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2} \quad (5.23)$$

Ejemplo 5.26

Se utilizó una máquina para llenar botellas de plástico con blanqueador. Una muestra de 18 botellas tenía una media de volumen de llenado de 2.007 l y desviación estándar de 0.010 l. Después el aparato se cambió de lugar a otra ubicación. Una muestra de diez botellas llenadas en la nueva ubicación tenía una media del volumen llenado de 2.001 l y desviación estándar de 0.012 l. Se creía que cambiar de lugar la máquina podría haber cambiado la media del volumen llenado, pero es poco probable haber cambiado la desviación estándar. Suponga que ambas muestras provienen de poblaciones aproximadamente normales. Determine un intervalo de confianza de 99% para la diferencia entre la media de los volúmenes de llenado en las dos ubicaciones.

Solución

Se tiene $\bar{X} = 2.007$, $s_X = 0.010$, $n_X = 18$, $\bar{Y} = 2.001$, $s_Y = 0.012$ y $n_Y = 10$. Debido a que se cree que las desviaciones estándares poblacionales son iguales, se estima su valor común con la desviación estándar combinada, utilizando la ecuación (5.23). Se obtiene

$$s_p = \sqrt{\frac{(18 - 1)(0.010^2) + (10 - 1)(0.012^2)}{18 + 10 - 2}} = 0.0107$$

El número de grados de libertad es $18 + 10 - 2 = 26$. Se utiliza la expresión (5.22) para determinar el intervalo de confianza de 99%. Consultando la tabla t con 26 grados de libertad, se encuentra que $t_{26, 0.005} = 2.779$. El intervalo de confianza de 99% es, por tanto,

$$2.007 - 2.001 \pm 2.779(0.0107)\sqrt{1/18 + 1/10}, \text{ o } 0.006 \pm 0.012.$$

No suponga que las varianzas poblacionales son exactamente iguales sólo porque las varianzas muestrales están cercanas

El intervalo de confianza dado por la expresión (5.22) requiere que las varianzas poblacionales sean iguales, o casi iguales. En situaciones donde las varianzas *muestrales* son casi igua-

les, es tentador suponer que las varianzas poblacionales son también casi iguales. Sin embargo, cuando los tamaños muestrales son pequeños, las varianzas muestrales no son necesariamente buenas aproximaciones a las varianzas poblacionales. Por tanto, es posible que las varianzas muestrales estén cercanas incluso cuando las varianzas poblacionales estén bastante alejadas. En general, se debe suponer que las varianzas poblacionales son iguales solamente cuando se tiene conocimiento acerca de los procesos que produjeron los datos que justifiquen esta suposición.

El intervalo de confianza dado por la expresión (5.21) produce buenos resultados en casi todos los casos, si las varianzas poblacionales son iguales o no. (Puede haber excepciones cuando las muestras son de tamaños muy diferentes.) Por tanto, cuando se duda acerca de si las varianzas poblacionales son iguales, utilice la expresión (5.21).

Ejercicios para la sección 5.6

1. Se midió cinco veces el contenido de carbono (en partes por millón) para cada una de las dos placas de silicio diferentes. Las mediciones fueron:

Placa A:	1.10	1.15	1.16	1.10	1.14
Placa B:	1.20	1.18	1.16	1.18	1.15

Determine un intervalo de confianza de 99% para la diferencia en contenido de carbono entre las dos placas.

2. En un estudio de la tasa en donde el pesticida hexaconazol es absorbido a través de piel, se expusieron muestras de piel a 24 μg de hexaconazol. Se expusieron cuatro muestras durante 30 minutos y otras cuatro fueron expuestas durante 60 minutos. Las cantidades (en μg) que fueron absorbidas fueron

30 minutos:	3.1	3.3	3.4	3.0
60 minutos:	3.7	3.6	3.7	3.4

Determine un intervalo de confianza de 95% para la media de la cantidad absorbida en el intervalo entre 30 y 60 minutos después de la exposición.

3. El artículo "Differences in Susceptibilities of Different Cell Lines to Bilirubin Damage" (K. Ngai, C. Yeung, y C. Leung, en *Journal of Pediatric Child Health*, 2000:36-45) describe una investigación respecto de la toxicidad de bilirrubina sobre algunas líneas de células. Diez conjuntos de células de hígado humanos y diez conjuntos de células de fibroblastos de ratón fueron puestos en soluciones de bilirrubina en albúmina a una proporción de 1.4 bilirrubina/albúmina molar durante 24 horas. En los diez conjuntos de celdas de hígado humanas, el porcentaje promedio de células que sobrevivían era de 53.9 con desviación estándar de 10.7. En los diez conjuntos de celdas de fibroblastos de ra-

tón, el porcentaje promedio de células que sobrevivían era de 73.1 con desviación estándar de 9.1. Determine un intervalo de confianza de 98% para la diferencia de porcentajes de supervivencia entre las dos líneas de células.

4. Un médico genetista está estudiando la frecuencia de cierta mutación genética en hombres de diferentes edades. Para diez hombres de 20-29 años, la media del número de secuencias mutantes por μg de ADN era 47.2 y la desviación estándar 15.1. Para 12 hombres de 60-69 años la media era 109.5 y la desviación estándar 31.2. Determine un intervalo de confianza de 99% para la diferencia en la media de la frecuencia mutante entre hombres de 20-29 años y otros de 60-69.

5. Durante el verano de 1999 los bombardeos destruyeron muchas instalaciones de almacenamiento de combustible en Serbia. Por consiguiente, importantes cantidades de productos de aceite fueron derramadas y quemadas, dando como resultado contaminación del suelo. El artículo "Mobility of Heavy Metals Originating from Bombing of Industrial Sites" (B. Škrbić, J. Novaković, y N. Miljević, en *Journal of Environmental Science and Health*, 2002:7-16) notifica mediciones de concentraciones de metal pesado en algunos sitios industriales en junio de 1999, justo después del bombardeo, y nuevamente en marzo de 2000. En el sitio de Smederevo, en la rivera del Danubio, se tomaron ocho muestras de tierra en 1999, que tenían una concentración promedio de plomo (en mg/kg) de 10.7 con una desviación estándar de 3.3. Cuatro muestras tomadas en 2000 tuvieron una concentración promedio de plomo de 33.8 con desviación estándar de 0.50. Determine un intervalo de confianza de 95% para el aumento en la concentración de plomo entre junio de 1999 y marzo de 2000.

6. El artículo "Quality of the Fire Clay Coal Bed. Southeastern Kentucky" (J. Hower, W. Andrews y cols., en *Journal of Coal Quality*, 1994:13-26) contiene mediciones acerca de muestras de carbón en algunos condados de Kentucky. En unidades de porcentajes de ceniza, cinco muestras del condado de Knott tenían un promedio de dióxido de aluminio (AlO_2) de 32.17 y desviación estándar de 2.23. Seis muestras de condado de Leslie tenían un promedio de contenido de AlO_2 de 26.48 y desviación estándar de 2.02. Determine un intervalo de confianza de 98% para la diferencia en el contenido de AlO_2 contenido entre las muestras de carbón de los dos condados.
10. El artículo "Permeability, Diffusion and Solubility of Gases" (B. Flaconnèche y colaboradores, en *Oil and Gas Science and Technology*, 2001:262-278) describe un estudio del efecto de la temperatura y otros factores en los coeficientes de transporte de gas en polímeros semicristalinos. El coeficiente de permeabilidad (en 10^{-6} (STP)/cm·s·MPa) de CO_2 fue medido para la densidad media extrudida de polietileno tanto a 60°C como 61°C. Los resultados fueron:

60°C:	54	51	61	67	57	69	60
	60	63	62				
61°C:	58	60	66	66	68	61	60

7. El artículo "The Frequency Distribution of Daily Global Irradiation at Kumasi" (F. Akuffo y A. Brew-Hammond, en *Solar Energy*, 1993:145-154) define el índice de claridad diaria para una ubicación que está en la proporción de irradiación global con la irradiación extraterrestre. Se tomó mediciones en la ciudad de Ibadan, Nigeria, en un periodo de cinco años. Para cinco meses de mayo, el promedio del índice de claridad era de 0.498 y tenía 0.036 de desviación estándar. Durante cinco meses de julio, el promedio era de 0.389 y la desviación estándar de 0.049. Encuentre un intervalo de confianza de 95% para la diferencia entre las medias de los índices de claridad de mayo y julio.
11. Una administradora de sistemas computacionales observa que las computadoras que corren en un sistema operativo especial parecen paralizarse más a menudo conforme pasa el tiempo desde la instalación del sistema operativo. Ella mide el tiempo (en minutos) antes de que se paralice para siete computadoras un mes después de la instalación, y para nueve computadoras siete meses después. Los resultados fueron:

Un mes después de la instalación:	207.4	233.1	215.9
	235.1	225.6	244.4
	245.3		
Siete meses después de la instalación:	84.3	53.2	127.3
	201.3	174.2	246.2
	149.4	156.4	103.3

8. Se probaron algunas muestras de carbón de cada una de dos minas, y se midió la capacidad calorífica (en kilocalorías por libra) para cada muestra. Los resultados fueron:

Mina 1:	4 167	4 268	4 159	4 285	4 229
	4 386	4 103			
Mina 2:	3 924	3 988	4 096	4 026	
	4 235	4 178			

Determine un intervalo de confianza de 90% para la diferencia en la capacidad calorífica entre carbón de la mina 1 y carbón de la mina 2.

9. La fuerza de ruptura de palos de hockey de dos compuestos de grafito-Kevlar diferentes producen los siguientes resultados (en newtons):

Compuesto A:	487.3	444.5	467.7	456.3	449.7
	459.2	478.9	461.5	477.2	
Compuesto B:	488.5	501.2	475.3	467.2	462.5
	499.7	470.0	469.5	481.5	485.2
	509.3	479.3	478.3	491.5	

Determine un intervalo de confianza de 98% para la diferencia entre las medias de la fuerza de ruptura de palos de hockey hechos de los dos materiales.

Determine un intervalo de confianza de 95% para la diferencia de las medias en el tiempo en que se paraliza entre el primero y el séptimo meses.

12. En el artículo "Bactericidal Properties of Flat Surfaces and Nanoparticles Derivatized with Alkylated Polyethylenimines" (J. Lin, S. Qiu y colaboradores, en *Biotechnology Progress*, 2002:1082-1086), se describen experimentos en los que se fijó polietileniminas alquiladas a superficies y a nanopartículas para hacerlas bactericidas. En una serie de experimentos, la eficiencia bactericida contra la bacteria *E. coli* fue comparada para un metilado contra un polímero no metilado. La media del porcentaje de células de bacterias muertas con el polímero metilado era de 95 con una desviación estándar de 1, y la media del porcentaje de células de bacterias muertas con el polímero no metilado era de 70 con una desviación estándar 6. Suponga que se hizo cinco mediciones independientes en cada tipo de polímero. Determine un intervalo de confianza de 95% para el aumento en la eficiencia bactericida del polímero metilado.

13. Una muestra de ocho acondicionadores de aire para habitación de cierto modelo tenía una media de presión de sonido de media de 52 decibeles (dB) y una desviación estándar de cinco decibeles, y una muestra de 12 acondicionadores de aire de un modelo diferente tenía una media de presión de so-

nido de 46 decibeles y desviación estándar de 2 decibeles. Determine un intervalo de confianza de 98% para la diferencia en las medias de presión de sonido entre los dos modelos.

5.7 Intervalos de confianza con datos apareados

Los métodos analizados hasta ahora para encontrar intervalos de confianza con base en dos muestras han requerido que las muestras sean independientes. En algunos casos, es mejor diseñar un experimento con el propósito de que cada elemento en una muestra se empareje con un elemento en la otra. A continuación se muestra un ejemplo.

Un fabricante de neumáticos desea comparar el desgaste de la huella de los neumáticos hechos de un nuevo material con el de los neumáticos de un material convencional. Un neumático de cada tipo se coloca en cada rueda delantera de cada uno de diez automóviles de unidad de disco de rueda delantera. La elección respecto a qué tipo de neumático va en la rueda derecha y cuál en la izquierda se hace lanzando al aire una moneda. Cada automóvil se conduce durante 40 000 millas, entonces se quita los neumáticos y se mide la profundidad de la huella de cada uno. En la figura 5.15 se presenta los resultados.

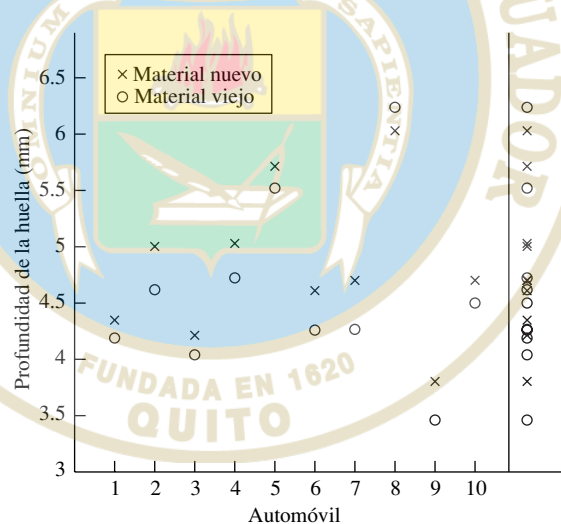


FIGURA 5.15 Profundidad de la huella para diez pares de neumáticos.

La columna del lado derecho de la figura 5.15 presenta los resultados para los 20 neumáticos. Hay coincidencia considerable en el desgaste de la huella para las dos muestras. Es difícil decir de la columna si hay una diferencia entre las clases vieja y nueva de neumático. Sin embargo, cuando los datos se revisan en pares, está claro que, en general, los neumáticos del nuevo tipo tienen más huella que los de la vieja clase. La razón de analizar los pares es presentar un esquema más claro del resultado, que los automóviles varían mucho en cuanto

al desgaste que tienen. Automóviles pesados, y los que tienen patrones de manejo que implican muchos arranques y paradas, generalmente producen más desgaste que otros. Los datos agregados en la columna de la derecha de la figura incluyen esta variabilidad entre los automóviles, así como variabilidad en el desgaste de las llantas. Cuando los datos se consideran en pares, la variabilidad entre los automóviles desaparece, porque ambas llantas en un par provienen del mismo automóvil.

La tabla 5.1 presenta, para cada automóvil, las profundidades de las huellas, así como la diferencia entre ellas. Se desea encontrar un intervalo de confianza de 95% para la media de la diferencia en el desgaste de la huella entre materiales viejos y nuevos en una forma que resulta ventajosa para reducir la variabilidad producida por el diseño apareado. La forma de hacer esto último es pensar en una población de pares de valores, en la cual cada par consiste de mediciones de un tipo viejo y de un nuevo tipo de neumático en el mismo automóvil. Para cada par en la población, hay una diferencia (nuevo-viejo), por lo que hay una población de diferencias. Los datos constituyen, entonces, una muestra aleatoria poblacional de pares y sus diferencias representan una muestra aleatoria poblacional de diferencias.

TABLA 5.1 Profundidades de la huella, en mm, para neumáticos hechos de materiales nuevos y viejos

	Automóvil									
	1	2	3	4	5	6	7	8	9	10
Material nuevo	4.35	5.00	4.21	5.03	5.71	4.61	4.70	6.03	3.80	4.70
Material viejo	4.19	4.62	4.04	4.72	5.52	4.26	4.27	6.24	3.46	4.50
Diferencia	0.16	0.38	0.17	0.31	0.19	0.35	0.43	-0.21	0.34	0.20

Con el propósito de poner lo anterior en notación estadística, sea $(X_1, Y_1), \dots, (X_{10}, Y_{10})$ los diez pares observados, con X_i representando la huella del neumático hecho del nuevo material del i -ésimo automóvil y Y_i representando la huella del neumático hecho del material viejo del i -ésimo automóvil. Sea $D_i = X_i - Y_i$ la diferencia entre las huellas para los neumáticos del i -ésimo automóvil. Sean μ_X y μ_Y las medias poblacional para X y Y , respectivamente. Se desea encontrar un intervalo de confianza de 95% para la diferencia $\mu_X - \mu_Y$. Sea μ_D la media poblacional de diferencias. Entonces $\mu_D = \mu_X - \mu_Y$. En consecuencia, un intervalo de confianza para μ_D también será un intervalo de confianza para $\mu_X - \mu_Y$.

Dado que la muestra $D_1, \dots, D_{10}$ es una muestra aleatoria de una población con media μ_D , es posible utilizar métodos para encontrar intervalos de confianza para μ_D . En este ejemplo, puesto que el tamaño muestral es pequeño, se usa el método t de Student de la sección 5.3. Los valores observados de la media muestral y la desviación estándar muestral son

$$\bar{D} = 0.232 \quad s_D = 0.183$$

El tamaño muestral es diez, por lo que hay nueve grados de libertad. El valor adecuado t es $t_{9, 0.025} = 2.262$. Por tanto, con la expresión (5.9) (de la sección 5.3) el intervalo de confianza es $0.232 \pm (2.262)(0.183)/\sqrt{10}$, o $(0.101, 0.363)$. Cuando el número de pares es grande, los métodos de grandes muestras de la sección 5.1, específicamente la expresión (5.1), se pueden utilizar.

Resumen

Sea $D_1, \dots, D_n$ una muestra aleatoria pequeña ($n \leq 30$) de diferencias de pares. Si la población de diferencias es aproximadamente normal, entonces un intervalo de confianza de nivel $100(1 - \alpha)\%$ para la media de la diferencia μ_D está dado por

$$\bar{D} \pm t_{n-1, \alpha/2} \frac{s_D}{\sqrt{n}} \tag{5.24}$$

Observe que este intervalo es el mismo que el que se obtiene con la expresión (5.9).

Si el tamaño muestral es grande, un intervalo de confianza de nivel $100(1 - \alpha)\%$ para la media de la diferencia μ_D está dado por

$$\bar{D} \pm z_{\alpha/2} \sigma_{\bar{D}} \tag{5.25}$$

En la práctica $\sigma_{\bar{D}}$ se aproxima con $s_D/\sqrt{n}$. Observe que este intervalo es el mismo que el que se obtiene con la expresión (5.1).

Ejercicios para la sección 5.7

1. El artículo “Simulation of the Hot Carbonate Process for Removal of CO₂ and H₂S from Medium Btu Gas” (K. Park y T. Edgar, en *Energy Progress*, 1984:174-180) presenta una ecuación que utilizó para calcular la presión de vapor en equilibrio del CO₂ en una solución de carbonato de potasio. Se midió la presión de equilibrio real (en kPa) en nueve reacciones diferentes y se comparó con el valor calculado por una ecuación. Los resultados se presentan en la tabla siguiente:
2. El artículo “Effect of Refrigeration on the Potassium Bitartrate Stability and Composition of Italian Wines” (A. Versan, D. Barbanti y colaboradores, en *Italian Journal of Food Science* 2002:45-52) notifica un estudio en el que los ocho tipos de vino blanco tenían su concentración de ácido tartárico (en g/l), medido tanto antes como después de un proceso de estabilización frío. Los resultados se presentan en la tabla siguiente:

Reacción	Estimado	Experimental	Diferencia	Tipo de vino	Antes	Después	Diferencia
1	45.10	42.95	2.15	1	2.86	2.59	0.27
2	85.77	79.98	5.79	2	2.85	2.47	0.38
3	151.84	146.17	5.67	3	1.84	1.58	0.26
4	244.30	228.22	16.08	4	1.60	1.56	0.04
5	257.67	240.63	17.04	5	0.80	0.78	0.02
6	44.32	41.99	2.33	6	0.89	0.66	0.23
7	84.41	82.05	2.36	7	2.03	1.87	0.16
8	150.47	149.62	0.85	8	1.90	1.71	0.19
9	253.81	245.45	8.36				

Determine un intervalo de confianza de 95% para la diferencia de medias entre las presiones estimadas y reales.

Determine un intervalo de confianza de 95% para la diferencia de medias entre las concentraciones de ácido tartárico antes y después del proceso de estabilización frío.

3. En un experimento para determinar si hay una diferencia sistemática entre los pesos obtenidos con dos balanzas diferentes, se pesaron diez ejemplares de rocas, en gramos, en cada balanza. Se obtuvieron los siguientes datos:

Ejemplar	Peso en la balanza 1	Peso en la balanza 2
1	11.23	11.27
2	14.36	14.41
3	8.33	8.35
4	10.50	10.52
5	23.42	23.41
6	9.15	9.17
7	13.47	13.52
8	6.47	6.46
9	12.40	12.45
10	19.38	19.35

Suponga que la diferencia entre las balanzas, si es que hay alguna, no depende del objeto pesado. Determine un intervalo de confianza de 98% para esta diferencia.

4. En un tipo específico de motor, el volante está sujeto al bloque de brida del cigüeñal mediante ocho pernos, que se deben apretar en secuencia. Un ingeniero quiere determinar qué diferencia existe entre el par de torsión del primero y el último perno que se apretó. En una muestra de siete motores, se mide el par de torsión en cada uno de estos pernos. Los siguientes son los valores, en N-m, para los pernos 1 y 8.

Motor	Perno 1	Perno 8
1	105.28	105.38
2	105.03	105.96
3	104.53	105.19
4	104.68	105.61
5	104.98	105.42
6	105.21	105.20
7	105.30	105.98

Determine un intervalo de confianza de 90% para la diferencia entre los pares de torsión.

5. En un estudio de tiempo de disolución de diferentes dulces, nueve personas disolvieron una pieza de chocolate y una pieza de dulce de azúcar y mantequilla. Los tiempos de disolución, en segundos, se presentan en la tabla siguiente:

Persona	Chocolate	Dulce de azúcar y mantequilla
1	51	53
2	47	40
3	90	155
4	65	90
5	27	33
6	105	68
7	90	72
8	54	52
9	93	77

Determine un intervalo de confianza de 98% para la diferencia entre la media de tiempos de disolución de chocolate y caramelo de azúcar y mantequilla.

6. Una muestra de diez camiones diesel fue operada tanto caliente como fría para calcular la diferencia en el ahorro de combustible. Los resultados, en milla/galón, se presentan en la tabla siguiente. (De "In-use Emissions from Heavy-Duty Diesel Vehicles," J. Yanowitz, tesis de doctorado, Escuela de Minas, de Colorado, 2001.)

Camión	Caliente	Frío
1	4.56	4.26
2	4.46	4.08
3	6.49	5.83
4	5.37	4.96
5	6.25	5.87
6	5.90	5.32
7	4.12	3.92
8	3.85	3.69
9	4.15	3.74
10	4.69	4.19

Determine un intervalo de confianza de 98% para la diferencia en la media del millaje de combustible entre motores calientes y fríos.

7. Para una muestra de nueve automóviles, se mide el millaje (en mil millas) de los patines de frenos frontales originales que se han desgastado 10% de su espesor original, así como el millaje de los patines de los frenos traseros originales que se han desgastado 10% de su espesor original. Los resultados están dados en la tabla siguiente:

Automóvil	Delanteros	Traseros
1	32.8	41.2
2	26.6	35.2
3	35.6	46.1
4	36.4	46.0
5	29.2	39.9
6	40.9	51.7
7	40.9	51.6
8	34.8	46.1
9	36.6	47.3

Automóvil	Huella A	Huella B
1	24.1	20.3
2	22.3	19.7
3	24.5	22.5
4	26.1	23.2
5	22.6	20.4
6	23.3	23.5
7	22.4	21.9
8	19.9	18.6
9	27.1	25.8
10	23.5	21.4
11	25.4	20.6
12	24.9	23.4
13	23.7	20.3
14	23.9	22.5
15	24.6	23.5
16	26.4	24.5
17	21.5	22.4
18	24.6	24.9

Determine un intervalo de confianza de 95% para la diferencia en la media de tiempos de vida entre los patines de frenos delanteros y traseros.

8. Con referencia al ejercicio 7, alguien sugiere que el diseño emparejado sea reemplazado con un diseño en el que se seleccionan 18 automóviles, se mide el tiempo de vida de los frenos delanteros en nueve de ellos, y se mide el tiempo de vida de los frenos traseros en los otros nueve. Después se construye un intervalo de confianza para la diferencia entre las medias usando la expresión (5.21) (de la sección 5.6). Él afirma que este diseño producirá un intervalo de confianza más preciso, ya que se utilizará 18 automóviles en lugar de nueve.

- a) ¿El nuevo diseño producirá un intervalo de confianza válido? Explique.
- b) ¿Es probable que el intervalo de confianza producido por el nuevo diseño será más preciso que, menos preciso que, o casi con la misma precisión que el intervalo de confianza producido por el diseño apareado? Explique. (Sugerencia: busque en la figura 5.15.)

9. Un fabricante de neumáticos está interesado en probar el ahorro de combustible para dos patrones de huellas diferentes. Los neumáticos de cada tipo de huella son conducidos durante mil millas en cada uno de 18 automóviles diferentes. Los millajes, en milla/galón, se presentan en la tabla siguiente:

- a) Determine un intervalo de confianza de 99% para la media de la diferencia en el ahorro de combustible.
- b) Un intervalo de confianza basado en los datos de la tabla tiene un ancho ± 0.5 milla/galón. ¿El nivel de este intervalo de confianza está más cerca de 80%, de 90% o de 95%?

10. Con referencia al ejercicio 9, en un experimento distinto se equipararon 18 automóviles con neumáticos con huellas del tipo A, y 18 del tipo B. Cada automóvil se condujo mil millas. Los automóviles con huellas del tipo A promediaban 23.93 milla/galón, con desviación estándar de 1.79 milla/galón. Los automóviles con huellas de tipo B promediaban 22.19 milla/galón, con desviación estándar de 1.95 milla/galón.

- a) ¿Qué método se debe usar para determinar un intervalo de confianza para la diferencia entre las medias de los millajes de dos tipos de huellas: la expresión (5.24) (de esta sección) o la expresión (5.21) (de la sección 5.6)?
- b) Usando el método adecuado, determine un intervalo de confianza de 99% para la diferencia entre las medias de los millajes de los dos tipos de huella.
- c) ¿Es el intervalo de confianza encontrado en el inciso b) más ancho que el encontrado en el ejercicio 9? ¿Por qué esto es así?

5.8 Uso de simulación para construir intervalos de confianza

Si $X_1, \dots, X_n$ son variables aleatorias independientes con desviaciones estándares conocidas, $\sigma_1, \dots, \sigma_n$, y $U = U(X_1, \dots, X_n)$ es una función de $X_1, \dots, X_n$, entonces el método de la propagación del error (véase el capítulo 3) se puede utilizar para calcular la desviación estándar, o la incertidumbre, en U . Si además las variables aleatorias $X_1, \dots, X_n$ están distribuidas aproximadamente normales, se da con frecuencia el caso (no siempre) de que también U se distribuye aproximadamente normal. En estos casos se puede utilizar la expresión (5.13) (de la sección 5.3) para calcular un intervalo de confianza para la media de μ_U de U . Con el propósito de determinar si U está distribuida aproximadamente normal, se puede utilizar la simulación.

Para dar un ejemplo concreto, suponga que se fabrican lavadoras de acero, cuyos radios están distribuidos normalmente con media μ_R desconocida y desviación estándar conocida $\sigma_R = 0.1$ cm. Se observa que una sola lavadora, seleccionada aleatoriamente, tiene un radio de $R = 2.5$ cm. Debido a que R proviene de una población normal con desviación estándar conocida, se puede utilizar la expresión (5.13) para encontrar un intervalo de confianza para la media del radio μ_R . Un intervalo de confianza de 95% para μ_R es $R \pm 1.96\sigma_R = 2.5 \pm 0.196$. Ahora se considera el área de la lavadora. Ésta es $A = \pi R^2 = 3.14(2.5^2) = 19.63$ cm². Se puede usar el método de la propagación del error (ecuación 3.10 de la sección 3.3) para calcular la desviación estándar, o la incertidumbre, σ_A , por

$$\sigma_A = \left| \frac{dA}{dR} \right| \sigma_R = 2\pi R \sigma_R = 2(3.14)(2.5)(0.10) = 1.57 \text{ cm}^2$$

Ahora se tiene un solo valor muestreado $A = 19.63$, poblacional de todas áreas posibles, y se tiene una estimación de la desviación estándar de esta población, $\sigma_A = 1.57$. ¿Se puede encontrar un intervalo de confianza para la media μ_A del área? Si la distribución de áreas estuviera distribuida normalmente, se podría encontrar un intervalo de confianza para μ_A con el mismo método usado para μ_R , obteniendo $A \pm 1.96\sigma_A = 19.63 \pm 1.96(1.57)$. La suposición de normalidad es necesaria para justificar este método. ¿Cómo se puede determinar si la distribución de las áreas es normal? Si se tuviera una muestra grande de áreas, se podría construir una gráfica de probabilidad normal. No se tiene una muestra grande de áreas, pero se puede simular una, como se muestra a continuación (véase la sección 4.11 para un análisis de los principios básicos de simulación).

Se empieza por generar una muestra grande de radios simulados R^* . Se quiere que la distribución poblacional de la cual se extrae la muestra esté tan cerca como sea posible a la distribución poblacional a partir de la cual se extrajo la observación $R = 2.5$. Se sabe que el valor $R = 2.5$ fue extraído de una población normal con $\sigma_R = 0.1$. No se conoce μ_R , pero se puede utilizar el valor observado $R = 2.5$ como una aproximación a μ_R para propósitos de la simulación. Por tanto, se generará una muestra grande de radios simulados $R_1^*, \dots, R_N^*$ de una distribución $N(2.5, 0.1^2)$. (La notación R_i^* indica que ello representa un valor simulado, en lugar de un valor observado en un experimento verdadero.)

Para entender exactamente qué se puede hacer con los valores simulados, imagine que se tenía una muestra grande de lavadoras verdaderas, y que sus radios $R_1, \dots, R_n$ se habían determinado. ¿Cuáles son las similitudes y las diferencias entre la muestra real $R_1, \dots, R_n$ y la muestra simulada $R_1^*, \dots, R_N^*$? La muestra verdadera proviene de una población que está

normalmente distribuida, se sabe que la desviación estándar es de 0.1, y su media μ_R es desconocida. La muestra simulada proviene de una población que también está distribuida normalmente, cuya desviación estándar es también igual a 0.1, y cuya media se ha establecido de 2.5 (el valor de una observación verdadera de R). Por tanto, la población simulada tiene la misma forma (normal) y la extensión (desviación estándar) poblacional real. Las poblaciones simuladas y reales tienen medias diferentes (2.5 para la población simulada y el valor desconocido μ_R para la población real).

Ahora para cada R_i^* se calcula un área simulada $A_i^* = \pi R_i^{*2}$. Imagine que se tenía una muestra verdadera de radios, $R_1, \dots, R_n$, y se calcula una muestra real $A_1, \dots, A_n$, de áreas de ésta. Debido a que la muestra de radios simulados $R_1^*, \dots, R_n^*$ proviene de una población cuya forma y extensión son las mismas que la de la población real de radios, es razonable suponer que la muestra simulada de áreas $A_1^*, \dots, A_n^*$ proviene de una población cuya forma y extensión son muy similares a la población verdadera de áreas. En otras palabras, es razonable suponer que la desviación estándar muestral simulada $A_1^*, \dots, A_n^*$ es cercana a la desviación estándar muestral σ_A , y es razonable suponer que si la muestra simulada $A_1^*, \dots, A_n^*$ proviene de una población que es aproximadamente normal, la población verdadera de áreas es también aproximadamente normal. Sin embargo, la media poblacional simulada de áreas diferirá de la media poblacional verdadera de las áreas. La razón de esto es que las medias de las poblaciones simuladas y reales de radios son diferentes. La media poblacional simulada de áreas será cercana a un valor observado verdadero de A , que es 19.63. La media poblacional verdadera de áreas es el valor desconocido μ_A .

Se construye una gráfica de probabilidad normal para $A_1^*, \dots, A_n^*$. Si se muestra que la población de áreas es aproximadamente normal, entonces se puede suponer que el verdadero observado proviene de una población aproximadamente normal, y se puede encontrar un intervalo de confianza para μ_A . La figura 5.16 presenta una gráfica de probabilidad normal para una muestra de 1 000 áreas. Con excepción de algunos puntos en cualesquiera de los extremos, la suposición de normalidad parece bien satisfecha.

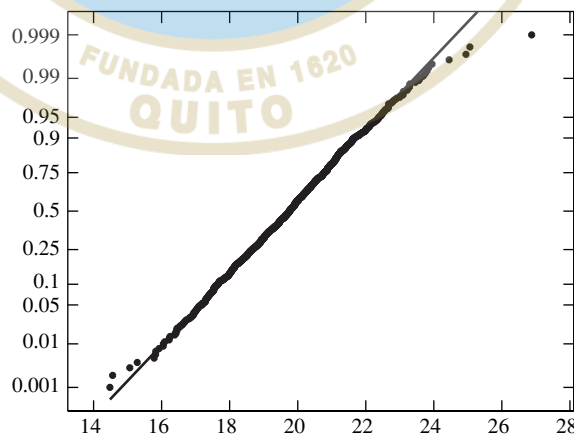


FIGURA 5.16 Gráfica de probabilidad normal para 1 000 áreas simuladas. La suposición de normalidad está justificada.

El valor observado de A en realidad es $A = 19.63$. La desviación estándar de la población de la que A fue extraída, calculada con el método de propagación del error, es $\sigma_A = 1.57$. Debido a que esta población es aproximadamente normal, un intervalo de confianza de 95% para la media μ_A del área es $19.63 \pm 1.96(1.57)$, o $(16.55, 22.71)$.

En el intervalo de confianza apenas calculado, la desviación estándar σ_A se obtuvo con el método de propagación del error. Es también aceptable usar la desviación estándar muestral de valores simulados $A_1^*, \dots, A_n^*$ para estimar σ_A . En los mil valores que se simularon, la desviación estándar muestral fue de 1.5891, que estaba cerca del valor 1.57 calculado por la propagación del error. Este acercamiento es típico.

Es importante observar que el centro del intervalo de confianza es el verdadero valor observado A , y no la media del valor simulado $\bar{A}^*$. La razón para esto último es que se está determinando un intervalo de confianza para la media poblacional verdadera de áreas μ_A , y el valor observado A se ha muestreado de esta población. Los valores simulados se han muestreado de una población cuya media es diferente a la población verdadera. Por tanto, $\bar{A}^*$ no es una elección adecuada para el centro del intervalo de confianza.

El método apenas descrito puede ser muy útil cuando se hacen mediciones cuyos errores de medición están distribuidos normalmente. Se presentan algunos ejemplos.

Ejemplo

5.27

La longitud y el ancho de un rectángulo se miden de $X = 3.0 \pm 0.1$ y $Y = 3.5 \pm 0.2$ cm, respectivamente. Suponga que las mediciones provienen de poblaciones normales y no tienen sesgos. Suponga que se conocen las desviaciones estándares $\sigma_X = 0.1$ y $\sigma_Y = 0.2$. Determine un intervalo de confianza de 95% para el área del rectángulo.

Solución

Sea $A = XY$ la medida del área de un rectángulo. El valor observado de A es $A = (3.0)(3.5) = 10.5$. Se usará el método de propagación del error para calcular σ_A . Entonces se utilizará la simulación para comprobar que la distribución de A es aproximadamente normal. Para calcular σ_A , primero se calculan las derivadas parciales de A con respecto a X y Y .

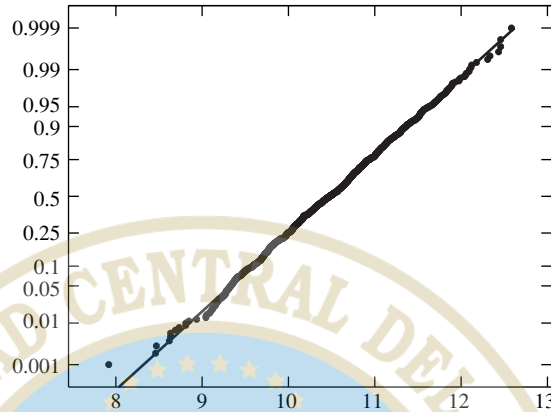
$$\frac{\partial A}{\partial X} = Y = 3.5 \quad \frac{\partial A}{\partial Y} = X = 3.0$$

La desviación estándar σ_A se estima con

$$\begin{aligned} \sigma_A &= \sqrt{\left(\frac{\partial A}{\partial X}\right)^2 \sigma_X^2 + \left(\frac{\partial A}{\partial Y}\right)^2 \sigma_Y^2} \\ &= \sqrt{(3.5)^2(0.1)^2 + (3.0)^2(0.2)^2} \\ &= 0.6946 \end{aligned}$$

Para comprobar la normalidad, se generan mil valores simulados $X_1^*, \dots, X_{1000}^*$ de una distribución $N(3.0, 0.1^2)$, y otros mil $Y_1^*, \dots, Y_{1000}^*$ de una distribución $N(3.5, 0.1^2)$. Observe que se utilizaron los valores observados 3.0 y 3.5 para hacer una aproximación de las medias desconocidas μ_X y μ_Y (que es la longitud y el ancho, verdadero, respectivamente) para propósitos

de simulación. Después se calcula 1 000 áreas simuladas $A_1^*, \dots, A_{1000}^*$. Una gráfica de probabilidad normal se presenta en la figura siguiente. La suposición de la normalidad está justificada. Un intervalo de confianza de 95% para el área del rectángulo es $10.5 \pm 1.96(0.6946)$.



Nota técnica: En el ejemplo 5.27 se está encontrando un intervalo de confianza para la media μ_A del área medida. El área verdadera del rectángulo es el producto de la longitud verdadera por el ancho verdadero, que es $\mu_X \mu_Y$. El valor μ_A es un poco diferente del producto $\mu_X \mu_Y$, pero la diferencia es despreciable para propósitos prácticos.

Ejemplo 5.28

Dos resistores, cuyas resistencias se miden de X y Y , están conectados en paralelo. Se calcula la resistencia total con $R = (XY)/(X + Y)$. Suponga que $X = 10.0 \pm 1.0 \Omega$, $Y = 20.0 \pm 2.0 \Omega$, y que X y Y provienen de poblaciones normales y no tienen sesgos. Encuentre un intervalo de confianza de 95% para la resistencia total.

Solución

El valor observado de R es $(10)(20)/(10 + 20) = 6.667 \Omega$. Para calcular σ_R , primero se calculan las derivadas parciales de R :

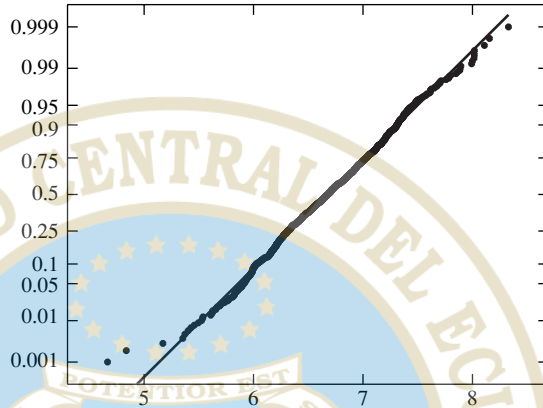
$$\frac{\partial R}{\partial X} = \left(\frac{Y}{X + Y} \right)^2 = 0.4444$$

$$\frac{\partial R}{\partial Y} = \left(\frac{X}{X + Y} \right)^2 = 0.1111$$

Ahora $\sigma_X = 1 \Omega$ y $\sigma_Y = 2 \Omega$. Por tanto,

$$\begin{aligned} \sigma_R &= \sqrt{\left(\frac{\partial R}{\partial X} \right)^2 \sigma_X^2 + \left(\frac{\partial R}{\partial Y} \right)^2 \sigma_Y^2} \\ &= \sqrt{(0.4444)^2 (1)^2 + (0.1111)^2 (2)^2} \\ &= 0.497 \Omega \end{aligned}$$

Para comprobar la normalidad, se han generado mil valores simulados $X_1^*, \dots, X_{1000}^*$ de una distribución $N(10, 1.0^2)$, y mil $Y_1^*, \dots, Y_{1000}^*$ de $N(20, 2.0^2)$. Observe que se utilizan los valores observados 10 y 20 para aproximar las medias μ_X y μ_Y para propósitos de simulación. Después se calculan mil valores simulados $R_1^*, \dots, R_{1000}^*$. En la siguiente figura se muestra una gráfica de probabilidad normal. La suposición de normalidad está justificada. Un intervalo de confianza de 95% para la resistencia total es $6.667 \pm 1.96(0.497)$.



La nota técnica que se muestra después del ejemplo 5.27 también se aplica al ejemplo 5.28. Se está determinando realmente un intervalo de confianza en realidad para la media μ_R de la resistencia total medida. Esta media μ_R es un poco diferente de la resistencia total verdadera $\mu_X\mu_Y/(\mu_X + \mu_Y)$, pero la diferencia es despreciable para propósitos prácticos.

En algunos casos la distribución de una función $U(X_1, \dots, X_n)$ no es normal, aun cuando $X_1, \dots, X_n$ son normales. Por esta razón es importante comprobar la normalidad con una simulación. El ejemplo 5.29 proporciona una muestra.

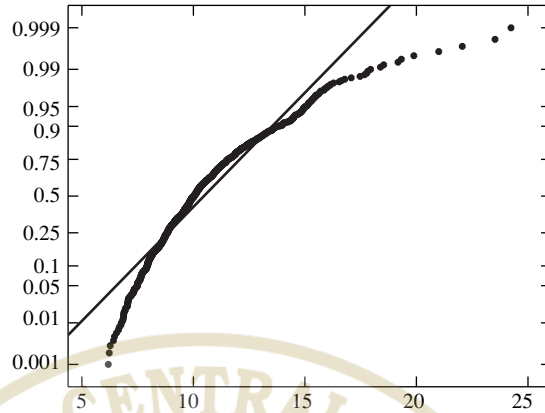
Ejemplo

5.29

La masa de una roca se mide de $M = 10 \pm 0.4$ g, y su volumen se mide de $V = 1.0 \pm 0.2$ ml. Se calcula la densidad con $D = M/V$. Se supone que M y V provienen de poblaciones normales y no tienen sesgos. ¿ D está distribuida normalmente? ¿Se puede utilizar el método descrito en los ejemplos 5.27 y 5.28, que está basado en la curva normal, para encontrar un intervalo de confianza de 95% para la densidad de la roca?

Solución

Se generan mil valores simulados $M_1^*, \dots, M_{1000}^*$ de una distribución de $N(10, 0.4^2)$, y mil más $V_1^*, \dots, V_{1000}^*$ de una distribución $N(1.0, 0.2^2)$. Después se calculan los valores $D_i^* = M_i^*/V_i^*$. Una gráfica de probabilidad normal de los D_i^* se muestra en la figura siguiente. La suposición de normalidad no está justificada. El método basado en la curva normal no se puede utilizar para encontrar un intervalo de confianza para la densidad de la roca.



Intervalos de confianza usando estimación bootstrap

Cuando se extrae una muestra de una población que no es normal, y se requiere un intervalo de confianza, se pueden utilizar los métodos basados en la estimación bootstrap (véase la sección 4.11). Existen muchos de estos métodos; aquí se presenta uno simple y se muestra cómo utilizarlo para construir un intervalo de confianza para una media poblacional.

Se considerará el siguiente ejemplo. El artículo “In-use Emissions from Heavy Duty Diesel Vehicles” (J. Yanowitz, tesis de doctorado, Escuela de Minas, de Colorado, 2001) presenta mediciones de eficiencia del combustible (en milla/galón) de una muestra de siete camiones. Los datos son los siguientes:

7.69	4.97	4.56	6.49	4.34	6.24	4.45

Se supone que ésta es una muestra aleatoria de una población de camiones, y que se desea construir un intervalo de confianza de 95% para la media μ de la eficiencia de combustible de esta población. Al observar la muestra se sugiere que hay una separación cerca de la mitad de la distribución, ya que no hay ningún camión en la muestra con valores entre 5 y 6. Por tanto, no se podría suponer que los datos fueran normales. La estimación bootstrap proporciona un método para construir un intervalo de confianza, cuyo nivel será aproximadamente de 95% (o cualquier otro valor que se podría especificar).

Para construir un intervalo de confianza de estimación bootstrap, se deben extraer *muestras de estimación bootstrap* de los datos. Una muestra de estimación bootstrap es del mismo tamaño que los datos, extraída con reemplazo. Para describir esto último con detalle, se denotan los valores en una muestra aleatoria por $X_1, \dots, X_n$. Imagine colocar estos valores en una caja y extraer uno al azar. Éste sería el primer valor de la muestra de estimación bootstrap; se le denota con X_1^* . Después se regresa a X_1^* a la caja, y se extrae otro valor, X_2^* . Se continúa de este modo hasta que se ha extraído n valores $X_1^*, \dots, X_n^*$. Ésta es una muestra de estimación bootstrap. Observe que cada valor en la muestra de estimación bootstrap se ha extraído de la muestra de datos completa, por lo que es probable que algunos valores aparezcan más de una vez mientras que otros no lo harán.

Se extrajeron mil muestras por estimación bootstrap de los datos específicos del millaje. Los primeros diez y el último de éstos se presentan en la tabla 5.2. La media muestral se calcula para cada muestra de estimación bootstrap.

TABLA 5.2 Muestras de estimación bootstrap de los datos del millaje

Muestra	Valores de la muestra								Media muestral
1	4.97	6.49	7.69	4.97	7.69	4.56	4.45		5.8314
2	6.24	4.97	4.56	4.97	4.56	6.24	7.69		5.6043
3	4.34	4.45	4.56	4.45	6.24	4.97	4.34		4.7643
4	4.45	6.49	7.69	6.24	4.97	4.45	4.34		5.5186
5	6.24	4.34	4.45	7.69	4.56	4.34	4.45		5.1529
6	4.34	4.97	7.69	4.97	6.24	6.24	6.24		5.8129
7	4.45	6.49	6.24	4.97	4.34	7.69	4.34		5.5029
8	6.49	7.69	4.97	6.49	6.49	4.34	4.56		5.8614
9	7.69	4.45	4.45	4.45	4.45	4.56	4.56		4.9443
10	6.24	4.56	4.97	6.49	4.45	4.97	6.24		5.4171
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮
1 000	4.34	7.69	4.45	4.56	7.69	4.45	7.69		5.8386

Para construir un intervalo de confianza de estimación bootstrap, se deben extraer muchas muestras de estimación bootstrap (mil como mínimo). Debido a que se quiere un intervalo de confianza para la media μ poblacional, se calcula la media muestral para cada muestra de estimación bootstrap. Sea $\bar{X}_i^*$ la media de la i -ésima muestra de estimación bootstrap. Debido a que se quiere que el nivel del intervalo de confianza esté tan cerca de 95% como sea posible, se determina el intervalo que contiene a 95% del intermedio de las medias muestrales de estimación bootstrap. Los puntos finales de este intervalo son 2.5 percentil y 97.5 percentil de la lista de las medias de estimación bootstrap. Se denotan estos percentiles por $\bar{X}_{.025}^*$ y $\bar{X}_{.975}^*$, respectivamente.

Se calcularon estos percentiles para los datos del millaje. Las siguientes son las 26 más pequeñas y 26 más grandes de las mil muestras de estimación bootstrap $\bar{X}_i^*$.

26 más pequeñas:	4.4929	4.4971	4.5357	4.5400	4.5514	4.5557	4.5557	4.5829
	4.5986	4.6143	4.6429	4.6457	4.6729	4.6729	4.6900	4.6943
	4.7014	4.7157	4.7257	4.7257	4.7329	4.7371	4.7414	4.7486
	4.7643	4.7643						
26 más grandes:	6.4757	6.4757	6.4800	6.4900	6.4986	6.5214	6.5443	6.5543
	6.5929	6.5929	6.6257	6.6257	6.6471	6.6671	6.6900	6.6929
	6.7057	6.7129	6.7514	6.7971	6.7971	6.8486	6.9329	6.9686
	7.0714	7.1043						

Utilizando el método del cálculo de percentiles presentado en el capítulo, el percentil 2.5 es el promedio de los 25o. y 26o. valores en la muestra ordenada de mil, y el percentil 97.5 es el promedio de los valores 975o. y 976o. Por tanto, en este caso $\bar{X}_{.025}^* = 4.7643$ y $\bar{X}_{.975}^* = 6.4757$.

Ahora hay dos métodos disponibles para construir el intervalo de confianza; cuál será el mejor es un tema de cierta controversia. En el método 1, el intervalo de confianza es $(\bar{X}_{.025}^*, \bar{X}_{.975}^*)$. El método 2 utiliza la media $\bar{X}$ de la muestra original, además de los percentiles; con el método 2 el intervalo de confianza es $(2\bar{X} - \bar{X}_{.975}^*, 2\bar{X} - \bar{X}_{.025}^*)$. Para los datos del millaje, el intervalo de confianza de 95% por el método 1 es (4.7643, 6.4757).

La media muestral para los datos del millaje es $\bar{X} = 5.5343$. Por tanto, el intervalo de confianza de 95% calculado por el método 2 es

$$(2(5.5343) - 6.4757, 2(5.5343) - 4.7643) = (4.5929, 6.3043)$$

Para este caso específico, los intervalos de confianza de los dos métodos son similares.

Resumen

Dada una muestra aleatoria $X_1, \dots, X_n$ de una población con media μ , un intervalo de confianza de estimación bootstrap para μ con nivel aproximado de $100(1 - \alpha)\%$ se puede calcular como se muestra a continuación:

- Extraiga un gran número m ($m \geq 1\,000$) de muestras de estimación bootstrap de tamaño n con reemplazo de $X_1, \dots, X_n$.
- Calcule la media de cada muestra de estimación bootstrap. Denote estas medias de estimación bootstrap por $\bar{X}_1^*, \dots, \bar{X}_m^*$.
- Calcule los percentiles $100\alpha/2$ y $100(1 - \alpha/2)$ de las medias de estimación. Denote estos valores $\bar{X}_{\alpha/2}^*$, $\bar{X}_{1 - \alpha/2}^*$.
- Hay dos métodos para calcular el intervalo de confianza.

Método 1: $(\bar{X}_{\alpha/2}^*, \bar{X}_{1 - \alpha/2}^*)$ Método 2: $(2\bar{X} - \bar{X}_{1 - \alpha/2}^*, 2\bar{X} - \bar{X}_{\alpha/2}^*)$

Aunque al principio no es obvio, existe una conexión entre el método de estimación bootstrap presentado aquí para el cálculo de los intervalos de confianza para una media poblacional y el de grandes muestras basado en la curva normal. En ambos casos el ancho del intervalo de confianza debe idealmente ser igual al ancho de 95% intermedio de la distribución de la media muestral $\bar{X}$. Cuando el tamaño muestral es grande, la distribución de $\bar{X}$ sigue la curva normal, por lo que el ancho del intervalo de confianza de 95% es puesto para igualar el ancho de 95% intermedio de la distribución normal (véase la figura 5.1 en la sección 5.1). Se utiliza la estimación bootstrap cuando la distribución de $\bar{X}$ no es necesariamente normal. La colección de las medias muestrales de estimación bootstrap $\bar{X}_i^*$ se aproxima a una muestra aleatoria de la distribución de $\bar{X}$, por lo que esta colección, en vez de la curva normal, constituye la base para el intervalo de confianza. El ancho del intervalo de confianza de estimación bootstrap es puesto para igualar el ancho de 95% intermedio de la media muestral de estimación bootstrap con el fin de aproximar el ancho de la distribución desconocida de $\bar{X}$.

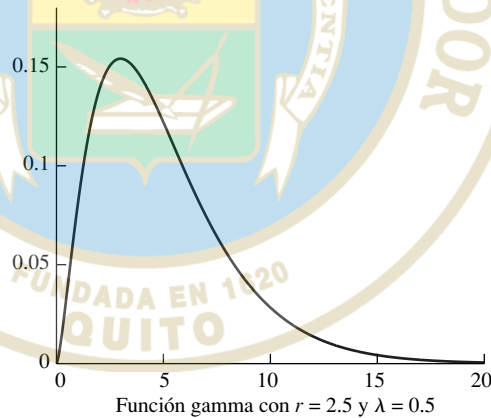
Hay muchos métodos diferentes para calcular intervalos de confianza de estimación bootstrap. Los métodos simples que aquí se presentan funcionan bien cuando la población de la cual se extrajo la muestra de estimación bootstrap es aproximadamente simétrica, pero no funcionan bien cuando está muy sesgada. Se han desarrollado métodos más sofisticados que producen buenos resultados con condiciones más generales. Se puede encontrar información adicional sobre este tema en Efron y Tibshirani (1993).

Uso de simulación para evaluar intervalos de confianza

Un intervalo con nivel de confianza del $100(1 - \alpha)\%$ es el que se calcula mediante un procedimiento que tiene éxito en contener el valor verdadero para el $100(1 - \alpha)\%$ de todas las

muestras que posiblemente se pudiera extraer. Cuando los supuestos que gobiernan el uso de un método son incumplidos, esta tasa de éxito (también llamada la **probabilidad de cobertura**) podría ser más baja. Con frecuencia, en la práctica, los supuestos no son válidos. Algunos métodos son muy sensibles a sus suposiciones, en lo que respecta a la probabilidad de cobertura puede ser mucho menos del $100(1 - \alpha)\%$, aun cuando los supuestos son solamente incumplidos. Otros métodos son “robustos”, ello significa que la probabilidad de cobertura no se va muy por debajo del $100(1 - \alpha)\%$ en tanto sus suposiciones se satisfagan de manera aproximada. La ventaja de un método robusto es que es útil en una amplia gama de condiciones y requiere menos preocupación acerca de sus suposiciones. Los experimentos de simulación proporcionan una buena manera de valorar la robustez de un procedimiento estadístico. En seguida se presenta un experimento que será instructivo respecto a la robustez del método t de Student para construir intervalos de confianza para una media poblacional (expresión 5.9 en la sección 5.3).

La distribución t de Student se puede utilizar para construir intervalos de confianza para una media poblacional, siempre que la muestra provenga de una población que sea “aproximadamente” normal. Se realizará un experimento de simulación para tener un poco de percepción de cómo se puede hacer burdamente esta aproximación. La siguiente figura muestra la función de densidad de probabilidad para la distribución $\Gamma(2.5, 0.5)$ (que también se conoce como distribución Ji-cuadrada con cinco grados de libertad). Está bastante sesgada y no se parece a la curva normal. La media de esta población es $\mu = 5$. Si el método t de Student se aplica a muestras de tamaño 5 de esta población, ¿qué proporción de veces el intervalo de confianza de 95% contendrá la media verdadera?



Para abordar esta pregunta, se generan 10 000 muestras de tamaño 5 de la distribución $\Gamma(2.5, 0.5)$. Se denota la i -ésima muestra mediante $X_{1i}^*, X_{2i}^*, X_{3i}^*, X_{4i}^*, X_{5i}^*$; denote su media muestral por medio de $\bar{X}_i^*$; y la desviación estándar muestral por s_i^* . Para cada muestra se calcula un intervalo de confianza usando la fórmula por un intervalo de confianza de 95% basado en la distribución t de Student (expresión 5.9 de la sección 5.3). El límite de confianza inferior es $L_i^* = \bar{X}_i^* - 2.776s_i^*/\sqrt{5}$ (observe que $t_{4, 0.025} = 2.776$). El límite de confianza superior es $U_i^* = \bar{X}_i^* + 2.776s_i^*/\sqrt{5}$. En la tabla 5.3 se presenta los resultados para las primeras diez y la última muestras. La columna derecha contiene un “1” si $L_i^* < 5 < U_i^*$, en otras palabras, si el i -ésimo intervalo de confianza contiene a la media verdadera de 5.

TABLA 5.3 Datos simulados de la distribución $\Gamma(2.5, 0.5)$

i	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*	$\bar{X}^*$	s^*	L^*	U^*	$L^* < 5 < U^*$
1	2.58	6.54	3.02	3.40	1.23	3.36	1.96	0.92	5.79	1
2	1.28	1.44	1.45	10.22	4.17	3.71	3.83	-1.05	8.47	1
3	7.26	3.28	2.85	8.94	12.09	6.88	3.89	2.05	11.72	1
4	6.11	3.81	7.06	11.89	3.01	6.38	3.49	2.04	10.72	1
5	4.46	9.70	5.14	2.45	4.99	5.35	2.66	2.05	8.65	1
6	2.20	1.46	9.30	2.00	4.80	3.95	3.26	-0.09	7.99	1
7	7.17	13.33	6.19	10.31	8.49	9.10	2.83	5.59	12.61	1
8	1.97	1.81	4.13	1.28	5.16	2.87	1.68	0.78	4.95	0
9	3.65	1.98	8.19	7.20	3.81	4.97	2.61	1.72	8.21	1
10	3.39	2.31	1.86	5.97	5.28	3.76	1.80	1.52	6.00	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
10 000	7.30	7.21	1.64	3.54	3.41	4.62	2.52	1.49	7.75	1

Nueve de los diez primeros intervalos de confianza contienen la media verdadera. Así que si se basaran en los resultados de las diez primeras muestras, se estimaría la probabilidad de cobertura del intervalo de confianza de 0.90. Por supuesto, diez muestras no son suficientes. De las 10 000 muestras, la media verdadera estaba contenida en 9 205 de ellas. Por tanto, se estima la probabilidad de cobertura de 0.9205. Mientras que ésta es menor de 95%, no es dramáticamente menor. Esta simulación indica que el procedimiento t de Student es bastante fuerte; en otras palabras, que los intervalos de confianza basados en la distribución t de Student contienen la media verdadera tan a menudo como la deben contener, incluso cuando la población es un poco diferente de la normal.

Si la población se desvía mucho de la normal, el método t de Student no funciona bien. (Véase el ejercicio 8.)

Ejercicios para la sección 5.8

1. La presión de aire (en MPa) que entra en un compresor que se mide de $X = 8.5 \pm 0.2$, y la que sale se mide de $Y = 21.2 \pm 0.3$. Por tanto, la presión intermedia se mide de $P = \sqrt{XY} = 13.42$. Suponga que X y Y provienen de poblaciones normales y no tienen sesgos.
 - a) ¿A partir de qué distribuciones es apropiado simular los valores X^* y Y^* ?
 - b) Genere muestras simuladas de los valores X^* , Y^* y P^* .
 - c) Utilice el método de propagación del error para calcular la desviación estándar de P . (Puede usar la muestra simulada.)
 - d) Construya una gráfica de probabilidad normal para los valores P^* . ¿Es razonable suponer que P está aproximadamente normal distribuida?
 - e) Si es apropiado, use la curva normal para encontrar un intervalo de confianza de 95% para la presión intermedia.
2. La masa (en kg) de una muestra de tierra se mide de $X = 1.18 \pm 0.02$. Después la muestra es deshidratada en un horno, la masa de la tierra deshidratada es de $Y = 0.85 \pm 0.02$. Suponga que X y Y provienen de poblaciones normales y no tienen sesgos. El contenido de agua de la tierra se mide de

$$W = \frac{X - Y}{X}$$

- a) ¿Con cuál distribución es adecuado simular los valores de X^* y Y^* ?
 - b) Genere muestras simuladas de valores X^* , Y^* y W^* .
 - c) Use el método de propagación del error para calcular la desviación estándar de W (puede usar la muestra simulada).
 - d) Construya una gráfica de probabilidad normal para los valores de W^* . ¿Es razonable suponer que W está distribuida normalmente?
 - e) Si es apropiado, utilice la curva normal para encontrar un intervalo de confianza de 95% para el contenido de agua.
3. Un estudiante mide la aceleración A de un carro que se está moviendo hacia abajo de un plano inclinado midiendo el tiempo T cuánto le toma al carro recorrer 1 m y usando la fórmula $A = 2/T^2$. Suponga que $T = 0.55 \pm 0.01$ s, y que la medición proviene de una población de normal y no tiene sesgos.
- a) Genere una muestra simulada adecuada de valores de A^* . Es razonable suponer que A se distribuye normalmente.
 - b) Utilice el método de propagación del error para calcular la desviación estándar de A . (Por otra parte, puede usar la muestra simulada.)
 - c) Si es adecuado, utilice la curva normal para encontrar un intervalo de confianza de 95% para la aceleración del carro.
4. La temperatura inicial de cierto recipiente se mide de $T_0 = 20^\circ\text{C}$. La temperatura ambiental se mide de $T_a = 4^\circ\text{C}$. Un ingeniero utiliza la ley de enfriamiento de Newton para calcular el tiempo necesario para enfriar el recipiente a una temperatura de 10°C . Tomando en cuenta las propiedades físicas del recipiente, este tiempo (en minutos) se calcula de
- $$T = 40 \ln \left(\frac{T_0 - T_a}{10 - T_a} \right)$$
- Suponga que las mediciones de temperatura T_0 y T_a no tienen sesgos y que provienen de poblaciones normales con desviación estándar de 0.1°C .
- a) Genere una muestra simulada adecuada de valores T^* . ¿Es razonable suponer que T está distribuida normalmente?
 - b) Utilice el método de propagación del error para calcular la desviación estándar de T (puede usar la muestra simulada).
 - c) Si es adecuado, utilice la curva normal para encontrar un intervalo de confianza de 95% para el tiempo necesario para enfriar el recipiente a una temperatura de 10°C .
5. En el artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, en *Water Environment Research*, 2002:177-186), se midieron concentraciones de amonio (en mg/l) en un gran número de pozos en Iowa. Éstas incluían 349 pozos aluviales y 143 pozos cuaternarios. Las concentraciones en los pozos de aluviales promediaban 0.27 con una desviación estándar de 0.40, y aquellos en los pozos cuaternarios determinaron el promedio de 1.62 con una desviación estándar de 1.70. Debido a que estas desviaciones estándares están basadas en muestras grandes, suponga que son despreciables a diferencia de las desviaciones estándares poblacionales. Una estimación para la razón de la media de la concentración en pozos cuaternarios con la media de la concentración en pozos aluviales es $R = 1.62/0.27 = 6.00$.
- a) Puesto que las medias muestrales 1.62 y 0.27 están basadas en muestras grandes, es razonable suponer que provienen de poblaciones normales. ¿Qué distribución se aproxima a la distribución de la media muestral de concentración de pozos aluviales, $N(0.27, 0.40^2)$ o $N(0.27, 0.40^2/349)$? ¿Qué distribución se aproxima a la distribución de la media muestral de concentración de pozos cuaternarios, $N(1.62, 1.70^2)$ o $N(1.62, 1.70^2/143)$? Explique.
 - b) Genere una muestra simulada de las medias muestrales y de las proporciones de las medias muestrales. ¿Es razonable suponer que la proporción R está distribuida normalmente?
 - c) Utilice el método de la propagación del error para estimar la desviación estándar de R (puede usar la muestra simulada).
 - d) Si es adecuado, utilice la curva normal para encontrar un intervalo de confianza de 95% para la proporción de las medias de las concentraciones.
6. En el ejemplo 5.20 (de la sección 5.3) se presentaron las mediciones siguientes para la fuerza compresiva cilíndrica (en MPa) para 11 vigas:

38.43	38.43	38.39	38.83	38.45	38.35
38.43	38.31	38.32	38.48	38.50	

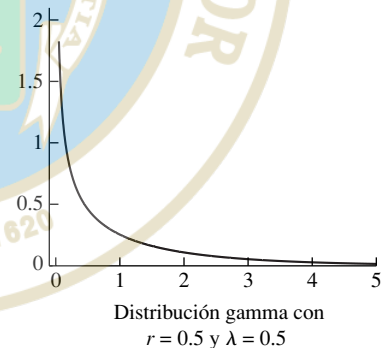
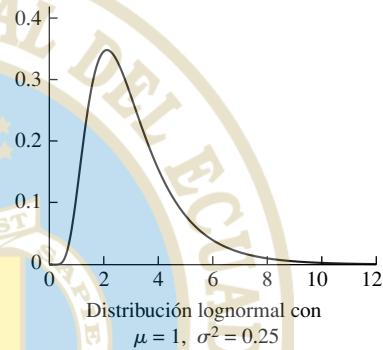
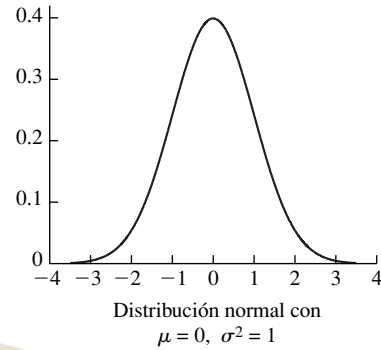
Se generaron 1 000 muestras de estimación bootstrap a partir de estos datos, y las medias muestrales de estimación bootstrap

se arreglaron en orden. Denotando al valor más pequeño con Y_1 , el segundo más pequeño como Y_2 , y así sucesivamente, siendo el más grande $Y_{1\ 000}$. Suponga que $Y_{25} = 38.3818$, $Y_{26} = 38.3818$, $Y_{50} = 38.3909$, $Y_{51} = 38.3918$, $Y_{950} = 38.5218$, $Y_{951} = 38.5236$, $Y_{975} = 38.5382$ y $Y_{976} = 38.5391$.

- Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 1 como se describe en la p. 358.
 - Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.
 - Calcule un intervalo de confianza de estimación bootstrap de 90% para la media de la fuerza compresiva, usando el método 1, como se describe en la p. 358.
 - Calcule un intervalo de confianza de estimación bootstrap de 90% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.
7. Con referencia al ejercicio 6.

- Genere 1 000 muestras de estimación bootstrap de estos datos. Determine los percentiles 2.5 y 97.5.
- Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 1, como se describe en la p. 358.
- Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.

8. Este ejercicio continúa con el estudio de la fortaleza del método t de Student para construir intervalos de confianza. La siguiente figura muestra gráficas de las funciones de densidad de probabilidad para la distribución de $N(0, 1)$, la distribución lognormal con $\mu = 1$ y $\sigma^2 = 0.25$, y la distribución gamma con $r = 0.5$ y $\lambda = 0.5$ (esta última también se conoce como distribución Ji-cuadrada con un grado de libertad). Para cada una de estas distribuciones, genere 10 000 muestras de tamaño 5, y para cada muestra calcule los límites superior e inferior de un intervalo de confianza de 95% con el método t de Student. [Si fuera necesario, es posible calcular valores aleatorios lognormal y gamma a partir de valores aleatorios de la normal. Específicamente, calcule un valor X de una distribución de lognormal con $\mu = 1$ y $\sigma^2 = 0.25$, genere $Y \sim N(1, 0.25)$ y calcule $X = e^Y$. Para generar un valor de X a partir de una distribución gamma con $r = 0.5$ y $\lambda = 0.5$, genere $Y \sim N(0, 1)$ y calcule $X = Y^2$].



- La media verdadera de la distribución $N(0, 1)$ es 0. Con base en los resultados de simulación, estime la probabilidad de cobertura (proporción muestral para las que el intervalo de confianza contiene la media verdadera) para muestras de tamaño 5 de la distribución $N(0, 1)$. (En virtud de que los supuestos subyacentes del método t del estudiante se satisfacen aquí, su respuesta debe ser muy cercana a 95%.)

- b) La media verdadera de la distribución lognormal con $\mu = 1$, $\sigma^2 = 0.25$ es 3.0802. Con base en los resultados de simulación, calcule la probabilidad de cobertura (la proporción de muestras para las que el intervalo de confianza contiene a la media verdadera) para muestras de tamaño 5 de la distribución lognormal con $\mu = 1$, $\sigma^2 = 0.25$.
- c) La media verdadera de la distribución gamma con $r = 0.5$ y $\lambda = 0.5$ es 1. Con base en los resultados de simulación, calcule la probabilidad de cobertura (la proporción de muestras para las cuales el intervalo de confianza contiene a la media verdadera) para muestras de tamaño 5 de la distribución gamma con $r = 0.5$ y $\lambda = 0.5$.
9. Este ejercicio está diseñado para comparar el desempeño del intervalo de confianza de Agresti-Coull para una proporción (expresión 5.5 en la p. 316) con la del intervalo de confianza tradicional (expresión 5.8 en la p. 319). Se usarán tamaños muestrales de $n = 10$, $n = 17$, y $n = 40$, con $p = 0.5$.
- a) Genere 10 000 observaciones X_i^* , cada una de una distribución binomial con $n = 10$ y $p = 0.5$. Para cada observación, calcule los límites superior e inferior tanto para el intervalo de confianza de Agresti-Coull de 95% como para el tradicional. Para cada intervalo de confianza, calcule su longitud (límite superior-límite inferior). Utilice los datos simulados para calcular la probabilidad de cobertura y la media del ancho tanto para el intervalo de confianza de Agresti-Coull como para el tradicional.
- b) Repita el inciso a), usando $n = 17$.
- c) Repita el inciso a), usando $n = 40$.
- d) El desempeño del intervalo de confianza tradicional no mejora regularmente conforme el tamaño muestral aumenta; en lugar de eso oscila, por lo que la probabilidad de cobertura puede ser mejor para una muestra más pequeña que para una más grande. Compare las probabilidades de cobertura con el método tradicional para tamaños muestrales de 17 y de 40. ¿Sus resultados confirman este hecho?
- e) ¿Para qué tamaños muestrales el intervalo Agresti-Coull tiene mayor probabilidad de cobertura que con el tradicional? ¿Para qué tamaños muestrales son las probabilidades de cobertura casi iguales?
- f) Otras cosas serán iguales, un intervalo de confianza más angosto es mejor que uno más amplio. ¿Qué método produce intervalos de confianza con media del ancho más angosta?
10. Un método general para encontrar un intervalo de confianza para la diferencia entre dos medias poblacionales normales está dado por la expresión (5.21) en la p. 341. Se puede usar un método combinado cuando se sabe que las varianzas de las poblaciones son iguales y está dado por la expresión (5.22) en la p. 343. Este ejercicio está diseñado para comparar las probabilidades de cobertura de estos métodos bajo una variedad de condiciones.
- a) Sea $n_X = 10$, $n_Y = 10$, $\sigma_X = 1$ y $\sigma_Y = 1$. Genere 10 000 pares de muestras: $X_{1i}^*, \dots, X_{1n_X}^*$ de una distribución $N(0, \sigma_X^2)$, y $Y_{1i}^*, \dots, Y_{1n_Y}^*$ de distribución $N(0, \sigma_Y^2)$. Para cada par de muestras, calcule un intervalo de confianza de 95% con el método general, y otro de 95% utilizando el método combinado. Observe que cada población tiene media 0, por lo que la diferencia verdadera entre medias es 0. Estime la probabilidad de cobertura para cada método calculando la proporción de intervalos de confianza que contienen al valor verdadero 0.
- b) Repita el inciso a), utilizando $n_X = 10$, $n_Y = 10$, $\sigma_X = 1$ y $\sigma_Y = 5$.
- c) Repita el inciso a), utilizando $n_X = 5$, $n_Y = 20$, $\sigma_X = 1$ y $\sigma_Y = 5$.
- d) Repita el inciso a), utilizando $n_X = 20$, $n_Y = 20$, $\sigma_X = 1$ y $\sigma_Y = 5$.
- e) La probabilidad de cobertura para el método general es diferente de 95% en cualesquiera de las condiciones de los incisos a) al d)? (Esto no se debe hacer.)
- f) Con base en los resultados de los incisos a) al d), ¿bajo qué condiciones el desempeño del método combinado sería más malo?
- i) Cuando los tamaños muestrales son iguales y las varianzas son diferentes.
- ii) Cuando tanto los tamaños muestrales como las varianzas son diferentes, y la muestra más grande proviene de la población con la varianza más grande.
- iii) Cuando tanto los tamaños muestrales como las varianzas son diferentes, y la muestra más pequeña proviene de la población con la varianza más grande.

Ejercicios adicionales para el capítulo 5

- Un biólogo molecular está estudiando la eficacia de una enzima especial para digerir cierta secuencia de nucleótidos del ADN. Divide seis muestras de ADN en dos partes, trata una parte con la enzima, y deja la otra parte sin tratar. Después analiza utilizando una reacción en cadena de polimerasa para contar el número de fragmentos de ADN que contienen la secuencia dada. Los resultados son los siguientes:

	Muestra					
	1	2	3	4	5	6
Enzima presente	22	16	11	14	12	30
Enzima ausente	43	34	16	27	10	40

Determine un intervalo de confianza del 95% para la diferencia entre la media de los números de fragmentos.

- Con referencia al ejercicio 1, otra bióloga molecular repite el estudio con un diseño diferente. Toma 12 muestras de ADN, y después elige seis aleatoriamente para tratarlas con la enzima y seis permanecen sin tratamiento. Los resultados son los siguientes:

Enzima presente:	12	15	14	22	22	20
Enzima ausente:	23	39	37	18	26	24

Determine un intervalo de confianza de 95% para la diferencia entre la media de los números de fragmentos.

- El artículo "Genetically Based Tolerance to Endosulfan, Chromium (VI) and Fluoranthene in the Grass Shrimp *Palaeomonetes pugio*" (R. Harper, Arabia, Tesis de doctorado, Escuela de Minas, de Colorado, 2002) informa que de 1 985 huevos producidos por el langostino en el sitio Diesel Creek en Charleston, Carolina del Sur, 1 919 incubaron, y en el sitio de Shipyard Creek, también en Charleston, de 4 561, 4 988 incubaron. Determine un intervalo de confianza de 99% para la diferencia entre las proporciones de huevos que se incubaron en los dos sitios.
- Se hicieron mediciones de resistencia en una muestra de 81 alambres de cierto tipo. La media muestral de resistencias era de 17.3 mΩ y la desviación estándar era de 1.2 mΩ.
 - Determine un intervalo de confianza de 98% para la media de la resistencia de este tipo de alambre.
 - ¿Cuál es el nivel del intervalo de confianza (17.1, 17.5)?

- ¿Cuántos alambres se deben probar con el propósito de que un intervalo de confianza de 98% especificará la media dentro de ± 0.1 mΩ?

- Una muestra de 125 piezas de hilo tuvo una media de fuerza de ruptura de 6.1 N y una desviación estándar de 0.7 N. Se hizo un nuevo grupo de hilo, usando nuevas materias primas de un distribuidor diferente. En una muestra de 75 piezas de hilo del nuevo grupo, la media de la fuerza de ruptura era de 5.8 N y la desviación estándar de 1.0 N. Encuentre un intervalo de confianza de 90% para la diferencia en la media de la fuerza de ruptura entre los dos tipos de hilos.

- Con referencia al ejercicio 5 se muestrean más piezas de hilo con el fin de mejorar la precisión del intervalo de confianza. ¿Cuál incrementaría más la precisión: muestrear 50 piezas más de hilo del grupo viejo, 50 piezas más provenientes del nuevo grupo, o 25 piezas más de cada uno de los grupos?

- La fuga de tanques de combustible subterráneos ha sido una fuente de contaminación del agua. En una muestra aleatoria de 87 estaciones de gasolina, se encontró que 13 tenían al menos un tanque subterráneo con fuga.

- Determine un intervalo de confianza de 95% para la proporción de estaciones de gasolina con, por lo menos, un tanque subterráneo con fuga.

- ¿Cuántas estaciones se deben muestrear para que un intervalo de confianza de 95% especifique la proporción dentro de ± 0.03 ?

- Un nuevo catalizador está siendo investigado para el uso en la producción de un plástico químico. Se producen diez grupos del producto químico. La media de la producción de los diez grupos es 72.5% y la desviación estándar es 5.8%. Suponga que las producciones son independientes y normalmente distribuidas. Encuentre un intervalo de confianza de 99% para la media de la producción cuando se usa un nuevo catalizador.

- Se calculan tres intervalos de confianza para la media de la fuerza de corte (en ksi) de pernos de anclaje de un tipo dado, todos de la misma muestra. Los intervalos son (4.01, 6.02), (4.20, 5.83) y (3.57, 6.46). Los niveles de los intervalos son 90, 95 y 99%. ¿Qué intervalo tiene cada nivel?

- Se realiza una encuesta en la que a una muestra aleatoria de residentes en cierta ciudad se les preguntará si están en fa-

vor o se oponen a la construcción de un nuevo estacionamiento en el centro. ¿A cuántos residentes se les debe preguntar para asegurar que un intervalo de confianza de 95% para la proporción que favorece la construcción específica que proporción está dentro de ± 0.05 ?

11. En el artículo “Groundwater Electromagnetic Imaging in Complex Geological and Topographical Regions: A Case Study of a Tectonic Boundary in the French Alps” (S. Houtot, P. Tarits y colaboradores, *Geophysics*, 2002:1048-1060), se midió el pH para varias muestras de agua en diferentes localizaciones cerca del lago Gittaz en los Alpes franceses. Los resultados para 11 localizaciones en el lado norte del lago y para seis localizaciones en el lado sur son las siguientes:

Lado norte:	8.1	8.2	8.1	8.2	8.2	7.4
	7.3	7.4	8.1	8.1	7.9	
Lado sur:	7.8	8.2	7.9	7.9	8.1	8.1

Determine un intervalo de confianza de 98% para la diferencia en pH entre el lado norte y el sur.

12. Los policlorobifenilos (PCB) son un grupo de químicos sintéticos parecidos al aceite que fueron en un tiempo ampliamente usados como aislantes de equipo eléctrico y se descargaban en ríos. Se descubrió que presentaban un riesgo para la salud y se prohibieron en la década de 1970. Desde entonces se ha dedicado mucho esfuerzo para dar seguimiento a las concentraciones de PCB en canales navegables. Suponga se están extrayendo muestras de agua de un canal navegable con el fin de calcular la concentración de PCB.

- Suponga que una muestra aleatoria de tamaño 80 tiene una media muestral de 1.69 ppb y desviación estándar muestral de 0.25 ppb. Determine un intervalo de confianza de 95% para la concentración de PCB.
- Calcule el tamaño muestral necesario con el propósito de que un intervalo de confianza de 95% especificará que la media poblacional está dentro de ± 0.02 ppb.

13. Se calcula un intervalo de confianza de 90% para una media poblacional basada en 144 observaciones de (2.7, 3.4). ¿Cuántas observaciones se debe hacer para que dicho intervalo de confianza de 90% especificara la media dentro de ± 0.2 ?

14. Se extrae una muestra de 100 componentes, y un intervalo de confianza de 95% para la proporción de componentes defectuosos especifica esta proporción dentro de ± 0.06 . Para obtener una estimación más precisa del número de com-

ponentes defectuosos, el tamaño muestral se incrementará a 400, y el intervalo de confianza será nuevamente calculado. ¿Cuál será el ancho aproximado del nuevo intervalo de confianza? Elija la mejor respuesta:

- ± 0.015
- ± 0.03
- ± 0.06
- ± 0.12
- ± 0.24

15. Un metalúrgico hace varias mediciones de la temperatura de fusión de cierta aleación y calcula un intervalo de confianza de 95% de $2\,038 \pm 2^\circ\text{C}$. Suponga que el proceso de medición de la temperatura no está sesgado. Verdadero o falso:

- Hay una probabilidad de 95% de que la temperatura de fusión verdadera esté en el intervalo $2\,038 \pm 2^\circ\text{C}$.
- Si el experimento fuera repetido, la probabilidad es de 95% de que la media de la medición de ese experimento estaría en el intervalo $2\,038 \pm 2^\circ\text{C}$.
- Si el experimento fuera repetido, y un intervalo de confianza de 95% se calcula, hay una probabilidad de 95% de que el intervalo de confianza contendría al punto de fusión verdadero.
- Si se hubiera hecho una medición más, la probabilidad es de 95% de que estaría en el intervalo $2\,038 \pm 2^\circ\text{C}$.

16. En un estudio de tiempos de vida de componentes electrónicos, una muestra aleatoria de 400 componentes es evaluado hasta que dejan de funcionar. La media muestral de los tiempos de vida era de 370 horas y la desviación estándar era de 650 horas. Verdadero o falso:

- Un intervalo de confianza de 95% aproximado para la media de los tiempos de vida de este tipo de componente es de 306.3 a 433.7 horas.
- Aproximadamente el 95% de los componentes de la muestra tenía tiempos de vida entre 306.3 y 433.7 horas.
- Si alguien toma una muestra aleatoria de 400 componentes, divide la desviación estándar muestral de sus tiempos de vida entre 20, y después suma y resta esa cantidad a la media muestral, existe una posibilidad de 68% de que el intervalo así construido contenga la media del tiempo de vida de este tipo de componente.
- La tabla z no se puede usar para construir intervalos de confianza aquí, porque los tiempos de vida de los componentes no siguen la curva normal.
- ¿Aproximadamente 68% de los componentes tenía tiempos de vida en el intervalo 370 ± 650 horas?

17. La temperatura de cierta solución se calcula tomando muchas mediciones independientes y determinando el promedio de ellas. El cálculo aproximado es de 37°C y la incertidumbre (desviación estándar) en esta estimación es 0.1°C .
- Determine un intervalo de confianza de 95% para la temperatura.
 - ¿Cuál es el nivel de confianza del intervalo $37 \pm 0.1^{\circ}\text{C}$?
 - Si sólo se había hecho una pequeña cantidad de mediciones independientes, ¿qué suposición adicional sería necesaria para calcular un intervalo de confianza?
 - Haciendo la suposición adicional, calcule un intervalo de confianza de 95% para la temperatura si hubieran hecho diez mediciones.
18. Cajas de clavos contienen 100 clavos cada una. Se extrae una muestra de diez cajas, y cada una de las cajas es pesada. El promedio del peso es de 1 500 g y la desviación estándar es de 5 g. Suponga que el peso de la caja misma es despreciable, por lo que todo el peso es atribuible a los clavos en la caja.
- Sea μ_{caja} la media del peso de una caja de clavos. Determine un intervalo de confianza de 95% para μ_{caja} .
 - Sea μ_{clavo} la media del peso de un clavo. Expresé μ_{caja} en función de μ_{clavo} .
 - Determine un intervalo de confianza de 95% para μ_{clavo} .
19. Sea X el número de eventos que se observa que ocurren en n unidades de tiempo o espacio, y suponga que $X \sim \text{Poisson}(n\lambda)$, donde λ es la media del número de eventos que ocurren en una unidad de tiempo o espacio. Suponga que X es grande, por lo que $X \sim N(n\lambda, n\lambda)$. Con los siguientes incisos del a) al d) se deduce un intervalo de confianza de nivel $100(1 - \alpha)\%$ para λ . Después, en el inciso e), se le pide que aplique el resultado encontrado en el inciso d).
- Muestre que para una proporción $1 - \alpha$ de todas las posibles muestras, $X - z_{\alpha/2}\sigma_X < n\lambda < X + z_{\alpha/2}\sigma_X$.
 - Sea $\hat{\lambda} = X/n$. Muestre que $\sigma_{\hat{\lambda}} = \sigma_X/n$.
 - Concluya que para una proporción $1 - \alpha$ de todas las muestras posibles, $\hat{\lambda} - z_{\alpha/2}\sigma_{\hat{\lambda}} < \lambda < \hat{\lambda} + z_{\alpha/2}\sigma_{\hat{\lambda}}$.
 - Utilice el hecho de que $\sigma_{\hat{\lambda}} \approx \sqrt{\hat{\lambda}/n}$ para deducir una expresión para un intervalo de confianza de nivel $100(1 - \alpha)\%$ para λ .
 - Una muestra de 5 ml de cierta suspensión contiene 300 partículas. Sea λ la media del número de partículas por ml en la suspensión. Determine un intervalo de confianza de 95% para λ .
20. Es necesaria la respuesta al inciso d) del ejercicio 19 para la realización de este ejercicio. Un geólogo cuenta 64 partículas emitidas en un minuto por cierta roca radiactiva.
- Determine un intervalo de confianza de 95% para la tasa de emisiones en unidades de partículas por minuto.
 - Después de cuatro minutos, se cuentan 256 partículas. Determine un intervalo de confianza de 95% para la tasa de emisiones en unidades de partículas por minuto.
 - ¿Por cuántos minutos se debe contar los errores para que el intervalo de confianza de 95% especifique la tasa dentro de una partícula por minuto?
21. En un flujo de Couette, dos placas planas grandes caen una encima de otra, separadas por una capa fina de fluido. Si se aplica una fuerza de corte a la placa superior, la viscosidad del fluido también produce movimiento en la placa inferior. La velocidad V en la placa superior con respecto a la placa inferior está dada por $V = \tau h/\mu$, donde τ es la tensión de corte aplicada en la placa superior, h es el espesor de la capa de fluido y μ es la viscosidad del fluido.
- Suponga que μ , h y τ son medidos por separado y que las mediciones no tienen sesgos y están distribuidas normalmente. Los valores medidos son $\mu = 1.6 \text{ Pa} \cdot \text{s}$, $h = 15 \text{ mm}$ y $\tau = 25 \text{ Pa}$. Las incertidumbres (desviaciones estándares) de estas mediciones son $\sigma_{\mu} = 0.05$, $\sigma_h = 1.0$ y $\sigma_{\tau} = 1.0$.
- Utilice el método de la propagación del error para calcular V y su incertidumbre σ_V .
 - Suponiendo que la estimación de V esté distribuida normalmente, determine un intervalo de confianza de 95% para V .
 - Realice una simulación para determinar si es o no válido el intervalo de confianza encontrado en el inciso b).
22. Una muestra de siete bloques de concreto tenía su fuerza de compresión medida en MPa. Los resultados fueron
- | | | | | |
|---------|---------|---------|---------|---------|
| 1 367.6 | 1 411.5 | 1 318.7 | 1 193.6 | 1 406.2 |
| 1 425.7 | 1 572.4 | | | |
- Diez mil muestras de estimación bootstrap se generaron a partir de estos datos, y las medias de la estimación bootstrap fueron arregladas en orden. Denotando a la media más pequeña por Y_1 , la segunda más pequeña por Y_2 , y así sucesivamente, siendo la más grande $Y_{10\,000}$. Suponga que
- | | | |
|---------------------------|---------------------------|---------------------------|
| $Y_{50} = 1\,283.4$, | $Y_{51} = 1\,283.4$, | $Y_{100} = 1\,291.5$, |
| $Y_{101} = 1\,291.5$, | $Y_{250} = 1\,305.5$, | $Y_{251} = 1\,305.5$, |
| $Y_{500} = 1\,318.5$, | $Y_{501} = 1\,318.5$, | $Y_{9\,500} = 1\,449.7$, |
| $Y_{9\,501} = 1\,449.7$, | $Y_{9\,750} = 1\,462.1$, | $Y_{9\,751} = 1\,462.1$, |
| $Y_{9\,900} = 1\,476.2$, | $Y_{9\,901} = 1\,476.2$, | $Y_{9\,950} = 1\,483.8$ y |
| $Y_{9\,951} = 1\,483.8$. | | |

- a) Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 1, como se describe en la p. 358.
 - b) Calcule un intervalo de confianza de estimación bootstrap de 95% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.
 - c) Calcule un intervalo de confianza de estimación bootstrap de 99% para la media de la fuerza compresiva, usando el método 1, como se describe en la p. 358.
 - d) Calcule un intervalo de confianza de estimación bootstrap de 99% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.
- 23.** Con referencia al ejercicio 22.
- a) Genere 10 000 muestras de estimación bootstrap de los datos del ejercicio 22. Encuentre la media de estimación bootstrap en la muestra de los percentiles que se usan para calcular un intervalo de confianza de 99 por ciento.
 - b) Calcule un intervalo de confianza de estimación bootstrap de 99% para la media de la fuerza compresiva, usando el método 1, como se describe en la p. 358.
 - c) Calcule un intervalo de confianza de estimación bootstrap de 99% para la media de la fuerza compresiva, usando el método 2, como se describe en la p. 358.



Capítulo 6

Pruebas de hipótesis

Introducción

En el ejemplo 5.2 (sección 5.1) una muestra de 50 microperforadoras tenía un tiempo de vida promedio de $\bar{X} = 12.68$ huecos perforados con una desviación estándar de $s = 6.83$. Supóngase que la cuestión principal es responder si la media poblacional μ de los tiempos de vida es o no mayor de 11. Esta cuestión se aborda cuando se examina el valor de la media muestral $\bar{X}$. Se nota que $\bar{X} > 11$, pero debido a la incertidumbre en $\bar{X}$ no se garantiza que $\mu > 11$. Nos gustaría conocer con qué certeza consideramos que $\mu > 11$. Un intervalo de confianza no es todo lo que se necesita. En el ejemplo 5.2 se calculó un intervalo de confianza de 95% para la media poblacional μ de (10.79, 14.57). Éste indica una confianza de 95% de que μ está entre 10.79 y 14.57. Pero el intervalo de confianza no indica directamente cuánta confianza se tiene de que $\mu > 11$.

El enunciado “ $\mu > 11$ ” constituye una **hipótesis** acerca de la media poblacional μ . Con el propósito de determinar cómo se logra la certeza de que una hipótesis semejante es verdadera se debe realizar una **prueba de hipótesis**. Esta última produce un número entre 0 y 1 que mide el grado de certeza que se puede tener de la validez de una hipótesis con respecto a una cantidad, como una media o proporción de la población. Lo anterior da como resultado que las pruebas de hipótesis estén estrechamente relacionadas con los intervalos de confianza. En general, siempre que se pueda calcular un intervalo de confianza, se puede realizar una prueba de hipótesis, y viceversa.

6.1 Pruebas de hipótesis para la media poblacional con muestras grandes

Se inicia con un ejemplo. Cierta tipo de motor de automóvil emite una media de 100 mg de óxidos de nitrógeno (NO_x) por segundo con 100 caballos de fuerza. Se ha propuesto una modificación al diseño del motor para reducir las emisiones de NO_x . El nuevo diseño se produ-

cirá si se demuestra que la media de su tasa de emisiones es menor de 100 mg/s. Se construye y se prueba una muestra de 50 motores modificados. La media muestral de emisiones de NO_x es de 92 mg/s, y la desviación estándar muestral es de 21 mg/s.

La población, en este caso, consta de las tasas de emisiones de los motores que serían contruidos si se fabricara este diseño modificado. Si no hubiera incertidumbre en la media muestral, entonces se podría concluir que la modificación reduciría las emisiones de 100 a 92 mg/s. Por supuesto, hay incertidumbre en la media muestral. La media poblacional en realidad será poco mayor o menor de 92.

Los fabricantes están preocupados de que los motores modificados no puedan reducir todas las emisiones; es decir, que la media poblacional pudiera ser 100 o mayor. Quieren saber si esta preocupación está justificada. La pregunta, por tanto, es: ¿Es factible que esta muestra, con media de 92, pudiera provenir de una población cuya media es 100 o mayor?

Éste es el tipo de preguntas que las pruebas de hipótesis están diseñadas para responder, y ahora se construirá una prueba de hipótesis para responderla. Se ha observado una muestra con media 92. Hay dos interpretaciones posibles de esta observación:

1. La media poblacional es realmente mayor que o igual a 100, y la media muestral es menor que ésta sólo debido a la variación aleatoria de la media poblacional. Por tanto, las emisiones no bajarán si el nuevo diseño de producciones produce, y la muestra es engañosa.
2. La media poblacional es en realidad menor que 100, y la media muestral refleja este hecho. Por tanto, la muestra representa una diferencia verdadera que se puede esperar si el nuevo diseño se fabrica.

Estas dos explicaciones tienen nombres comunes. La primera se llama **hipótesis nula**. En la mayoría de las situaciones, la hipótesis nula establece que el efecto que indica la muestra es atribuible solamente a la variación aleatoria entre la muestra y la población. La segunda se denomina **hipótesis alternativa**. Esta hipótesis alternativa establece que el efecto que indica la muestra es verdadero, ya que representa a toda la población.

En este ejemplo, los fabricantes de motores están preocupados de que la hipótesis nula pudiera ser verdadera. Una prueba de hipótesis asigna una medida cuantitativa a la factibilidad de la hipótesis nula. Después de realizar una prueba de hipótesis, se podría decir a los fabricantes, en términos numéricos, qué tan válida es su preocupación.

Para hacer las cosas más precisas, todo se expresa mediante símbolos. La hipótesis nula se denota con H_0 . La hipótesis alternativa se denota mediante H_1 . Como es usual, la media poblacional es μ . Por tanto, se tiene,

$$H_0: \mu \geq 100 \quad \text{contra} \quad H_1: \mu < 100$$

Esencialmente, para realizar una prueba de hipótesis se pone la hipótesis nula en juicio. Se empieza suponiendo que H_0 es verdadera, de la misma manera como se empieza un juicio bajo el supuesto de que un acusado es inocente. La muestra aleatoria proporciona la evidencia. La prueba de hipótesis implica medir la fuerza del desacuerdo entre la muestra y H_0 para producir un número entre 0 y 1, llamado **P-valor**. Éste mide la factibilidad de H_0 . Entre menor sea el P-valor, más fuerte será la evidencia en contra de H_0 . Si el P-valor es suficientemente pequeño, se puede estar dispuesto a abandonar la suposición de que H_0 es verdadera y creer, en su lugar, que H_1 es verdadera. Lo anterior se llama **rechazar** la hipótesis nula.

En este ejemplo, sean $X_1, \dots, X_{50}$ las tasas de emisiones medidas de los 50 motores de la muestra. El valor observado de la media muestral es $\bar{X} = 92$. También se necesitará conocer la desviación estándar muestral, que es $s = 21$. Se debe evaluar la factibilidad de H_0 , que dice que la media poblacional es de 100 o mayor, dado que se ha observado una muestra de esta población, cuya media es solamente 92. Dicho procedimiento se realizará en dos pasos, de la siguiente manera:

1. Se calculará a la distribución de $\bar{X}$ bajo el supuesto de que H_0 es verdadera. Esta distribución se llama **distribución nula** de $\bar{X}$.
2. Se calculará el P -valor. Éste es la probabilidad, bajo el supuesto que H_0 sea verdadera, de observar un valor de $\bar{X}$, cuyo desacuerdo con H_0 sea al menos tan grande como el valor observado de 92.

Para realizar el paso 1, observe que $\bar{X}$ es la media de una muestra grande, así que el teorema del límite central especifica que proviene de una distribución normal, cuya media es μ y su varianza es $\sigma^2/50$, donde σ^2 es la varianza poblacional y 50 el tamaño de muestra. Se deben especificar los valores para μ y para σ con el fin de determinar la distribución nula. En virtud de que se está suponiendo que H_0 es verdadera, se supone que $\mu \geq 100$. Esto último no proporciona un valor específico para μ . Se toma como el valor supuesto para μ el valor más cercano a la hipótesis alternativa H_1 , por razones que se explicarán posteriormente en esta sección. Por tanto, se supone que $\mu = 100$. No se conoce la desviación estándar poblacional σ . Sin embargo, debido a que la muestra es grande, se puede aproximar a σ con la desviación estándar muestral $s = 21$. Por tanto, se ha determinado que bajo H_0 , $\bar{X}$ tiene una distribución normal con media 100 y desviación estándar $21/\sqrt{50} = 2.97$. La distribución nula es $\bar{X} \sim N(100, 2.97^2)$.

Ahora se está listo para el paso 2. La figura 6.1 muestra la distribución nula. El número 92 indica el punto de la distribución correspondiente al valor observado de $\bar{X}$. ¿Qué tan factible es que un número muestreado de esta distribución sea tan pequeño como 92? Lo anterior se mide con el P -valor. Éste constituye la probabilidad de que un número extraído de la distribución nula esté en desacuerdo con H_0 al menos tan intensamente como el valor observado de $\bar{X}$ que de 92. Dado que H_0 especifica que la media de $\bar{X}$ es mayor que o igual a 100, los valores menores de 92 están en desacuerdo mayor con H_0 . Por tanto, el P -valor es la probabilidad de que un número extraído de una distribución $N(100, 2.97^2)$ sea menor o igual a 92. Esta probabilidad se determina al calcular el puntaje z :

$$z = \frac{92 - 100}{2.97} = -2.69$$

De la tabla z , la probabilidad de que una variable aleatoria normal estándar z sea menor que o igual a -2.69 es 0.0036. El P -valor para esta prueba es 0.0036.

Como promete, el P -valor proporciona una medida cuantitativa de la factibilidad de H_0 . ¿Pero cómo se interpreta esta cantidad? La interpretación adecuada es algo sutil. El P -valor indica que si H_0 fuera verdadera, la probabilidad de extraer una muestra cuya media esté tan lejos de H_0 como el valor observado de 92 es solamente 0.0036. Por tanto, es posible una de las siguientes dos conclusiones:

- H_0 es falsa.
- H_0 es verdadera, lo que implica que de todas las muestras que pudieron haber extraído, solamente el 0.36% tiene una media igual o más pequeña que la muestra que en realidad se extrajo. En otras palabras, la media muestral está en 0.36% más extremo de su distribución.

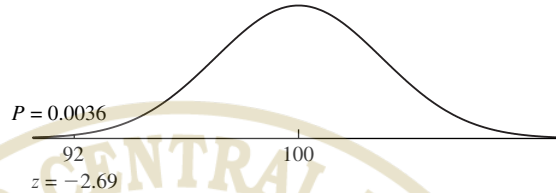


FIGURA 6.1 La distribución nula de $\bar{X}$ es $N(100, 2.97^2)$. Por tanto, si H_0 es verdadera, la probabilidad de que $\bar{X}$ tome un valor tan extremo o más extremo que el valor observado de 92 es 0.0036. Éste es el P -valor.

En la práctica, los eventos en el 0.36% más extremo de sus distribuciones muy raramente ocurren. Por tanto, se rechaza H_0 y se concluye que los nuevos motores tendrán emisiones menores.

La hipótesis nula en este caso especifica solamente que $\mu \geq 100$. Suponiendo que H_0 es verdadera, ¿por qué se eligió el valor $\mu = 100$, que es el más cercano a H_1 ? Para dar una prueba justa a la H_0 , se debe evaluar en su forma más factible. El valor más factible para μ , suponiendo que H_0 sea verdadera, es el valor más cercano a H_1 . Para observar lo anterior, vea la figura 6.1. Suponga que se ha elegido un valor para μ mayor a 100 para representar a H_0 . Entonces la distribución nula tendría que estar corrida a la derecha. Esto último causaría que la media muestral de 92 esté aún más lejos en la cola. Entonces el P -valor tendría que ser aún menor. Por tanto, entre los valores de μ consistentes con H_0 , el más cercano a H_1 tiene el P -valor mayor, y, por tanto, es el más factible. Por esta razón, cuando se supone que H_0 es verdadera, siempre se utiliza el valor del parámetro más cercano a H_1 cuando se realiza una prueba de hipótesis.

Es natural preguntar qué tan pequeño debe ser el P -valor con la finalidad de rechazar H_0 . Algunas personas usan la “regla del 5%”; ellas rechazan H_0 si $P \leq 0.05$. Sin embargo, no hay ninguna justificación científica para ésta o cualquier otra regla. Esta cuestión se analiza más profundamente en la sección 6.2.

Observe que dicho método usa el teorema del límite central. Por eso para que este método sea válido, el tamaño muestral debe ser razonablemente grande, de 30 o mayor. En la sección 6.4 se presentan las pruebas de hipótesis que algunas veces son válidas para muestras pequeñas.

Finalmente, observe que el cálculo del P -valor se hizo al calcular un puntaje z . Por tanto, este puntaje se llama **estadístico de prueba**. Una prueba que usa un puntaje z como un estadístico de prueba se denomina prueba z .

Hay muchas clases de pruebas de hipótesis. Todas siguen una serie básica de pasos, que se ilustran en el cuadro de la página siguiente.

Pasos para la realización de una prueba de hipótesis

1. Defina H_0 y H_1 .
2. Suponga que H_0 es verdadera.
3. Calcule un **estadístico de prueba**. Éste constituye un estadístico que se usa para evaluar la fuerza de la evidencia en contra de H_0 .
4. Calcule el P -valor del estadístico de prueba. El P -valor es la probabilidad, suponiendo que H_0 es verdadera, de que el estadístico de prueba tenga un valor cuya diferencia con H_0 sea tan grande o mayor que el realmente observado. El P -valor también se llama **nivel de significancia observado**.

Ejemplo

6.1

El artículo “Wear in Boundary Lubrication” (S. Hsu, R. Munro y M. Shen, en *Journal of Engineering Tribology*, 2002:427-441) analiza algunos experimentos que implican diferentes lubricantes. En un experimento, 45 bolas de acero, lubricadas con parafina purificada, estaban sujetas a una carga de 40 kg a 600 rpm durante 60 minutos. El promedio de desgaste, medido por la reducción en el diámetro, era de $673.2 \mu\text{m}$, y la desviación estándar era de $14.9 \mu\text{m}$. Suponga que la especificación para un lubricante es que la media del desgaste sea menor de $675 \mu\text{m}$. Determine un P -valor para probar $H_0: \mu \geq 675$ contra $H_1: \mu < 675$.

Solución

Primero se traduce el problema al lenguaje estadístico. Se tiene una muestra aleatoria simple $X_1, \dots, X_{45}$ de diámetros de desgaste. La media muestral y la desviación estándar son $\bar{X} = 673.2$ y $s = 14.9$. La media poblacional es desconocida y se denota con μ . Antes de entrar a la construcción de la prueba, se indicará nuevamente que la idea básica es la incertidumbre en la media muestral. Si no hubiera incertidumbre se concluiría que el lubricante satisfaría la especificación, ya que $673.2 < 675$. La cuestión es si la incertidumbre en la media muestral es suficientemente grande como para que pudiera ser factible que la media poblacional fuera tan grande como 675.

Para realizar la prueba de hipótesis se siguen los pasos anteriores. La hipótesis nula es que el lubricante no satisface la especificación, y que la diferencia entre la media muestral de 673.2 y 675 es consecuencia de la aleatoriedad. La hipótesis alternativa es que el lubricante efectivamente satisface la especificación.

Se supone que H_0 es verdadera, ya que la muestra se extrajo de una población con media $\mu = 675$ (el valor más cercano a H_1). Se calcula la desviación estándar poblacional σ con la desviación estándar muestral $s = 14.9$. La prueba está basada en $\bar{X}$. Bajo H_0 , $\bar{X}$ proviene de una población normal con media 675 y desviación estándar $14.9/\sqrt{45} = 2.22$. El P -valor es la probabilidad de observar una media muestral menor o igual a 673.2. El estadístico de prueba es el puntaje z , que es

$$z = \frac{673.2 - 675}{2.22} = -0.81$$

El P -valor es 0.209 (véase la figura 6.2). Por tanto, si H_0 es verdadera, hay una probabilidad de 20.9% de observar una muestra cuyo desacuerdo con H_0 es al menos tan grande como la que en realidad fue observada. Puesto que 0.209 no es una probabilidad muy pequeña, H_0 no se rechaza. En vez de esto último, se concluye que H_0 es factible. Los datos no muestran que el lubricante satisfaga la especificación. Observe que *no* se está concluyendo que la H_0 sea *verdadera*, solamente que es *factible*. En la sección 6.2 se analizará esta diferencia con mayor detalle.

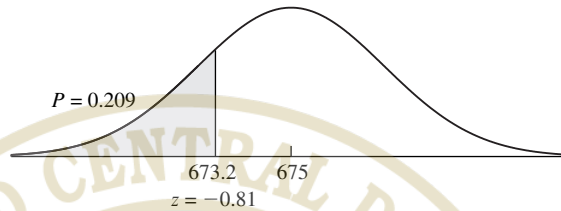


FIGURA 6.2 La distribución nula de $\bar{X}$ es $N(675, 2.22^2)$. Por tanto, si H_0 es verdadera, la probabilidad que $\bar{X}$ tome un valor tan extremo o más extremo que el valor observado de 673.2 es 0.209. Éste es el P -valor.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.1.

One-Sample Z: Wear							
Test of $\mu = 675$ vs < 675							
The assumed standard deviation = 14.9							
					95%		
					Upper		
Variable	N	Mean	StDev	SE Mean	Bound	Z	P
Wear	45	673.200	14.9	2.221	676.853	-0.81	0.209

El resultado establece la hipótesis nula como $\mu = 675$ en vez de $\mu \geq 675$. Esto último refleja que para construir la distribución nula se utilizó el valor $\mu = 675$. La cantidad “SE Mean” representa la desviación estándar de $\bar{X}$, calculado con $s/\sqrt{n}$. El resultado también proporciona un límite superior de confianza de 95% para μ .

En los ejemplos anteriores, la hipótesis nula especifica que la media poblacional era menor que o igual a algo, o mayor que o igual a algo. En algunos casos, una hipótesis nula especifica que la media poblacional es igual a un valor específico. El ejemplo 6.2 ilustra el caso.

Ejemplo

6.2

Se está calibrando una balanza al pesar una pesa de prueba de 1 000 g 60 veces. Las 60 lecturas de la balanza tienen una media de 1 000.6 g y desviación estándar de 2 g. Determine el P -valor para la prueba $H_0: \mu = 1\,000$ contra $H_1: \mu \neq 1\,000$.

Solución

Sea μ la media poblacional de las lecturas. La hipótesis nula dice que la balanza está calibrada, por lo que la media poblacional μ es igual al peso verdadero de 1 000 g, y la diferencia entre la media muestral de las lecturas y el peso verdadero es atribuible completamente al aspecto aleatorio. La hipótesis alternativa dice que la balanza no está calibrada.

En este ejemplo la hipótesis nula especifica que μ es *igual* a un valor específico, en lugar de ser mayor o igual que o menor o igual que un valor específico. Por esta razón, los valores de $\bar{X}$, que son mucho mayores o mucho menores que μ , proporcionarán evidencia en contra de H_0 . En los ejemplos anteriores, sólo los valores de $\bar{X}$ en un lado de μ proporcionan evidencia en contra de H_0 .

Se supone que H_0 es verdadera, y que, por tanto, las lecturas de la muestra se extrajeron de una población con media $\mu = 1\,000$. La desviación estándar poblacional σ se aproxima con $s = 2$. La distribución nula de $\bar{X}$ es normal con media de 1 000 y desviación estándar $2/\sqrt{60} = 0.258$. El puntaje z del valor observado $\bar{X} = 1\,000.6$ es

$$z = \frac{1\,000.6 - 1\,000}{0.258} = 2.32$$

Dado que H_0 especifica que $\mu = 1\,000$, las regiones en ambas colas de la curva están en mayor desacuerdo con H_0 que el valor observado de 1 000.6. El P -valor es la suma de las áreas de las dos colas, que es 0.0204 (véase la figura 6.3). Por tanto, si H_0 es verdadera, la probabilidad de un resultado mayor o igual que el valor observado es sólo de 0.0204. Las evidencias en contra de H_0 son muy fuertes. Sería prudente rechazar H_0 y recalibrar la balanza.

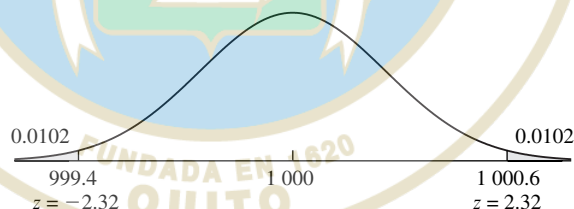


FIGURA 6.3 La distribución nula de $\bar{X}$ es $N(1\,000, 0.258^2)$. Por lo que si H_0 es verdadera, la probabilidad de que $\bar{X}$ tome un valor tan extremo o más extremo que el valor observado de 1 000.6 es 0.0204. Éste es el P -valor.

Cuando H_0 especifica un solo valor para μ , como en el ejemplo 6.2, ambas colas contribuyen al P -valor, y se dice que la prueba es una prueba de **dos lados** o de **dos colas**. Cuando H_0 sólo especifica que μ es mayor o igual que, o menor que o igual a un valor, sólo contribuye una cola al P -valor, y la prueba se llama de **un lado** o de **una cola**.

Se concluye esta sección con el resumen del procedimiento utilizado para realizar una prueba de hipótesis para la media poblacional con muestras grandes.

Resumen

Sea $X_1, \dots, X_n$ una muestra *grande* (por ejemplo, $n > 30$) de una población con la media μ y desviación estándar σ .

Para probar una hipótesis nula de la forma $H_0: \mu \leq \mu_0$, $H_0: \mu \geq \mu_0$, o $H_0: \mu = \mu_0$:

- Calcule el puntaje z : $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$.

Si σ es desconocida se puede aproximar con s .

- Calcule el P -valor. Éste constituye un área bajo la curva normal, que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: \mu > \mu_0$$

$$H_1: \mu < \mu_0$$

$$H_1: \mu \neq \mu_0$$

P -valor

Área a la derecha de z

Área a la izquierda de z

Suma de áreas en las colas correspondientes a z y $-z$

Ejercicios para la sección 6.1

- Recientemente muchas compañías han experimentado con las “telecomunicaciones”, al permitir que sus empleados trabajen en su casa en sus computadoras. Entre otros factores, se supone que la telecomunicación reduce las faltas por enfermedad. Suponga que en una compañía se sabe que en años pasados los empleados faltaron una media de 5.4 días por enfermedad. Este año la compañía introduce las telecomunicaciones. La dirección elige una muestra aleatoria simple de 80 empleados para estudiarlos en detalle, y, al final del año, éstos promedian 4.5 días de faltas por enfermedad con desviación estándar de 2.7 días. Sea μ la media del número de días de faltas por enfermedad para todos los empleados de la compañía.
 - Determine el P -valor para probar $H_0: \mu \geq 5.4$ contra $H_1: \mu < 5.4$.
 - Ya sea que la media del número de días de faltas por enfermedad haya disminuido desde la introducción de las telecomunicaciones, o que la muestra esté en el _____% más extremo de esta distribución.
- Una muestra aleatoria simple consta de 65 longitudes de alambre de piano que se probaron para la cantidad de alargamiento bajo una carga de 30 N. El alargamiento promedio para los 65 alambres era de 1.102 mm, y la desviación estándar, de 0.020 mm. Sea μ la media del alargamiento para todas las unidades de alambre de piano.
 - Determine el P -valor para probar $H_0: \mu \leq 1.1$ contra $H_1: \mu > 1.1$.
 - Ya sea que la media del alargamiento para este tipo de alambre sea mayor que 1.1 mm, o la media muestral está en _____% más extremo de esta distribución.
- El artículo “Evaluation of Mobile Mapping Systems for Roadway Data Collection” (H. Karimi, A. Khattak y J. Hummer, en *Journal of Computing in Civil Engineering*, 2000:168-173) describe un sistema para medir a distancia elementos de avenidas, como el ancho de los carriles y las alturas de las señales de tránsito. Para una muestra de 160 de esos elementos, el error promedio (en porcentaje) en las mediciones era de 1.90, con desviación estándar de 21.20. Sea μ la media del error en este tipo de medición.
 - Determine el P -valor para probar $H_0: \mu = 0$ contra $H_1: \mu \neq 0$.
 - Ya sea que la media del error para este tipo de medición es diferente a cero, o que la muestra esté en _____% más extremo de su distribución.
- Un inspector midió el volumen de llenado de una muestra aleatoria simple de 100 latas de jugo cuya etiqueta afirmaba que contenían 12 oz. La muestra tenía una media de volumen de 11.98 oz y desviación estándar de 0.19 oz. Sea μ la media del volumen de llenado para todas las latas de jugo.

recientemente llenadas con esta máquina. El inspector probará $H_0: \mu = 12$ contra $H_1: \mu \neq 12$.

- a) Determine el P -valor.
 - b) ¿Piensa que es factible que la media del volumen de llenado es de 12 oz? Explique su razonamiento.
5. Cuando está operando adecuadamente, una planta química tiene una media de producción diaria de cuando menos 740 toneladas. La producción se mide en una muestra aleatoria simple de 60 días. La muestra tenía una media de 715 toneladas/día y desviación estándar de 24 toneladas/día. Sea μ la media de la producción diaria de la planta. Un ingeniero prueba que $H_0: \mu \geq 740$ contra $H_1: \mu < 740$.
- a) Determine el P -valor.
 - b) ¿Piensa que es factible que la planta esté operando adecuadamente o está convencido de que la planta no funciona en forma adecuada? Explique su razonamiento.
6. En un proceso de fabricación de placas de silicio cubiertas con tungsteno la resistencia objetivo para una placa es de 85 m Ω . En una muestra aleatoria simple de 50 placas, la media de la resistencia de la muestra era de 84.8 de m Ω , y la desviación estándar, 0.5 m Ω . Sea μ la media de la resistencia de las placas fabricadas mediante este proceso. Un ingeniero de control de calidad prueba que $H_0: \mu = 85$ contra $H_1: \mu \neq 85$.
- a) Determine el P -valor.
 - b) ¿Piensa que es factible que la media esté en el objetivo, o está convencido de que la media no está en el objetivo? Explique su razonamiento.
7. Un filtro de aire ultrabajo de partículas se utiliza para mantener uniforme el flujo de aire en áreas de producción en una habitación limpia. Ahí la media de la velocidad del aire debe ser al menos de 40 cm/s. Cierta distribuidor probó una muestra aleatoria simple de 58 filtros. La media de la velocidad de la muestra era de 39.6 cm/s, con desviación estándar de 7 cm/s. Sea μ la media de la velocidad del aire obtenida por filtros suministrados por este distribuidor. Se hace una prueba de $H_0: \mu \geq 40$ contra $H_1: \mu < 40$.
- a) Determine el P -valor.

b) ¿Piensa que es factible que la media de velocidad es de al menos 40 cm/s, o está convencido de que la media es menor de 40 cm/s? Explique su razonamiento.

8. Se está preparando una nueva mezcla de concreto con el fin de proporcionar una fuerza compresiva adecuada para bloques de concreto. La especificación para una aplicación determinada pide que los bloques tengan una media de la fuerza compresiva μ mayor de 1 350 kPa. Se produce y se prueba una muestra de 100 bloques. Su media de fuerza compresiva es de 1 356 kPa, y su desviación estándar, de 70 kPa. Se hace una prueba de $H_0: \mu \leq 1\,350$ contra $H_1: \mu > 1\,350$.
 - a) Determine el P -valor.
 - b) ¿Piensa que es factible que los bloques no satisfagan la especificación, o está convencido de que sí lo hacen? Explique su razonamiento.
9. Complete el enunciado: Si la hipótesis nula es $H_0: \mu \leq 5$, entonces la media de $\bar{X}$ bajo la distribución nula es _____.
 - i) 0
 - ii) 5
 - iii) Cualquier número menor o igual que 5.
 - iv) No se puede decir a menos que se conozca H_1 .
10. Complete el enunciado: en una prueba de $H_0: \mu \geq 10$ contra $H_1: \mu < 10$, la media muestral fue de $\bar{X} = 8$ y el P -valor era de 0.04. Esto significa que si $\mu = 10$, y el experimento se repitiera 100 veces, se esperaría obtener un valor de $\bar{X}$ de 8 o menor aproximadamente _____ veces.
 - i) 8
 - ii) 0.8
 - iii) 4
 - iv) 0.04
 - v) 80
11. Un ingeniero toma gran número de mediciones independientes de la longitud de un componente y obtiene $\bar{X} = 5.2$ mm y $\sigma_{\bar{X}} = 0.1$ mm. Utilice esta información para encontrar el P -valor que pruebe que $H_0: \mu = 5.0$ contra $H_1: \mu \neq 5.0$.

12. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para una media poblacional μ .

One-Sample Z: X

Test of mu = 73.5 vs not = 73.5

The assumed standard deviation = 2.3634

Variable	N	Mean	StDev	SE Mean	95% CI	Z	P
X	145	73.2461	2.3634	0.1963	(72.8614, 73.6308)	-1.29	0.196

- a) ¿Es ésta una prueba de una cola o de dos colas?
- b) ¿Cuál es la hipótesis nula?
- c) ¿Cuál es el P -valor?
- d) Utilice el resultado y una tabla adecuada para calcular el P -valor para la prueba de $H_0: \mu \geq 73.6$ contra $H_1: \mu < 73.6$
- e) Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 99% para μ .
13. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para una media poblacional μ . Faltan algunos de los números. Complete los números del (a) al (c).

One-Sample Z: X

Test of mu = 3.5 vs > 3.5

The assumed standard deviation = 2.00819

Variable	N	Mean	StDev	SE Mean	95% Lower Bound	Z	P
					Bound		
X	87	4.07114	2.00819	(a)	3.71700	(b)	(c)

6.2 Concluir a partir de las pruebas de hipótesis

Se revisarán con más detalle las conclusiones obtenidas en los ejemplos 6.1 y 6.2 (de la sección 6.1). En el ejemplo 6.2 se rechazó H_0 ; en otras palabras, se concluyó que H_0 era falsa. En el ejemplo 6.1 H_0 no se rechazó. Sin embargo, tampoco se concluyó que H_0 fuera verdadera. Sólo se puede concluir que H_0 es factible.

En efecto, sólo se pueden tener dos conclusiones en una prueba de hipótesis: que H_0 es falsa o que H_0 es factible. En particular, nunca se puede llegar a la conclusión de que H_0 es verdadera. Para comprender por qué, piense nuevamente en el ejemplo 6.1. La media muestral era $\bar{X} = 673.2$ y la hipótesis nula era $\mu \geq 675$. La conclusión era que 673.2 está suficientemente cerca de 675, por lo que la hipótesis nula *podría* ser verdadera. Pero una media muestral de 673.2 obviamente no conduce a que se concluya que $\mu \geq 675$ es verdadera, ya que 673.2 es menor que 675. Éste es un caso típico en muchas situaciones de interés. El estadístico de prueba es consistente con la hipótesis alternativa y está un poco en desacuerdo con la

nula. La única cuestión es si el nivel de desacuerdo medido con el P -valor es lo suficientemente grande para presentar la hipótesis nula como no factible.

¿Cómo se sabe cuándo rechazar H_0 ? Entre menor sea el P -valor, menos factible es H_0 . Una regla general consiste en dibujar una línea en el 5%. De acuerdo con esta regla, si $P \leq 0.05$, H_0 se rechaza; de otra manera, H_0 no se rechaza. Efectivamente, no hay una línea que divida con nitidez las evidencias concluyentes en contra de H_0 y las evidencias no concluyentes, así como no hay una línea que divida nítidamente el clima caluroso y el frío. Así que aunque esta regla general es conveniente, no tiene ninguna justificación científica verdadera.

Resumen

- Entre menor sea el P -valor, se puede tener más certeza de que H_0 es falsa.
- Entre mayor sea el P -valor, es más factible H_0 , pero nunca se puede tener la certeza de que H_0 sea verdadera.
- Una regla general indica rechazar H_0 cada vez que $P \leq 0.05$. Aunque esta regla es conveniente, no tiene ninguna base científica.

Significancia estadística

Cada vez que el P -valor es menor que un umbral específico, el resultado indica que es “significativo estadísticamente” a ese nivel. Por lo que, por ejemplo, si $P \leq 0.05$, el resultado es estadísticamente significativo a un nivel de 5%; si $P \leq 0.01$, el resultado es estadísticamente significativo a un nivel de 1%, y así sucesivamente. Si un resultado es estadísticamente significativo a un nivel de $100\alpha\%$, también se puede decir que la hipótesis nula se “rechaza a un nivel de $100\alpha\%$ ”.

Ejemplo

6.3

Se realiza una prueba de la hipótesis nula $H_0: \mu = 0$. El P -valor da un resultado de 0.03. ¿Es el resultado estadísticamente significativo a un nivel de 10%? ¿A un nivel de 5%? ¿A un nivel de 1%? ¿La hipótesis nula es rechazada a un nivel de 10%? ¿A un nivel de 5%? ¿A un nivel de 1%?

Solución

El resultado es estadísticamente significativo en cualquier nivel igual o mayor a 3%. Por tanto, es estadísticamente significativo en los niveles 10 y 5%, pero no a un nivel de 1%. De forma semejante, se puede rechazar la hipótesis nula en cualquier nivel igual o mayor al 3%, así H_0 se rechaza en los niveles 10 y 5%, pero no a un nivel de 1%.

A veces las personas sólo informan que un resultado de prueba era estadísticamente significativo en un nivel específico, sin dar el P -valor. Por ejemplo, es común leer que un resultado era “estadísticamente significativo a un nivel de 5%” o “estadísticamente significativo ($P < 0.05$)”. Ésta es una mala práctica, por tres razones. Primera, no proporciona ninguna manera de decir si el P -valor era sólo apenas menor que 0.05, o si era mucho menor. Segunda,

notificar que un resultado era estadísticamente significativo a un nivel de 5% implica que hay gran diferencia entre un P -valor justo abajo de 0.05 y uno justo arriba de 0.05, cuando efectivamente hay una diferencia pequeña. Tercera, un trabajo como éste no permite que los lectores decidan por ellos mismos si el P -valor es lo suficientemente pequeño para rechazar la hipótesis nula. Si un lector cree que la hipótesis nula no debe rechazarse a menos que $P < 0.01$, entonces informar solamente que $P < 0.05$ no permite que el lector determine si se rechaza o no H_0 .

Notificar el P -valor proporciona mayor información acerca de la fuerza de la evidencia contra la hipótesis nula y permite que cada lector decida por sí mismo si se debe rechazar. Los resultados de software siempre muestran los P -valores que obtienen; éstos se deben incluir siempre que se notifican los resultados de una prueba de hipótesis.

Resumen

Sea α cualquier valor entre 0 y 1. Entonces, si $P \leq \alpha$.

- Se dice que el resultado de la prueba es estadísticamente significativo a un nivel de $100\alpha\%$.
- La hipótesis nula es rechazada a un nivel de $100\alpha\%$.
- Cuando se informa el resultado de una prueba de hipótesis, informe el P -valor, en lugar de sólo compararlo con el 5 o el 1%.

El P -valor no es la probabilidad de que H_0 sea verdadera

Como consecuencia de que P -valor es una probabilidad, y que los P -valores pequeños indican que H_0 es improbable que sea verdadera, es tentador pensar que el P -valor representa la probabilidad de que H_0 sea verdadera. Éste no es enfáticamente el caso. El concepto de probabilidad analizado aquí es útil solamente cuando se aplica a los resultados que pueden resultar en formas diferentes cuando se repiten experimentos. Ésta tiene sentido para definir el P -valor como la probabilidad de observar un valor extremo de un estadístico como $\bar{X}$, ya que el valor de $\bar{X}$ podría ser diferente si el experimento se repitiera. Por otra parte, la hipótesis nula es o no verdadera. La verdad o falsedad de H_0 no se puede cambiar mediante la repetición del experimento. No es, por tanto, correcto hablar de la “probabilidad” de que H_0 sea verdadera.

En este momento se debe mencionar que hay un concepto de probabilidad, diferente del que se analiza en este libro, en el cual se puede calcular una probabilidad de que un enunciado, tal como que una hipótesis nula sea verdadera. Esta clase de probabilidad se llama probabilidad **subjettiva**, y es importante en la teoría de la **estadística bayesiana**. La clase de probabilidad que se analiza en este libro se llama probabilidad **frecuentista**. Una buena referencia para estadística bayesiana es Lee (1997).

Elija H_0 para responder la pregunta correcta

Cuando se realiza una prueba de hipótesis es importante elegir adecuadamente H_0 y H_1 con el propósito de que los resultados de la prueba sean útiles para establecer una conclusión. Los ejemplos 6.4 y 6.5 muestran esto último.

Ejemplo

6.4

Las especificaciones para un tubo de agua piden una media de fuerza de ruptura μ mayor de 2 000 lb por pie de longitud. Los ingenieros realizarán una prueba de hipótesis para determinar si se usa o no cierto tipo de tubo. Seleccionan una muestra aleatoria de secciones del tubo de un pie de longitud para medir sus fuerzas de ruptura y realizar una prueba de hipótesis. El tubo no se utilizará a menos que los ingenieros pueden concluir que $\mu > 2\,000$. Suponga que se prueba $H_0: \mu \leq 2\,000$ contra $H_1: \mu > 2\,000$. ¿Los ingenieros decidirán utilizar el tubo si se rechaza H_0 ? ¿Qué pasa si H_0 no es rechazada?

Solución

Si H_0 es rechazada, los ingenieros concluirán que $\mu > 2\,000$ y utilizarán el tubo. Si H_0 no es rechazada, los ingenieros concluirán que μ podría ser menor o igual a 2 000, y no usarán el tubo.

En el ejemplo 6.4 la acción de los ingenieros con respecto al uso del tubo diferirá dependiendo de si H_0 se rechaza o no. Por tanto, es útil realizar esta prueba, y H_0 y H_1 se han especificado correctamente.

Ejemplo

6.5

En el ejemplo 6.4, suponga que los ingenieros prueban que $H_0: \mu \geq 2\,000$ contra $H_1: \mu < 2\,000$. ¿Los ingenieros decidirán utilizar el tubo si se rechaza H_0 ? ¿Qué pasa si H_0 no es rechazada?

Solución

Si H_0 es rechazada, los ingenieros concluirán que $\mu < 2\,000$ y no utilizarán el tubo. Si H_0 no es rechazada, los ingenieros llegarán a la conclusión de que μ podría ser mayor o igual a 2 000, pero también podría no serlo, por lo que nuevamente no utilizarán el tubo.

En el ejemplo 6.5, la acción con respecto al uso del tubo será la misma, no lo utilizarán, si H_0 es o no rechazada. No tiene sentido realizar esta prueba. Las hipótesis de H_0 y H_1 no se han especificado correctamente.

Nota final: En una prueba de una cola la igualdad siempre va con la hipótesis nula. Por tanto, si μ_0 es el punto que divide H_0 de H_1 , se podría tener $H_0: \mu \leq \mu_0$ o $H_0: \mu \geq \mu_0$, pero nunca $H_0: \mu < \mu_0$ o $H_0: \mu > \mu_0$. La razón para esto último consiste en que cuando se define la distribución nula H_0 se representa con el valor de μ más cercano a H_1 . Sin la igualdad, no hay un valor especificado por H_0 que esté cercano a H_1 . Por tanto, la igualdad debe ir con H_0 .

Significancia estadística no es lo mismo que significancia práctica

Cuando un resultado tiene un P -valor pequeño se dice que es “estadísticamente significativo”. En el uso cotidiano, la palabra *significativo* quiere decir “importante”. Resulta, por tanto, tentador pensar que los resultados estadísticamente significativos siempre deben ser importantes.

Éste no es el caso. A veces los resultados estadísticamente significativos no tienen importancia científica o práctica. Se mostrará esto último con un ejemplo. Suponga que en un proceso con el que se fabrican fibras sintéticas se sabe que éstas tienen una fuerza de ruptura media de 50 N. Se ha desarrollado un nuevo proceso que requeriría instrumentar el rechazo de mucha maquinaria. En una muestra de mil fibras producidas por este nuevo método, el promedio de la fuerza de ruptura era de 50.1 N, y la desviación estándar, de 1 N. ¿Se puede concluir que el nuevo proceso produce fibras con una media mayor de la fuerza de ruptura?

Para responder esta pregunta, sea μ la fuerza de ruptura media de las fibras producidas con el nuevo proceso. Se necesita probar $H_0: \mu \leq 50$ contra $H_1: \mu > 50$. De este modo, si se rechaza H_0 se llega a la conclusión de que el nuevo proceso es mejor. Bajo H_0 , la media muestral $\bar{X}$ tiene una distribución normal con una media de 50 y una desviación estándar de $1/\sqrt{1\,000} = 0.0316$. El puntaje z es

$$z = \frac{50.1 - 50}{0.0316} = 3.16$$

El P -valor es 0.0008. Ésta es una evidencia muy fuerte contra H_0 . El nuevo proceso produce fibras con una fuerza de ruptura media mayor.

¿Qué conclusión práctica se debe extraer de este resultado? Con base en la prueba de hipótesis, se tiene la seguridad de que el nuevo proceso es mejor. ¿Valdría la pena instrumentar el nuevo proceso? Probablemente no. La razón es que la diferencia entre el proceso viejo y el nuevo, aunque estadísticamente muy significativa, es de sólo 0.1 N. Es improbable que esta diferencia sea lo suficientemente grande para ser importante.

La lección que aquí se obtiene es que un resultado puede ser estadísticamente significativo sin ser lo suficientemente grande para tener importancia práctica. ¿Cómo puede ocurrir esto? Una diferencia es estadísticamente significativa cuando es grande comparada con su desviación estándar. En el ejemplo, una diferencia de 0.1 N era estadísticamente significativa, ya que la desviación estándar era solamente de 0.0316 N. Cuando la desviación estándar es muy pequeña, aun una diferencia pequeña puede ser estadísticamente significativa.

El P -valor no mide la significancia práctica. Lo que mide es el grado de confianza que se puede tener de que el valor verdadero es muy diferente del valor especificado por la hipótesis nula. Cuando el P -valor es pequeño, entonces se puede tener confianza de que el valor verdadero es en verdad muy diferente. Esto no necesariamente implica que la diferencia sea lo bastante grande para que tenga importancia práctica.

Relación entre las pruebas de hipótesis y los intervalos de confianza

Tanto los intervalos de confianza como las pruebas de hipótesis están dedicados a determinar los valores creíbles para una cantidad como una media poblacional μ . En una prueba de hipótesis para una media poblacional μ , se especifica un valor particular de μ (la hipótesis nula) y se determina si ese valor es factible. En contraparte, un intervalo de confianza para una media poblacional μ se puede pensar como el conjunto de los valores de μ que satisfacen a un cierto criterio de factibilidad, especificado por el intervalo de confianza con un nivel de $100(1 - \alpha)\%$. Efectivamente, la relación entre los intervalos de confianza y las pruebas de hipótesis es muy cercana.

Para ser más específicos, los valores contenidos dentro de un intervalo de confianza de dos lados a nivel de $100(1 - \alpha)\%$, son precisamente aquellos valores para los cuales el P -valor de una prueba de hipótesis de dos lados será mayor que α . Para ilustrar esto último, considere el ejemplo siguiente (presentado como el ejemplo 5.2 de la sección 5.1). La media muestral de los tiempos de vida de 50 microperforadoras era de $\bar{X} = 12.68$ huecos perforados y la desviación estándar de $s = 6.83$. Haciendo α igual a 0.05 (5%), el intervalo de confianza al 95% para la media poblacional de los tiempos de vida μ se ha calculado de (10.79, 14.57). Suponga que se quiere probar la hipótesis de que μ era igual a uno de los puntos finales del intervalo de confianza. Por ejemplo, considere probar que $H_0: \mu = 10.79$ contra $H_1: \mu \neq 10.79$. Bajo H_0 , el valor observado $\bar{X} = 12.68$ proviene de una distribución normal con media 10.79 y desviación estándar $6.83/\sqrt{50} = 0.9659$. El puntaje z es $(12.68 - 10.79)/0.9659 = 1.96$. Puesto que H_0 especifica que μ es igual a 10.79, ambas colas contribuyen al P -valor, el cual es 0.05, y, por consiguiente, igual a α (véase la figura 6.4).

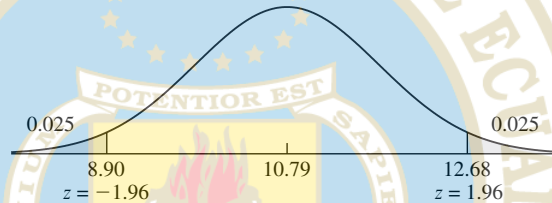


FIGURA 6.4 La media muestral $\bar{X}$ es igual a 12.68. Debido a que 10.79 es un punto final de un intervalo de confianza de 95% basado en $\bar{X} = 12.68$, el P -valor para evaluar $H_0: \mu = 10.79$ es igual a 0.05.

Ahora considere probar la hipótesis $H_0: \mu = 14.57$ contra $H_1: \mu \neq 14.57$, donde 14.57 representa el otro punto final del intervalo de confianza. Esta vez se obtendrá $z = (12.68 - 14.57)/0.9659 = -1.96$, y nuevamente el P -valor es 0.05. Es fácil comprobar que si se elige cualquier valor de μ_0 en el intervalo (10.79, 14.57) y se prueba $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$, el P -valor será mayor que 0.05. Por otra parte, si se elige $\mu_0 < 10.79$ o $\mu_0 > 14.57$, el P -valor será menor que 0.05. Por tanto, el intervalo de confianza de 95% consta precisamente de los valores de μ cuyos P -valores son mayores que 0.05 en una prueba de hipótesis. En este sentido, el intervalo de confianza contiene todos los valores que son creíbles para la media poblacional μ .

Es fácil comprobar que un intervalo de confianza a un nivel de $100(1 - \alpha)\%$ de un sólo lado consta de todos los valores para los cuales el P -valor en una prueba de una cola serían mayores de α . Por ejemplo, con $\bar{X} = 12.68$, $s = 6.83$ y $n = 50$, el límite inferior de confianza de 95% para los tiempos de vida de las perforadoras es 11.09. Si $\mu_0 > 11.09$, entonces el P -valor para probar $H_0: \mu \leq \mu_0$ será mayor de 0.05. De forma semejante, el límite superior de confianza de 95% para los tiempos de vida de las perforadoras es 14.27. Si $\mu_0 < 14.27$, entonces el P -valor para probar $H_0: \mu \geq \mu_0$ será mayor que 0.05.

Ejercicios para la sección 6.2

- ¿Para qué P -valor es la hipótesis nula más factible, $P = 0.5$ o $P = 0.05$?
- Verdadero o falso:
 - Si se rechaza H_0 , entonces se concluye que H_0 es falsa.
 - Si no se rechaza H_0 , entonces se llega a la conclusión de que H_0 es verdadera.
 - Si se rechaza H_0 , entonces se llega a la conclusión de que H_1 es verdadera.
 - Si no se rechaza H_0 , entonces se concluye que H_1 es falsa.
- Si $P = 0.01$, ¿cuál es la mejor conclusión?
 - H_0 es definitivamente falsa.
 - H_0 es definitivamente verdadera.
 - Hay probabilidad de 1% de que H_0 sea verdadera.
 - H_0 podría ser verdadera, pero es improbable.
 - H_0 podría ser falsa, pero es improbable.
 - H_0 es factible.
- Si $P = 0.50$, ¿cuál es la mejor conclusión?
 - H_0 es definitivamente falsa.
- ii) H_0 es definitivamente verdadera.
- iii) Hay una probabilidad de 50% de que H_0 sea verdadera.
- iv) H_0 es factible, y H_1 es falsa.
- v) Tanto H_0 como H_1 son creíbles.
- Verdadero o falso: Si $P = 0.02$, entonces
 - El resultado es estadísticamente significativo a un nivel de 5%.
 - El resultado es estadísticamente significativo a un nivel de 1%.
 - La hipótesis nula es rechazada a un nivel de 5%.
 - La hipótesis nula es rechazada a un nivel de 1%.
- Una hipótesis nula es rechazada a un nivel de 5%. Verdadero o falso:
 - El P -valor es mayor que 5%.
 - El P -valor es menor o igual a 5%.
 - El resultado es estadísticamente significativo a un nivel de 5%.
 - El resultado es estadísticamente significativo a un nivel de 10%.
- El siguiente resultado de MINITAB (mostrado por vez primera en el ejercicio 12 de la sección 6.1) presenta los resultados de una prueba de hipótesis para una media poblacional μ .

One-Sample Z: X

Test of $\mu = 73.5$ vs not = 73.5

The assumed standard deviation = 2.3634

Variable	N	Mean	StDev	SE Mean	95% CI	Z	P
X	145	73.2461	2.3634	0.1963	(72.8614, 73.6308)	-1.29	0.196

- ¿Puede H_0 ser rechazada a un nivel de 5%? ¿Cómo justifica su respuesta?
 - Alguien le pregunta si la hipótesis nula $H_0: \mu = 73$ contra $H_1: \mu \neq 73$ puede ser rechazada a un nivel de 5%. ¿Puede responder sin hacer cálculos? ¿Cómo?
-
- Sea μ el nivel de radiación al que un trabajador de esta área se expone durante el transcurso de un año. La agencia de protección ambiental ha establecido el nivel máximo de seguridad de la exposición en 5 rem por año. Si se realiza una prueba de hipótesis para determinar si un puesto de trabajo es seguro, ¿cuál es la hipótesis nula más adecuada: $H_0: \mu \leq 5$, $H_0: \mu \geq 5$, o $H_0: \mu = 5$? Explique.
 - En cada uno de los siguientes casos establezca la hipótesis nula más adecuada, tomando en consideración la media poblacional μ .
 - Se instalará un nuevo tipo de batería en marcapasos de corazón si se puede demostrar que tienen una media de las duraciones mayor a ocho años.

- b) Se utilizará un nuevo material para fabricar neumáticos si se puede mostrar que la media de la duración de los neumáticos es al menos de sesenta mil millas.
- c) Un inspector de control de calidad recalibra un medidor de flujo si la media de la tasa de flujo difiere de 10 ml/s.
10. Se recomienda la instalación de un dispositivo de disminución de radón en cualquier casa donde la media de la concentración de radón es de 4.0 picocuries por litro (pCi/l) o mayor, porque se ha pensado que a la larga esa exposición representa una dosis suficientemente grande que puede aumentar el riesgo de cáncer. Se hicieron 75 mediciones en una casa específica. La media de las concentraciones era de 3.72 pCi/l y la desviación estándar de 1.93 pCi/l.
- a) El inspector de casas que realizó la prueba informó que debido a que la media de las mediciones es menor de 4.0 es innecesaria la disminución de radón. Explique por qué este razonamiento es incorrecto.
- b) Debido a motivos de salud, se recomienda la disminución de radón siempre que sea factible que la concentración media de radón pueda ser de 4.0 pCi/l o mayor. Establezca las hipótesis nula y alternativa adecuadas para determinar si la disminución de radón es adecuada.
- c) Calcule el P -valor. ¿Usted recomendaría la disminución de radón? Explique.
11. Es deseable comprobar la calibración de una balanza pesando un peso patrón de 10 g 100 veces. Sea μ la media poblacional de lecturas en la balanza, por lo que esta última estará calibrada si $\mu = 10$. Se hace una prueba de hipótesis $H_0: \mu = 10$ contra $H_1: \mu \neq 10$. Considere tres conclusiones posibles: (i) la balanza está calibrada. (ii) la balanza no está calibrada. (iii) la balanza podría estar calibrada.
- a) ¿Cuál de las tres conclusiones es mejor si H_0 es rechazada?
- b) ¿Cuál de las tres conclusiones es mejor si H_0 no es rechazada?
- c) ¿Es posible realizar una prueba de hipótesis en una forma que sea posible demostrar concluyentemente que la balanza está calibrada? Explique.
12. Se supone que una máquina que llena cajas de cereal está calibrada, por lo que la media del peso de llenado es de 12 oz. Sea μ la media verdadera del peso de llenado. Suponga que en una prueba de hipótesis $H_0: \mu = 12$ contra $H_1: \mu \neq 12$, el P -valor es 0.30.
- a) ¿Se debe rechazar H_0 con base en esta prueba? Explique.
- b) ¿Puede concluir que la máquina está calibrada para proporcionar una media del peso de llenado de 12 oz? Explique.
13. Un método de aplicación de zinc para blindar acero se supone que produce una capa cuyo espesor medio no es mayor de 7 micras. Una inspectora de calidad mide el espesor de 36 unidades de las capas y prueba que $H_0: \mu \leq 7$ contra $H_1: \mu > 7$. Obtiene un P -valor de 0.40. Dado que $P > 0.05$, concluye que la media del espesor está dentro de la especificación. ¿Esta conclusión es correcta? Explique.
14. Complete la respuesta: un intervalo de confianza al 95% para μ es (1.2, 2.0). Con base en los datos a partir de los cuales se construyó el intervalo de confianza, alguien quiere probar que $H_0: \mu = 1.4$ contra $H_1: \mu \neq 1.4$. El P -valor será _____.
- i) Mayor que 0.05
- ii) Menor que 0.05
- iii) Igual a 0.05
15. Con referencia al ejercicio 14, ¿para qué hipótesis nula será $P = 0.05$?
- i) $H_0: \mu = 1.2$
- ii) $H_0: \mu \leq 1.2$
- iii) $H_0: \mu \geq 1.2$
16. Una científica calcula un intervalo de confianza del 90% de (4.38, 6.02). Con los mismos datos también calcula que un intervalo de confianza del 95% de (4.22, 6.18), y un intervalo de confianza del 99% de (3.91, 6.49). Ahora quiere probar $H_0: \mu = 4$ contra $H_1: \mu \neq 4$. Considerando el P -valor, ¿cuál de los siguientes enunciados es verdadero?
- i) $P > 0.10$
- ii) $0.05 < P < 0.10$
- iii) $0.01 < P < 0.05$
- iv) $P < 0.01$
17. La fuerza de cierto tipo de caucho se prueba sujetando pedazos de éste a una prueba de desgaste. Para que el caucho sea aceptable, la media del peso μ debe ser menor de 3.5 mg. Muchos pedazos de caucho fueron curados en cierta manera mientras estaban sujetos a la prueba de desgaste. Un límite superior de confianza del 95% para la media del peso perdido que se calculó a partir de estos datos era de 3.45 mg. Alguien sugiere utilizar estos datos para probar $H_0: \mu \geq 3.5$ contra $H_1: \mu < 3.5$.
- a) ¿Es posible determinar a partir del límite de confianza si $P < 0.05$? Explique.

b) ¿Es posible determinar a partir del límite de confianza si $P < 0.01$? Explique.

b) ¿Es posible determinar a partir del límite de confianza si $P < 0.05$? Explique.

18. Una remesa de fibras no es aceptable si la fuerza de ruptura media de las fibras es menor de 50 N. Se probó una gran muestra de fibras de esta remesa, y se calculó un límite inferior de confianza de 98% para la fuerza de ruptura media de 50.1 N. Alguien sugiere utilizar estos datos para probar la hipótesis $H_0: \mu \leq 50$ contra $H_1: \mu > 50$.

a) ¿Es posible determinar a partir del límite de confianza si $P < 0.01$? Explique.

19. Con referencia al ejercicio 17, se descubre que la media muestral utilizada para calcular el límite de confianza es $\bar{X} = 3.40$. ¿Es posible determinarla si $P < 0.01$? Explique.

20. Con referencia al ejercicio 18, se descubre que la desviación estándar muestral utilizada para calcular el intervalo de confianza es 5 N. ¿Es posible determinarla si $P < 0.01$? Explique.

6.3 Pruebas de hipótesis para la proporción poblacional

Una proporción poblacional es simplemente una media poblacional para una población de 0 y 1: una población de Bernoulli. Por esta razón, las pruebas de hipótesis para proporciones son similares a las analizadas en la sección 6.1 para medias poblacionales. He aquí un ejemplo.

Un proveedor de placas semiconductoras afirma que de todas las placas que suministra, sólo 10% son defectuosas. Se prueba una muestra de 400 placas, y 50 de ellas, 12.5%, presentan defectos. ¿Se puede concluir que esta afirmación es falsa?

Para realizar la prueba de hipótesis se procede en forma muy parecida a las que se observaron en la sección 6.1. Lo que hace diferente este problema es que la muestra consiste de éxitos y fracasos, con “éxito” se indica una placa defectuosa. Si la proporción poblacional de placas defectuosas se denota mediante p , entonces la afirmación del proveedor es que $p \leq 0.1$. Dado que esta hipótesis se ocupa de una proporción poblacional, es natural basar la prueba en la proporción muestral $\hat{p}$. Con base en la suposición razonable de que las placas se muestrean de manera independiente, y al usar el teorema del límite central, puesto que el tamaño muestral es grande, se obtiene que

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \quad (6.1)$$

donde n es el tamaño de muestra, igual a 400.

Se debe definir la hipótesis nula. La pregunta es si los datos permiten concluir que la afirmación del proveedor es falsa. Por tanto, la afirmación del proveedor, que es $p \leq 0.1$, debe ser H_0 . Por otra parte, sería imposible probar la falsedad de la afirmación, sin que importara lo que los datos indicaban.

Las hipótesis nula y alternativa son

$$H_0: p \leq 0.1 \quad \text{contra} \quad H_1: p > 0.1$$

Para realizar la prueba de hipótesis se supone que H_0 es verdadera y se toma $p = 0.1$. Al sustituir $p = 0.1$ y $n = 400$ en la expresión (6.1) se obtiene la distribución nula de $\hat{p}$:

$$\hat{p} \sim N(0.1, 2.25 \times 10^{-4})$$

La desviación estándar de $\hat{p}$ es $\sigma_{\hat{p}} = \sqrt{2.25 \times 10^{-4}} = 0.015$. El valor observado de $\hat{p}$ es $50/400 = 0.125$. El puntaje z de $\hat{p}$ es

$$z = \frac{0.125 - 0.100}{0.015} = 1.67$$

La tabla z indica que la probabilidad de que una variable aleatoria normal estándar tenga un valor mayor de 1.67 es 0.0475. El P -valor es, por tanto, 0.0475 (véase la figura 6.5).

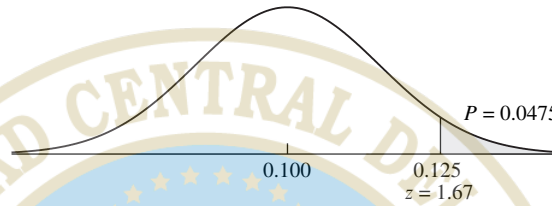


FIGURA 6.5 La distribución nula de $\hat{p}$ es $N(0.1, 0.015^2)$. Por tanto, si H_0 es verdadera, la probabilidad de que $\hat{p}$ tome un valor tan extremo o más extremo que el valor observado de 0.125 es 0.0475. Éste es el P -valor.

¿Qué se concluye acerca de H_0 ? Ya sea que la afirmación del proveedor sea falsa, o que se haya observado una muestra que sea tan extrema como las demás, que se podrían haber extraído 4.75% de las muestras. Tal muestra sería anormal, pero no muy improbable. Hay una razón para estar particularmente escéptico con respecto a la afirmación, pero tal vez no se debe aún condenar al proveedor. Si es posible, sería una buena idea muestrear más placas.

Observe que bajo la comúnmente utilizada regla práctica se rechazaría H_0 y se condenaría al proveedor, debido a que P es menor que 0.05. Este ejemplo ilustra la debilidad de esta regla. Si usted hace los cálculos, encontrará que si sólo 49 de las placas de la muestra hubieran estado defectuosas en vez de 50, el P -valor se elevaría a 0.0668, y el proveedor estaría salvado. Por tanto, el destino del proveedor depende del resultado de una sola placa de las 400. No tiene sentido marcar tal línea nítida. Es mejor sólo notificar el P -valor y esperar a tener más evidencias antes de obtener una conclusión final.

El tamaño muestral debe ser grande

La prueba descrita antes requiere que la proporción muestral esté normalmente distribuida. Esta suposición estará justificada siempre que $np_0 > 10$ y $n(1 - p_0) > 10$, donde p_0 es la proporción poblacional que se especificó en la distribución nula. Entonces el puntaje z se puede utilizar como el estadístico de prueba, haciendo a ésta una prueba z .

Ejemplo

6.6

El artículo “Refinement of Gravimetric Geoid Using GPS and Leveling Data” (W. Thurston, en *Journal of Surveying Engineering*, 2000:27-56) presenta un método para medir las alturas ortométricas arriba del nivel del mar. Para una muestra de 1 225 puntos de partida, 926 dieron

resultados que están dentro del espíritu de la clase C nivelando los límites de tolerancia. ¿Se puede llegar a la conclusión de que este método produce resultados dentro de los límites de tolerancia más de 75% de las veces?

Solución

Sea p la probabilidad de que el método produzca un resultado dentro de los límites de tolerancia. Las hipótesis nula y alternativa son:

$$H_0: p \leq 0.75 \text{ contra } H_1: p > 0.75$$

La proporción muestral es $\hat{p} = 926/1\,225 = 0.7559$. Bajo la hipótesis nula, $\hat{p}$ está distribuido normalmente con media de 0.75 y desviación estándar $\sqrt{(0.75)(1 - 0.75)/1\,225} = 0.0124$. El puntaje z es

$$z = \frac{0.7559 - 0.7500}{0.0124} = 0.48$$

El P -valor es 0.3156 (véase la figura 6.6). No se puede concluir que el método produzca buenos resultados más de 75% de las veces.

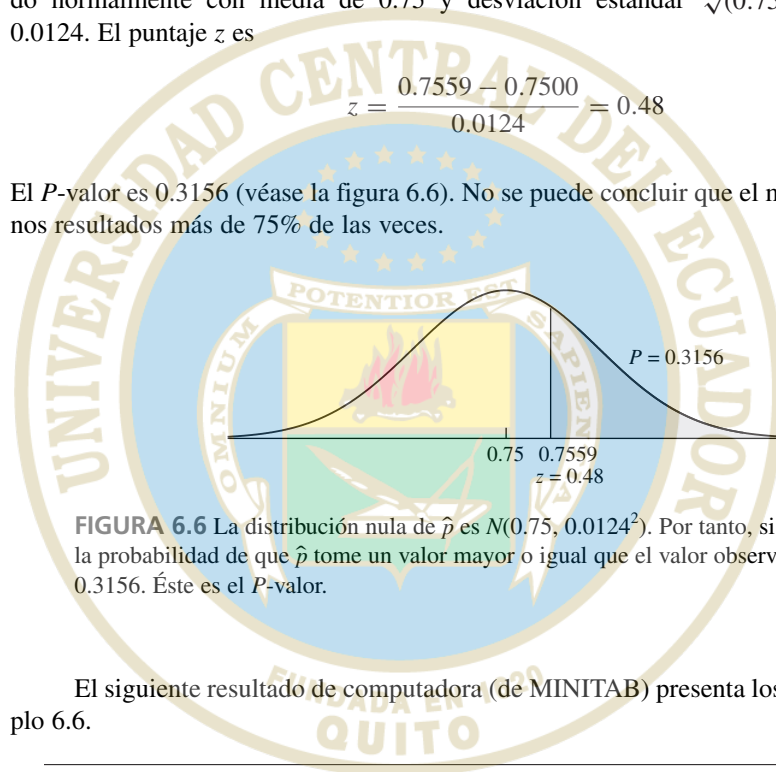


FIGURA 6.6 La distribución nula de $\hat{p}$ es $N(0.75, 0.0124^2)$. Por tanto, si H_0 es verdadera, la probabilidad de que $\hat{p}$ tome un valor mayor o igual que el valor observado de 0.7559 es 0.3156. Éste es el P -valor.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.6.

Test and CI for One Proportion: GPS

Test of $p = 0.75$ vs $p > 0.75$

Variable	X	N	Sample p	95%	Z-Value	P-Value
				Lower Bound		
GPS	926	1225	0.755918	0.735732	0.48	0.316

El resultado contiene un límite inferior de confianza de 95%, así como el P -valor. Observe que este límite inferior se calculó mediante el método tradicional (la expresión 5.8 de la sección 5.2 presenta la versión de dos lados de este método).

La relación con los intervalos de confianza para una proporción

Un intervalo de confianza a un nivel de $100(1 - \alpha)\%$ contiene los valores correspondientes a un parámetro para el cual el P -valor de una prueba de hipótesis será mayor que α . Respecto a los intervalos de confianza para la proporción presentados en la sección 5.2 y la prueba de hipótesis presentada aquí, este enunciado es aproximadamente verdadero. La razón de esto es que los métodos presentados en la sección 5.2 son versiones ligeramente modificadas (y mucho más fáciles de calcular) de métodos de intervalos de confianza más complicados para los cuales el enunciado es verdadero.

Resumen

Sea X el número de éxitos en n ensayos independientes de Bernoulli, cada uno con probabilidad de éxito p ; en otras palabras, sea $X \sim \text{Bin}(n, p)$.

Para probar una hipótesis nula de la forma $H_0: p \leq p_0$, $H_0: p \geq p_0$, o $H_0: p = p_0$, suponiendo que tanto np_0 como $n(1 - p_0)$ son mayores que 10:

- Calcule el puntaje z : $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$.
- Calcule el P -valor. Este último constituye un área bajo la curva normal, que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: p > p_0$$

$$H_1: p < p_0$$

$$H_1: p \neq p_0$$

P -valor

Área a la derecha de z

Área a la izquierda de z

Suma de áreas en las colas correspondientes a z y $-z$

Ejercicios para la sección 6.3

- Una muestra aleatoria de 300 componentes electrónicos fabricados mediante un proceso específico se muestrea y se encuentra que 25 están defectuosos. Sea p la proporción de componentes fabricados mediante este proceso que presentan defectos. El ingeniero responsable de la producción afirma que $p \leq 0.05$. ¿La muestra proporciona suficientes evidencias para rechazar la afirmación?
- Una muestra aleatoria de 100 pernos de la producción de un día se muestrea y se encuentra que dos de ellos tienen diámetros por debajo de la especificación. Se afirma que la proporción de pernos defectuosos entre los fabricados en ese día era menor que 0.05. ¿Es adecuado utilizar los métodos de esta sección para determinar si se puede rechazar esta afirmación? Si es así, establezca las hipótesis nula y alternativa adecuadas y calcule el P -valor. Si no, explique por qué.
- Una compañía de telecomunicaciones proporcionó a sus suscriptores de televisión por cable acceso libre a un nuevo canal de deportes durante un mes. Se eligió una muestra de 400 teleespectadores y se les preguntó si estarían dispuestos a pagar 10 dólares extra por mes para continuar teniendo acceso al canal. Solamente 25 respondieron que estarían dispuestos a pagar.
¿La compañía puede concluir que más de 5% de sus suscriptores pagarían el canal?
- Las incineradoras pueden ser una fuente de emisiones peligrosas en la atmósfera. Se recolectaron muchas muestras de gases de una muestra de 50 incineradoras en una ciudad. De las 50, sólo 18 satisfacían una norma ambiental para la concentración de un compuesto peligroso. ¿Se puede concluir que menos de la mitad de las incineradoras en la ciudad satisfacen la norma?

5. Las piezas de grava se clasifican como pequeñas, medianas o grandes. Una distribuidora afirma que al menos 10% de las piezas de grava de su planta son grandes. En una muestra aleatoria de 1 600 piezas, 150 se clasificaron como grandes. ¿Representa esto suficiente evidencia para rechazar la afirmación?
6. Una máquina pulidora será calificada para una tarea especial si puede demostrar que produce menos de 8% de partes defectuosas. En una muestra aleatoria de 300 partes, 12 estaban defectuosas. ¿Con base en estos datos la máquina puede ser calificada?
7. Un fabricante de estaciones de trabajo de computadora está probando un nuevo proceso de ensamble automatizado. El proceso actual tiene una tasa de defectos de 5%. En una muestra de 400 estaciones de trabajo ensambladas con el nuevo proceso, 15 tenían defectos. ¿Se puede concluir que el nuevo proceso tiene una tasa menor de defectos?
8. Con referencia al ejercicio 1 de la sección 5.2. ¿Se puede concluir que más de 60% de las mediciones hechas con el instrumento serán satisfactorias?
9. Con referencia al ejercicio 2 de la sección 5.2, ¿se puede concluir que menos de 40% de los fusibles fabricados ese día tenían amperajes de fundición mayores de 15 A?

10. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para una proporción poblacional p .

Test and CI for One Proportion: X

Test of $p = 0.4$ vs $p < 0.4$

Variable	X	N	Sample p	95% Upper Bound	Z-Value	P-Value
X	73	240	0.304167	0.353013	-3.03	0.001

- a) ¿Esta es una prueba de una o de dos colas?
 - b) ¿Cuál es la hipótesis nula?
 - c) ¿Puede H_0 ser rechazada a un nivel de 2%? ¿Cómo justifica su respuesta?
 - d) Alguien le pregunta si la hipótesis nula $H_0: p \geq 0.45$ contra $H_1: p < 0.45$ se puede rechazar a un nivel de 2%. ¿Puede responder sin hacer ningún cálculo? ¿Cómo?
 - e) Utilice el resultado y una tabla adecuada para calcular el P -valor para la prueba de $H_0: p \leq 0.25$ contra $H_1: p > 0.25$.
 - f) Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 90% para p .
11. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para una proporción poblacional p . Falta algunos números. Complete éstos del (a) al (c).

Test and CI for One Proportion: X

Test of $p = 0.7$ vs $p < 0.7$

Variable	X	N	Sample p	95% Upper Bound	Z-Value	P-Value
X	345	500	(a)	0.724021	(b)	(c)

6.4 Pruebas de hipótesis para la media poblacional con muestras pequeñas

En la sección 6.1 se describió un método con el cual se prueba una hipótesis para la media poblacional basada en una muestra grande. Un paso clave en el método es aproximar a la desviación estándar σ con la desviación estándar muestral s . Después se utiliza la curva normal para encontrar el P -valor. Cuando el tamaño muestral es pequeño, s puede no estar cerca de σ , ello invalida este método de muestras grandes. Sin embargo, cuando la población es aproximadamente normal, se puede utilizar la distribución t de Student. Esto se ilustra con un ejemplo.

Los separadores de anillos para una transmisión de eje tienen una especificación de espesor de 38.98-39.02 mm. El proceso con el que se fabrican los anillos se supone que está calibrado para que la media de los espesores sea de 39 mm en el centro de la ventana de especificación. Se extrae una muestra de seis anillos y se mide su espesor. Las seis capas son 39.030, 38.997, 39.012, 39.008, 39.019 y 39.002. Suponga que la población de los espesores de los anillos es aproximadamente normal. ¿Se puede concluir que el proceso necesita recalibración?

Se denota la media poblacional con μ , las hipótesis nula y alternativa son

$$H_0: \mu = 39.00 \text{ contra } H_1: \mu \neq 39.00$$

Observe que H_0 especifica un solo valor para μ , puesto que la calibración requiere que la media sea igual al valor correcto. Para construir el estadístico de prueba, observe que debido a que se supone que la población tiene una distribución normal, la cantidad

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

tiene una distribución t de Student con $n - 1 = 5$ grados de libertad. Éste es el estadístico de prueba.

En este ejemplo los valores observados de la media muestral y la desviación estándar son $\bar{X} = 39.01133$ y $s = 0.011928$. El tamaño muestral es $n = 6$. La hipótesis nula especifica que $\mu = 39$. Por tanto, el valor del estadístico de prueba es

$$t = \frac{39.01133 - 39.00}{0.011928/\sqrt{6}} = 2.327$$

El P -valor es la probabilidad de observar un valor del estadístico de prueba cuyo desacuerdo con H_0 es igual o mayor que el realmente observado. Puesto que H_0 especifica que $\mu = 39.00$, ésta es una prueba de dos colas, por lo que ambos valores arriba y abajo de 39.00 no concuerdan con H_0 . Por tanto, el P -valor es la suma de las áreas bajo la curva correspondiente a $t > 2.327$ y a $t < -2.327$.

La figura 6.7 ilustra la distribución nula e indica la ubicación del estadístico de prueba. De la tabla t (tabla A.3 en el apéndice A) el renglón que corresponde a 5 grados de libertad indica que el valor $t = \pm 2.015$ tiene un área de 0.05 en cada cola, para un total de 0.10, y que el valor $t = \pm 2.571$ corta un área de 0.025 en cada cola, para un total de 0.05. Por tanto, el P -valor está entre 0.05 y 0.10. Aunque no se puede establecer en forma concluyente que el proceso no está calibrado, no parece demasiado bueno. Sería prudente recalibrarlo.

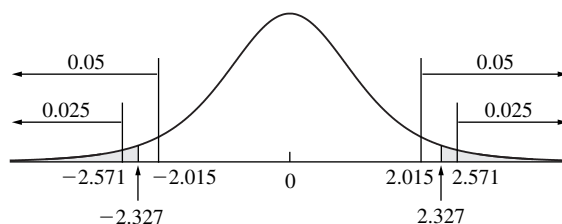


FIGURA 6.7 La distribución nula de $t = (\bar{X} - 39.00)/(s/\sqrt{6})$ es la t de Student con cinco grados de libertad. El valor observado de t , correspondiente a los valores observados $\bar{X} = 39.01133$ y $s = 0.011928$, es 2.327. Si H_0 es verdadera, la probabilidad de que t tome un valor mayor o igual que el observado está entre 0.05 y 0.10. Debido a que H_0 especificó que μ era igual a un valor específico, ambas colas de la curva contribuyen al P -valor.

En este ejemplo, el estadístico de prueba era una estadística t en vez de un puntaje z . Por esta razón, esta prueba se denomina prueba t .

Ejemplo

6.7

Antes de que una sustancia se pueda considerar segura para enterrarse como residuo se deben caracterizar sus propiedades químicas. El artículo “Landfilling Ash/Sludge Mixtures” (J. Benoit, T. Eighmy y B. Crannell, en *Journal of Geotechnical and Geoenvironmental Engineering* 1999:877-888) notifica que en una de seis muestras de lodo de una planta de tratamiento de agua residual de Nueva Hampshire, la media del pH era 6.68 con desviación estándar de 0.20. ¿Se puede concluir que la media del pH es menor de 7.0?

Solución

Sea μ la media del pH para este tipo de lodo. Las hipótesis nula y alternativa son

$$H_0: \mu \geq 7.0 \quad \text{contra} \quad H_1: \mu \leq 7.0$$

Bajo H_0 , el estadístico de prueba

$$t = \frac{\bar{X} - 7.0}{s/\sqrt{n}}$$

tiene una distribución t de Student con cinco grados de libertad. Al sustituir $\bar{X} = 6.68$, $s = 0.20$, y $n = 6$, el valor del estadístico de prueba es

$$t = \frac{6.68 - 7.00}{0.20/\sqrt{6}} = -3.919$$

Al observar la tabla t se tiene que al valor $t = -3.365$ le corresponde un área de 0.01 en la cola de la izquierda, y al valor $t = -4.033$ le corresponde un área de 0.005 (véase la figura 6.8). Se concluye que el P -valor se encuentra entre 0.005 y 0.01. Existe evidencia de que la media del pH es menor que 7.0.

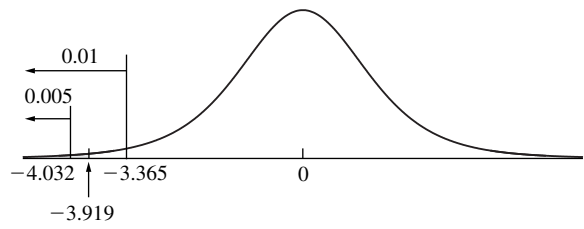


FIGURA 6.8 Solución al ejemplo 6.7. La distribución nula es una *t* de Student con cinco grados de libertad. El valor observado de *t* es -3.919 . Si H_0 es verdadera, la probabilidad de que *t* tenga un valor mayor o igual que el observado se encuentra entre 0.005 y 0.01.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.7.

One-Sample T: pH							
Test of mu = 7 vs < 7							
					95%		
					Upper		
Variable	N	Mean	StDev	SE Mean	Bound	T	P
pH	6	6.680	0.200	0.081665	6.84453	-3.92	0.006

Observe que el límite superior de confianza de 95% dado en el resultado es consistente con la hipótesis alternativa. Esto último indica que el *P*-valor es menor a 5%.

Utilice *z*, no *t*, si se conoce σ

En ocasiones se puede tomar una muestra pequeña de una población normal cuya desviación estándar σ es conocida. En estos casos no se utiliza la curva *t* de Student, porque no se está aproximando a σ con *s*. Sino que se utiliza la tabla *z* y se realiza una prueba *z*. El ejemplo 6.8 muestra el método.

Ejemplo
6.8

Al principio de esta sección se describió una muestra de seis separadores de anillos, cuyos espesores (en mm) fueron 39.030, 38.997, 39.012, 39.008, 39.019 y 39.002. Se denota la media poblacional de los espesores mediante μ y se prueban las hipótesis

$$H_0: \mu = 39.00 \quad \text{contra} \quad H_1: \mu \neq 39.00$$

Ahora suponga que estos seis separadores de anillos se fabricaron justo después de que se cambió de lugar la máquina que los produjo. Suponga que, con base en gran número de anillos fabricados antes del cambio, se tenía conocimiento de que la población de espesores de los anillos se aproximaba mucho a la normal, con desviación estándar $\sigma = 0.010$ mm, y es razonable suponer que el cambio no afectó esto. Con base en los datos, ¿es posible rechazar H_0 ?

Se calcula $\bar{X} = 39.01133$. No se necesita el valor de s , ya que se sabe que $\sigma = 0.010$. Dado que la población es normal, $\bar{X}$ también lo es, aunque el tamaño de muestra sea pequeño. Por tanto, la distribución nula es

$$\bar{X} \sim N(39.00, 0.010^2)$$

El puntaje z es

$$z = \frac{39.01133 - 39.000}{0.010/\sqrt{6}} = 2.78$$

El P -valor es 0.0054, por lo que H_0 se puede rechazar tranquilamente.

Resumen

Sea $X_1, \dots, X_n$ una muestra de una población *normal* con media μ y desviación estándar σ , donde σ es desconocida.

Para probar una hipótesis nula de la forma $H_0: \mu \leq \mu_0$, $H_0: \mu \geq \mu_0$, o $H_0: \mu = \mu_0$:

- Calcule el estadístico de prueba $t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$.
- Calcule el P -valor. Éste es un área bajo la curva t de Student con $n - 1$ grados de libertad, que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: \mu > \mu_0$$

$$H_1: \mu < \mu_0$$

$$H_1: \mu \neq \mu_0$$

P -valor

Área a la derecha de t

Área a la izquierda de t

Suma de áreas en las colas correspondientes a t y $-t$

- Si se conoce σ , el estadístico de prueba es $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$, y se debe hacer una prueba z .

Ejercicios para la sección 6.4

- Cada uno de los siguientes conjuntos de datos hipotéticos representan algunos de los pesos repetidos de un patrón estándar de peso del que se sabe que tiene una masa de 100 g. Suponga que las lecturas son una muestra aleatoria de una población que sigue una curva normal. Si es posible, realice una prueba t para ver si la balanza está calibrada adecuadamente. Si no es posible, explique por qué.
 - 100.02, 99.98, 100.03
 - 100.01
- Un geólogo está haciendo mediciones repetidas (en gramos) de la masa de una roca. No se conoce si las mediciones son una muestra aleatoria de una población aproximadamente normal. A continuación se muestran tres conjuntos de repeticiones de las mediciones, listadas en el orden en que se realizaron. Para cada conjunto de mediciones establezca si las suposiciones necesarias para la validez de la prueba t parecen satisfacerse. Si las suposiciones no se satisfacen, explique por qué.
 - | | | | | |
|--------|--------|--------|--------|--------|
| 213.03 | 212.95 | 213.04 | 213.00 | 212.99 |
| 213.01 | 221.03 | 213.05 | | |
 - | | | | | |
|--------|--------|--------|--------|--------|
| 213.05 | 213.00 | 212.94 | 213.09 | 212.98 |
| 213.02 | 213.06 | 212.99 | | |

c) 212.92 212.95 212.97 213.00 213.01
213.04 213.05 213.06

3. Se está estudiando un nuevo proceso para sintetizar el metanol a partir del metano con el fin de evaluar su viabilidad técnica. Las simulaciones de diseño indican que el reactor debe tener una conversión de metano media mayor que el 35% para que el proceso sea viable. En un estudio inicial se realizaron seis operaciones. La conversión promedio fue de 39% y la desviación estándar de 4%. Si se puede concluir que la media de la conversión μ es mayor de 35%, debe realizarse una evaluación adicional del proceso.

- a) Establezca las hipótesis nula y alternativa adecuadas.
b) Determine el P -valor.
c) ¿Debe realizarse una evaluación adicional del proceso? Explique.

4. Se supone que un tipo especial de gasolina tiene una media de octanaje mayor de 90%. Se realizaron cinco mediciones de la tasa de octanaje de la siguiente manera:

90.1 88.8 89.5 91.0 92.1

¿Se puede concluir que la media del octanaje es mayor a 90%?

5. Se requiere que el espesor de la pared de las botellas de dos litros de policarbonato sea en promedio de 4.0 milésimas de pulg. Un ingeniero responsable del control de calidad extrae una muestra de siete botellas de policarbonato de dos litros de una población grande y mide (en milésimas de pulg) los espesores de las paredes de cada una. Se obtienen los siguientes resultados.

4.065 3.967 4.028 4.008 4.195 4.057 4.010

¿Se puede concluir que la media de los espesores de las paredes es diferente a 4.0 milésimas de pulg?

6. Como parte del programa de control de calidad para la línea de producción de un catalizador, se prueba la pureza de las materias primas (alúmina y un espesante). El proceso requiere que la pureza de la alúmina sea mayor a 85%. Una muestra aleatoria de un envío reciente de alúmina produjo estos resultados (en %):

93.2 87.0 92.1 90.1 87.3 93.6

Se realizará una prueba de hipótesis para determinar si se acepta o no el embarque.

- a) Establezca las hipótesis nula y alternativa adecuadas.
b) Calcule el P -valor.
c) ¿Debe aceptarse el embarque? Explique.

7. Una muestra de 18 piezas de material laminado tenía una media de deformación de 1.88 mm y una desviación estándar de 0.21 mm. ¿Se puede concluir que la media de la deformación de este tipo de laminado es menor a 2 mm?

8. Con referencia al ejercicio 12 de la sección 5.3, ¿se puede concluir que la media de la cantidad de tolueno eliminado en el enjuague es menor a 8%?

9. Con referencia al ejercicio 13 en la sección 5.3, ¿se puede concluir que la media de la cantidad de uniconazol absorbida es menor a 2.5 μg ?

10. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para una media poblacional μ .

One-Sample T: X

Test of $\mu = 5.5$ vs > 5.5

					95%		
					Lower		
Variable	N	Mean	StDev	SE Mean	Bound	T	P
X	5	5.92563	0.15755	0.07046	5.77542	6.04	0.002

- a) ¿Es ésta una prueba de una o de dos colas?
b) ¿Cuál es la hipótesis nula?
c) ¿Puede rechazarse H_0 a un nivel de 1%? ¿Cómo justifica su respuesta?
d) Utilice el resultado y una tabla adecuada para calcular el P -valor para la prueba de $H_0: \mu \geq 6.5$ contra $H_1: \mu < 6.5$.
e) Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 99% para μ .

11. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para una media poblacional μ . Faltan algunos de los números. Complételos.

One-Sample T: X

Test of $\mu = 16$ vs not = 16

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
X	11	13.2874	(a)	1.8389	((b), (c))	(d)	0.171

6.5 Pruebas de hipótesis para la diferencia entre dos medias con muestras grandes

Ahora se analizarán ejemplos en los que se desea determinar si las medias de dos poblaciones son iguales. Los datos los conformarán dos muestras, una para cada población. La idea básica es simple. Se calculará la diferencia de las medias muestrales. Si la diferencia se encuentra alejada de 0, se concluirá que las medias poblacionales son diferentes. Si la diferencia se aproxima a 0, se concluirá que las medias poblacionales podrían ser iguales.

Como ejemplo, suponga que a la directora de producción de una fábrica de maquinaria industrial le interesa que los cojinetes de bolas producidos a temperaturas ambientales bajas tengan diámetros menores que los producidos a temperaturas mayores. Para analizar esta cuestión, extrae una muestra de 120 cojinetes que se fabricaron durante la mañana, antes de que en toda la fábrica se incrementara la temperatura y encontró que la media de su diámetro es de 5.068 mm, y que su desviación estándar, 0.011 mm. En forma independiente prueba 65 cojinetes fabricados durante la tarde; halló que la media de su diámetro es de 5.072 mm, y su desviación estándar, 0.007 mm. ¿Puede la directora concluir que los cojinetes fabricados en la mañana tienen, en promedio, diámetros menores que los fabricados durante la tarde?

Se empieza transfiriendo el problema al lenguaje estadístico. Se tiene una muestra aleatoria simple $X_1, \dots, X_{120}$ de diámetros de cojinetes de bolas fabricados por la mañana, y otra muestra aleatoria simple $Y_1, \dots, Y_{65}$ de diámetros de cojinetes elaborados por la tarde. Se denota la media poblacional de los diámetros de los primeros cojinetes fabricados por la mañana con μ_X , y la media poblacional de los diámetros de los segundos elaborados en la tarde con μ_Y . Las desviaciones estándar correspondientes se denotan con σ_X y σ_Y . Estas medias y las desviaciones estándar poblacionales no se conocen. Los tamaños muestrales son $n_X = 120$ y $n_Y = 65$. Se tiene interés en la diferencia $\mu_X - \mu_Y$.

Ahora se deben determinar las hipótesis nula y alternativa. La cuestión es si se puede concluir que la media poblacional de los cojinetes matutinos es menor que la de los vespertinos. Por consiguiente, las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y \geq 0 \quad \text{contra} \quad H_1: \mu_X - \mu_Y < 0$$

La prueba está basada en $\bar{X} - \bar{Y}$. Debido a que ambos tamaños de muestra son grandes, $\bar{X}$ y $\bar{Y}$ se distribuyen aproximadamente como una normal. En virtud de que las muestras son independientes, se tiene que la distribución nula de $\bar{X} - \bar{Y}$ es

$$\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2) = N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right) \quad (6.2)$$

Los valores observados son $\bar{X} = 5.068$ y $\bar{Y} = 5.072$ para las medias muestrales, y $s_X = 0.011$ y $s_Y = 0.007$ para las desviaciones estándar muestrales. Bajo H_0 , $\mu_X - \mu_Y = 0$ (el valor más cercano a H_1). Se aproximan las varianzas poblacionales σ_X^2 y σ_Y^2 con las varianzas muestrales $s_X^2 = 0.011^2$ y $s_Y^2 = 0.007^2$, respectivamente, y se sustituye $n_X = 120$ y $n_Y = 65$, para calcular la desviación estándar de la distribución nula, obteniendo $\sqrt{0.011^2/120 + 0.007^2/65} = 0.001327$. Por tanto, la distribución nula de $\bar{X} - \bar{Y}$ es

$$\bar{X} - \bar{Y} \sim N(0, 0.001327^2)$$

El valor observado de $\bar{X} - \bar{Y}$ es $5.068 - 5.072 = -0.004$. El puntaje z es

$$z = \frac{-0.004 - 0}{0.001327} = -3.01$$

La figura 6.9 muestra la distribución nula y la ubicación del estadístico de prueba. El P -valor es 0.0013. La sospecha de la directora es correcta. Los cojinetes fabricados en la mañana tienen menor media de los diámetros.

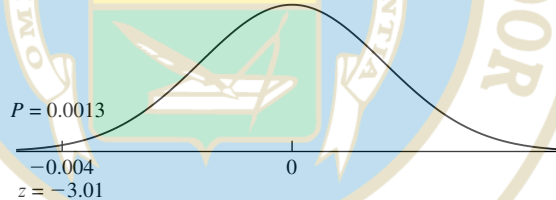


FIGURA 6.9 La distribución nula de $\bar{X} - \bar{Y}$ es $N(0, 0.001327^2)$. Por lo que si H_0 es verdadera, la probabilidad de que $\bar{X} - \bar{Y}$ tenga un valor mayor o igual que los valores observados de -0.004 es de 0.0013. Éste es el P -valor.

Observe que se utilizó la suposición de que las muestras eran independientes cuando se calculó la varianza de $\bar{X} - \bar{Y}$. Ésta es una condición que a menudo se obtiene fácilmente en la práctica. A menos que exista una conexión bastante obvia entre los elementos de las dos muestras, por lo general se supone con certeza que son independientes.

Ejemplo

6.9

El artículo “Effect of Welding Procedure on Flux Cored Steel Wire Deposits” (N. Ramini de Rissone, I. de S. Bott y cols., en *Science and Technology of Welding and Joining*, 2003:113-122) compara las propiedades de soldaduras hechas con dióxido de carbono como gas de protección con respecto a las de soldaduras hechas mediante una mezcla de argón y dióxido de

carbono. Una propiedad estudiada era el diámetro de inclusiones, que son partículas incrustadas en la soldadura. Una muestra de 544 inclusiones en soldaduras hechas al usar argón como protección tiene un diámetro promedio de $0.37 \mu\text{m}$, con desviación estándar de $0.25 \mu\text{m}$. Una muestra de 581 inclusiones en soldaduras hechas al emplear dióxido de carbono como protección tiene diámetro promedio de $0.40 \mu\text{m}$, con desviación estándar de $0.26 \mu\text{m}$. (Las desviaciones estándar se calcularon con una gráfica.) ¿Se puede concluir que las medias de los diámetros de las inclusiones son diferentes entre los dos gases de protección?

Solución

Sea $\bar{X} = 0.37$ la media del diámetro de la muestra para soldaduras de argón. Entonces $s_X = 0.25$ y el tamaño de muestra es $n_X = 544$. Sea $\bar{Y} = 0.40$ la media del diámetro de la muestra de soldaduras de dióxido de carbono. Entonces $s_Y = 0.26$ y el tamaño muestral es $n_Y = 581$. Sea μ_X la media poblacional del diámetro para las soldaduras de argón, y μ_Y la media poblacional del diámetro para soldaduras de dióxido de carbono. Las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y = 0 \quad \text{contra} \quad H_1: \mu_X - \mu_Y \neq 0$$

Se ha observado que $\bar{X} - \bar{Y} = 0.37 - 0.40 = -0.03$. Este valor se extrajo de una población normal con media $\mu_X - \mu_Y$, y varianza aproximada por $s_X^2/n_X + s_Y^2/n_Y$. Bajo H_0 se supone que $\mu_X - \mu_Y = 0$. Al sustituir los valores de s_X , s_Y , n_X y n_Y la desviación estándar es $\sqrt{0.25^2/544 + 0.26^2/581} = 0.01521$. Por tanto, la distribución nula de $\bar{X} - \bar{Y}$ es

$$\bar{X} - \bar{Y} \sim N(0, 0.01521^2)$$

El puntaje z es

$$z = \frac{-0.03 - 0}{0.01521} = -1.97$$

Ésta es una prueba de dos colas, y el P -valor es 0.0488 (véase la figura 6.10). De acuerdo con la hipótesis nula se rechazaría la regla de 5% . Es razonable ser escéptico sobre la veracidad de H_0 .

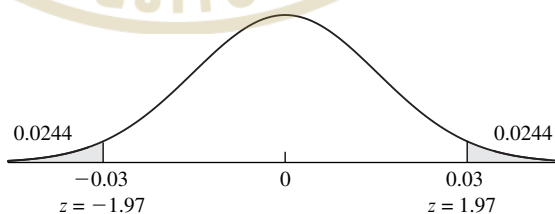


FIGURA 6.10 Solución del ejemplo 6.9.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.9.

Two-sample T for Argon vs CO2

	N	Mean	StDev	SE Mean
Argon	544	0.37	0.25	0.010719
CO2	581	0.40	0.26	0.010787

Difference = mu (Argon) - mu (CO2)

Estimate for difference: 0.030000

95% confidence bound for difference:

(-0.0598366, -0.000163)

T-Test of difference = 0 (vs not = 0):

T-Value = -1.97 P-Value = 0.049 DF = 1122

Observe que para esta prueba la computadora utiliza el estadístico t en lugar del estadístico z . Muchos softwares emplean el estadístico t siempre que se utiliza una desviación estándar muestral para estimar la desviación estándar poblacional. Cuando el tamaño de muestra es grande, la diferencia entre t y z es insignificante para fines prácticos. Cuando se utilizan tablas en vez de una computadora, el puntaje z tiene la ventaja de que el P -valor se puede determinar con mayor precisión mediante una tabla z que con una t .

Los métodos descritos en esta sección se pueden utilizar para evaluar la hipótesis de que dos medias poblacionales son diferentes por una constante específica. El ejemplo 6.10 muestra cómo.

Ejemplo

6.10

Con referencia al ejemplo 6.9, ¿se puede concluir que la media del diámetro para soldaduras de dióxido de carbono (μ_Y) es mayor para las soldaduras de argón (μ_X) en más de $0.015 \mu\text{m}$?

Solución

Las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y \geq -0.015 \quad \text{contra} \quad H_1: \mu_X - \mu_Y < -0.015$$

Se observa $\bar{X} = 0.37$, $\bar{Y} = 0.40$, $s_X = 0.25$, $s_Y = 0.26$, $n_X = 544$ y $n_Y = 581$. Bajo H_0 se tiene que $\mu_X - \mu_Y = -0.015$. La distribución nula de $\bar{X} - \bar{Y}$ dada por la expresión (6.2) es

$$\bar{X} - \bar{Y} \sim N(-0.015, 0.01521^2)$$

Se observa que $\bar{X} - \bar{Y} = 0.37 - 0.40 = -0.03$. El puntaje z es

$$z = \frac{-0.03 - (-0.015)}{0.01521} = -0.99$$

Ésta es una prueba de una cola. El P -valor es 0.1611. No se puede concluir que la media del diámetro de inclusiones de soldaduras de dióxido de carbono sea mayor que la de las soldaduras de argón en más de $0.015 \mu\text{m}$.

Resumen

Sean $X_1, \dots, X_{n_x}$ y $Y_1, \dots, Y_{n_y}$ muestras *grandes* (por ejemplo, $n_x > 30$ y $n_y > 30$) de las poblaciones con medias μ_X y μ_Y y las desviaciones estándar σ_X y σ_Y , respectivamente. Suponga que las muestras se extraen en forma independiente una de la otra.

Para probar una hipótesis nula de la forma $H_0: \mu_X - \mu_Y \leq \Delta_0$, $H_0: \mu_X - \mu_Y \geq \Delta_0$, o $H_0: \mu_X - \mu_Y = \Delta_0$.

- Calcule el puntaje z : $z = \frac{(\bar{X} - \bar{Y}) - \Delta_0}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}}$. Si σ_X y σ_Y son desconocidas se pueden aproximar con s_X y s_Y , respectivamente.
- Calcule el P -valor. El P -valor es un área debajo de la curva normal que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: \mu_X - \mu_Y > \Delta_0$$

$$H_1: \mu_X - \mu_Y < \Delta_0$$

$$H_1: \mu_X - \mu_Y \neq \Delta_0$$

P -valor

Área a la derecha de z

Área a la izquierda de z

Suma de las áreas en las colas correspondientes a z y $-z$

Ejercicios de la sección 6.5

- El artículo "Measurement of Complex Permittivity of Asphalt Paving Materials" (J. Shang, J. Umana y cols., en *Journal of Transportation Engineering*, 1999:347-356) compara las constantes dieléctricas entre dos tipos de asfalto, HL3 y HL8, utilizados comúnmente en los pavimentos. Para 42 elementos del asfalto HL3 el promedio de la constante dieléctrica fue de 5.92 con desviación estándar de 0.15, y para 37 elementos del asfalto HL8 el promedio de la constante dieléctrica fue 6.05 con desviación estándar de 0.16. ¿Se puede concluir que la media de las constantes dieléctricas difiere entre los dos tipos de asfalto?
- Para determinar el efecto del grado de combustible en la eficiencia del combustible, 80 nuevos automóviles de la misma marca, con motores idénticos, fueron conducidos cada uno durante mil millas. Cuarenta de los automóviles funcionaron con combustible regular y otros 40 con combustible de grado premium; los primeros tenían una media de 27.2 milla/galón, con desviación estándar de 1.2 milla/galón; los segundos tenían una media de 28.1 milla/galón y una desviación estándar de 2.0 milla/galón. ¿Puede concluir que este tipo de automóvil tiene mejor millaje con combustible premium?
- Se comparan dos métodos utilizados para purificar una proteína. En 50 operaciones del método A la media de recuperación fue de 60% y la desviación estándar de 15%, mientras que en 60 operaciones del método B la media de recuperación fue de 65% y la desviación estándar de 20%. ¿Puede concluir que hay una diferencia en las dos tasas de recuperación?
- Se comparan dos máquinas utilizadas para llenar envases de refresco. Se cuenta durante 60 minutos el número de envases llenados cada minuto por cada máquina. Durante 60 minutos, la máquina número 1 llenó en promedio 73.8 latas por minuto con desviación estándar de 5.2 latas por minuto, y la máquina número 2 llenó en promedio 76.1 latas por minuto con desviación estándar de 4.1 latas por minuto.
 - Si los recuentos se hacían cada minuto durante 60 minutos consecutivos, ¿qué suposición es necesaria para que pueda violarse la validez de una prueba de hipótesis?
 - Suponiendo que se satisfacen todas las suposiciones necesarias, realice una prueba de hipótesis. ¿Puede concluir que la máquina número 2 es más rápida que la número 1?
- Un profesor de estadística que imparte una clase a 160 estudiantes quiere determinar si los estudiantes tienen más dificultad con las pruebas de hipótesis de una cola o con las de dos colas. En el siguiente examen, 80 de los estudiantes, elegidos aleatoriamente, reciben una versión del examen con una pregunta de 10 puntos que requiere una prueba de

una cola. Los otros 80 estudiantes reciben una pregunta idéntica, excepto que requiere una prueba de dos colas. Los estudiantes de la pregunta de una cola tienen calificación promedio de 7.79 puntos, y su desviación estándar es 1.06 puntos. Los estudiantes de la pregunta de dos colas tienen una calificación promedio de 7.64 puntos, y su desviación estándar es de 1.31 puntos.

- a) ¿Puede concluir que la calificación promedio μ_1 en la pregunta de una prueba de hipótesis de una cola es mayor que la calificación promedio μ_2 de la pregunta de la prueba de hipótesis de dos colas? Establezca las hipótesis nula y alternativa adecuadas, y después calcule el P -valor.
 - b) ¿Puede concluir que la calificación promedio μ_1 de la pregunta de una prueba de hipótesis de una cola difiere de la calificación promedio μ_2 de la pregunta de la prueba de hipótesis de dos colas? Establezca las hipótesis nula y alternativa adecuadas, y después calcule el P -valor.
6. Se probó la velocidad en cierta aplicación de 50 chips nuevos de computadora, con otra cantidad igual de diseño viejo. La velocidad promedio, en MHz, de los nuevos chips fue de 495.6, y la desviación estándar de 19.4. La velocidad promedio de los chips viejos fue de 481.2, y la desviación estándar fue de 14.3.
- a) ¿Se puede concluir que la media de la velocidad de los nuevos chips es mayor que la de los chips viejos? Establezca las hipótesis nula y alternativa adecuadas, y después encuentre el P -valor.
 - b) Una muestra de 60 chips aún más viejos tenía velocidad promedio de 391.2 MHz, con desviación estándar de 17.2 MHz. Alguien afirma que los nuevos chips tienen una velocidad promedio mayor a 100 MHz que los más viejos. ¿Los datos proporcionan evidencias convincentes para esta afirmación? Establezca las hipótesis nula y alternativa, y después determine el P -valor.
7. Se consideran dos métodos en un proceso de fabricación de pintura con la finalidad de aumentar la producción. En una muestra aleatoria de 100 días, la media de producción diaria con el primer método fue de 625 toneladas, y la desviación estándar era de 40 toneladas. En una muestra aleatoria de 64 días, la media de la producción diaria que usaba el segundo método era de 640 toneladas, y la desviación estándar, de 50 toneladas. Suponga que las muestras son independientes.
- a) ¿Puede concluir que el segundo método produce la media de la producción diaria mayor?
 - b) ¿Puede concluir que la media de la producción diaria con el segundo método es mayor que la del primer método en más de diez toneladas?
8. Con referencia al ejercicio 7 de la sección 5.4, ¿puede concluir que la media de la dureza de las soldaduras enfriadas a una tasa de 10°C/s es mayor que la de las soldaduras enfriadas a una tasa de 30°C/s ?
9. La Corporación Eléctrica Subterránea (CES) afirma que sus fuentes de energía suministradas para computadoras personales duran más que la de su competidor, Productos para Aparatos Zircon (PAZ). Se toman muestras aleatorias independientes de 75 unidades de cada una de las fuentes de energía de los dos fabricantes y se calculan las medias y las desviaciones estándar muestrales
- CES: $\bar{X}_1 = 4\,387\text{ h}$ $s_1 = 252\text{ h}$
 PAZ: $\bar{X}_2 = 4\,260\text{ h}$ $s_2 = 231\text{ h}$
- ¿Puede concluir que las fuentes de potencia CES duran más que las de PAZ? ¿Cuál es el P -valor para esta prueba?

10. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales:

Two-sample T for X vs Y

	N	Mean	StDev	SE Mean
X	135	3.94	2.65	0.23
Y	180	4.43	2.38	0.18

Difference = $\mu(X) - \mu(Y)$

Estimate for difference: -0.484442

95% upper bound for difference: -0.007380

T-Test of difference = 0 (vs <): T-Value = -1.68 P-Value = 0.047 DF = 270

- ¿Esta es una prueba de una o de dos colas?
- ¿Cuál es la hipótesis nula?
- ¿Se puede rechazar H_0 a un nivel de 5%? ¿Cómo justifica su respuesta?
- El resultado presenta una prueba t de Student. Calcule el P -valor mediante una prueba z . ¿Los dos resultados son similares?
- Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 99% para $\mu_X - \mu_Y$ basado en el estadístico z .

11. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales. Faltan algunos de los números.

Two-sample T for X vs Y

	N	Mean	StDev	SE Mean
X	78	23.3	(i)	1.26
Y	63	20.63	3.02	(ii)

Difference = $\mu(X) - \mu(Y)$

Estimate for difference: 2.670

95% CI for difference: (0.05472, 5.2853)

T-Test of difference = 0 (vs not =): T-Value = 2.03 P-Value = 0.045 DF = 90

- Complete los números faltantes en (i) y en (ii).
- El resultado presenta una prueba t de Student. Calcule el P -valor mediante una prueba z . ¿Los dos resultados son similares?
- Utilice el producto y una tabla adecuada para calcular un intervalo de confianza de 98% para $\mu_X - \mu_Y$ basado en el estadístico z .

6.6 Pruebas de hipótesis para la diferencia entre dos proporciones

El procedimiento para probar la diferencia entre dos proporciones es similar al que se utiliza cuando se prueba la diferencia entre dos medias. Se ilustra con un ejemplo.

Una red de computadoras móviles consta de un grupo de éstas que mantienen comunicación inalámbrica entre sí dentro de un área específica. Un protocolo de direccionamiento representa un algoritmo que determina cómo se transmitirá de una máquina a otra en la red, para tener la mayor posibilidad de que lleguen a su destino los mensajes. El artículo "Performance Comparison of two Location Based Routing Protocols" (T. Camp, J. Boleng y cols., en *Proceedings of the IEEE International Conference on Communications*, 2002:3318-3324) compara la eficacia de dos protocolos de direccionamiento en muchas métricas, incluyendo la tasa de envíos exitosos. Suponga que mediante el protocolo A se enviaron 200 mensajes, y 170 de ellos, 85%, fueron recibidos con éxito. Con el protocolo B, se enviaron 150 mensajes, y 123 de ellos, 82%, fueron recibidos con éxito. ¿Se puede concluir que el protocolo A tiene mayor tasa de éxito?

En este ejemplo las muestras constan de éxitos y fracasos. Sea X el número de mensajes enviados con éxito al usar el protocolo A, y sea Y el número de mensajes enviados con éxito empleando el protocolo B. Los valores observados en este ejemplo son $X = 170$ y $Y = 123$.

Sea p_X la proporción de mensajes que se enviaron con éxito mediante el protocolo A y p_Y la proporción correspondiente del protocolo B. Los valores p_X y p_Y son desconocidos.

Las variables aleatorias X y Y tienen distribuciones binomiales, con $n_X = 200$ y $n_Y = 150$ ensayos, respectivamente. Las probabilidades de éxito son p_X y p_Y . Las proporciones muestrales de los valores observados son $\hat{p}_X = 170/200 = 0.85$ y $\hat{p}_Y = 123/150 = 0.82$.

Las hipótesis nula y alternativa son

$$H_0: p_X - p_Y \leq 0 \quad \text{contra} \quad H_1: p_X - p_Y > 0$$

La prueba está basada en el estadístico $\hat{p}_X - \hat{p}_Y$. Se debe determinar la distribución nula de este estadístico. Por el teorema del límite central, puesto que n_X y n_Y son grandes,

$$\hat{p}_X \sim N\left(p_X, \frac{p_X(1-p_X)}{n_X}\right) \quad \hat{p}_Y \sim N\left(p_Y, \frac{p_Y(1-p_Y)}{n_Y}\right)$$

Por tanto,

$$\hat{p}_X - \hat{p}_Y \sim N\left(p_X - p_Y, \frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}\right) \quad (6.3)$$

Obtener la distribución nula requiere sustituir los valores para $p_X - p_Y$ y la varianza $p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y$. La media se obtiene fácilmente. La hipótesis nula específica que $p_X - p_Y \leq 0$, por lo que se tiene que $p_X - p_Y = 0$. La obtención de la varianza es un poco más difícil. A primera vista, podría parecer razonable aproximar la desviación estándar al sustituir las proporciones muestrales $\hat{p}_X$ y $\hat{p}_Y$ para las proporciones poblacionales p_X y p_Y . Sin embargo, la hipótesis nula H_0 especifica que las proporciones poblacionales son iguales. Por tanto, se deben calcular ambas con un valor común. El valor adecuado es la **proporción combinada**, obtenida al dividir el número total de éxitos de ambas muestras entre el tamaño muestral total. Este valor es

$$\hat{p} = \frac{X + Y}{n_X + n_Y}$$

En consecuencia, la distribución nula de $\hat{p}_X - \hat{p}_Y$ se estima cuando se sustituye la proporción combinada $\hat{p}$ tanto para p_X como para p_Y en la expresión (6.3). Se obtiene

$$\hat{p}_X - \hat{p}_Y \sim N\left(0, \hat{p}(1-\hat{p})\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)\right) \quad (6.4)$$

En este ejemplo $\hat{p} = (170 + 123)/(200 + 150) = 0.837$. Bajo H_0 se tiene que $p_X - p_Y = 0$. Por tanto, la distribución nula de $\hat{p}_X - \hat{p}_Y$ es normal con media 0 y desviación estándar $\sqrt{0.837(1-0.837)(1/200 + 1/150)} = 0.0399$. El valor observado de $\hat{p}_X - \hat{p}_Y$ es $0.85 - 0.82 = 0.03$. Por consiguiente, el puntaje z es

$$z = \frac{0.03 - 0}{0.0399} = 0.75$$

El P -valor es 0.2266. La figura 6.11 muestra la distribución nula y señala la ubicación del estadístico de prueba. Con base en este P -valor no se puede concluir que el protocolo B tenga una tasa de éxito mayor. Observe que para que el teorema del límite central sea válido, ambas muestras deben ser razonablemente grandes. Una buena regla general es que haya al menos diez éxitos y diez fracasos en cada muestra.

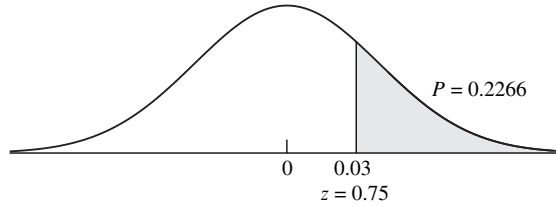


FIGURA 6.11 La distribución nula de $\hat{p}_X - \hat{p}_Y$ es $N(0, 0.0399^2)$. Por tanto, si H_0 es verdadera, la probabilidad que $\hat{p}_X - \hat{p}_Y$ adquiere un valor mayor o igual que el observado de 0.03 es 0.2266. Éste es el P -valor.

Ejemplo

6.11

A menudo, las compañías industriales emplean métodos de “transferencia de riesgo”, un seguro o las cláusulas de indemnización en los contratos, como una técnica de administración de riesgos. El artículo “Survey of Risk Management in Major U.K. Companies” (S. Baker, K. Ponniah y S. Smith, en *Journal of Professional Issues in Engineering Education and Practice*, 1999:94-102) informa los resultados de una encuesta en la cual se les pregunta a los directivos acerca de los métodos importantes en la estrategia de administración de riesgos de sus compañías. En una muestra de 43 compañías petroleras, 22 indicaban que la transferencia de riesgos fue determinante, mientras que en una muestra de 93 compañías constructoras 55 confirmaron lo anterior. (Estas cifras se obtuvieron a partir de una gráfica.) ¿Se puede concluir que la proporción de compañías petroleras que emplean el método de transferencia de riesgos es menor que la proporción de compañías constructoras que lo hacen?

Solución

Sea $\hat{p}_X = 22/43 = 0.5116$ la proporción muestral de compañías petroleras que emplean métodos de transferencia de riesgos, y $\hat{p}_Y = 55/93 = 0.5914$ la proporción muestral correspondiente de las compañías constructoras. Los tamaños muestrales son $n_X = 43$ y $n_Y = 93$. Sean p_X y p_Y las proporciones poblacionales de las compañías petroleras y constructoras, respectivamente. Las hipótesis nula y alternativa son

$$H_0: p_X - p_Y \geq 0 \quad \text{contra} \quad H_1: p_X - p_Y < 0$$

La prueba se basa en $\hat{p}_X - \hat{p}_Y$. Dado que ambas muestras son grandes, la distribución nula de $\hat{p}_X - \hat{p}_Y$ está dada por la expresión (6.4). La proporción combinada es

$$\hat{p} = \frac{22 + 55}{43 + 93} = 0.5662$$

La distribución nula es normal con media 0 y desviación estándar

$\sqrt{0.5662(1 - 0.5662)(1/43 + 1/93)} = 0.0914$. El valor observado de $\hat{p}_X - \hat{p}_Y$ es $0.5116 - 0.5914 = -0.0798$. El puntaje z es

$$z = \frac{-0.0798 - 0}{0.0914} = -0.87$$

El P -valor es 0.1922 (véase la figura 6.12). No se puede concluir que la proporción de compañías petroleras que emplean métodos de transferencia de riesgos sea menor que la proporción de compañías constructoras que lo hacen.

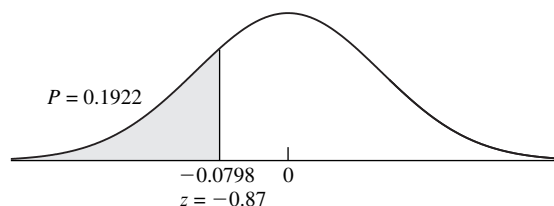


FIGURA 6.12 Solución al ejemplo 6.11.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.11.

```

Test and CI for Two Proportions: Oil, Indus.
Variable  X    N    Sample p
Oil       22   43    0.511628
Indus.    55   93    0.591398
Difference = p (Oil) - p (Indus.)
Estimate for difference: -0.079770
95% Upper Bound for difference: 0.071079
Test for difference = 0 (vs < 0): Z = -0.87  P-Value = 0.192

```

El resultado se explica por sí mismo. Observe que el límite superior de confianza de 95% se calcula con el método tradicional (la expresión 5.19 de la sección 5.5 presenta las dos versiones de este método).

Resumen

Sea $X \sim \text{Bin}(n_X, p_X)$ y $Y \sim \text{Bin}(n_Y, p_Y)$. Suponga que tanto n_X como n_Y son grandes, y que X y Y , son independientes.

Para probar una hipótesis nula de la forma $H_0: p_X - p_Y \leq 0$, $H_0: p_X - p_Y \geq 0$, o $H_0: p_X - p_Y = 0$:

■ Calcule $\hat{p}_X = \frac{X}{n_X}$, $\hat{p}_Y = \frac{Y}{n_Y}$, y $\hat{p} = \frac{X + Y}{n_X + n_Y}$.

■ Calcule el puntaje z : $z = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1 - \hat{p})(1/n_X + 1/n_Y)}}$.

■ Calcule el P -valor. Éste es un área bajo la curva normal que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: p_X - p_Y > 0$$

$$H_1: p_X - p_Y < 0$$

$$H_1: p_X - p_Y \neq 0$$

P -valor

Área a la derecha de z

Área a la izquierda de z

Suma de las áreas de las colas correspondientes a z y $-z$

Ejercicios de la sección 6.6

- Se comparan dos máquinas de extrusión que fabrican varillas de acero. En una muestra de mil varillas tomadas de la máquina número 1, 960 satisfacían las especificaciones de longitud y diámetro. En otra muestra de 600 varillas tomadas de la máquina número 2, 582 cumplían las especificaciones. La máquina número 2 tiene un costo de operación más alto, por lo que se decide que se usará la máquina número 1 a menos que se demuestre claramente que la máquina número 2 produce mayor proporción de varillas que satisfacen las especificaciones.
 - Establezca las hipótesis nula y alternativa adecuadas para tomar la decisión respecto de qué máquina se utilizará.
 - Calcule el P -valor.
 - ¿Qué máquina se debe utilizar?
- Se compran resistores etiquetados con $100\ \Omega$ a dos distribuidores diferentes. La especificación para este tipo de resistor es que su resistencia verdadera esté dentro del 5% de su resistencia etiquetada. En una muestra de 180 resistores del distribuidor A, 150 de éstos satisfacían la especificación. En otra muestra de 270 resistores comprados al distribuidor B, 233 cumplían la especificación. El distribuidor A es el proveedor actual, pero si los datos demuestran convincentemente que una proporción mayor de los resistores del distribuidor B satisface la especificación, se hará el cambio.
 - Establezca las hipótesis nula y alternativa adecuadas.
 - Determine el P -valor.
 - ¿Se debe hacer el cambio?
- El artículo “Strategic Management in Engineering Organizations” (P. Chinowsky, en *Journal of Management in Engineering*, 2001:60-68) presenta los resultados de una encuesta de estilos de dirección que se realizó tanto a compañías constructoras privadas como entidades públicas. De un total de 400 compañías privadas contactadas, 133 contestaron completamente la encuesta, mientras que de 100 entidades públicas sólo 50 respondieron. ¿Puede llegar a la conclusión de que la tasa de respuesta es diferente entre las compañías privadas y las entidades públicas?
- El artículo “Training Artificial Neural Networks with the Aid of Fuzzy Sets” (C. Juang, S. Ni y C. Lu, en *Computer-Aided Civil and Infrastructure Engineering*, 1999:407-415) describe el desarrollo de redes neurales artificiales diseñadas para predecir la plegabilidad de suelos. Un modelo con una capa oculta tiene una predicción de éxito en 48 de los 60 casos, mientras que otro con dos capas ocultas tiene una predicción de éxito en 44 de los 60 casos. Suponiendo que estas muestras son independientes, ¿puede concluir que el modelo con una capa oculta tiene una tasa de éxito mayor?
- En una encuesta de 100 usuarios de una tarjeta de crédito elegidos aleatoriamente, 57 dijeron que sabían que empleando la tarjeta podían ganar millas de viajero frecuente en cierta aerolínea. Después de una campaña publicitaria para difundir este beneficio, se realizó una encuesta independiente entre 200 usuarios de tarjeta de crédito, y 135 dijeron que conocían el beneficio. ¿Puede concluir que el conocimiento de este beneficio aumentó después de la campaña publicitaria?
- El artículo “Modeling the Inactivation of Particle-Associated Coliform Bacteria” (R. Emerick, F. Loge y cols., en *Water Environment Research*, 2000:432-438) presenta los recuentos de números de partículas de diferentes tamaños en muestras de agua residual que contenían bacterias coliformes. De las 161 partículas de $75\text{--}80\ \mu\text{m}$ de diámetro, 19 contenían bacterias coliformes, y de las 95 partículas de $90\text{--}95\ \mu\text{m}$ de diámetro, 22 contenían bacterias coliformes. ¿Puede concluir que las partículas más grandes tienen una probabilidad mayor de tener bacterias coliformes?
- Para probar la eficacia de los empaques de protección, una compañía envió 1 200 órdenes con un empaque ligero común y 1 500 órdenes con un empaque de gran resistencia. De las órdenes enviadas con el empaque ligero, 20 llegaron deterioradas, mientras que de las órdenes enviadas con el otro empaque, 15 llegaron deterioradas. ¿Puede concluir que el empaque de gran resistencia reduce la proporción de órdenes deterioradas?
- En una muestra de 100 lotes de un producto químico comprado al distribuidor A, 70 satisfacen una especificación de pureza. En una muestra de 70 lotes comprada al distribuidor B, 61 satisfacen la especificación. ¿Puede concluir que una proporción mayor de los lotes del distribuidor B satisface la especificación?
- En el artículo “Nitrate Contamination of Alluvial Groundwaters in the Nakdong River Basin, Korea” (J. Min, S. Yun y cols., en *Geosciences Journal*, 2002: 35-46) se describen 41 muestras de agua tomadas de pozos en el área Daesan, y se encontró que 22 satisfacían las normas de calidad de potabilidad. Se tomaron 31 muestras del área de Yongdang, y se encontró que 18 satisfacían las normas. ¿Puede concluir que la proporción de pozos que satisfacían las normas es diferente entre las dos áreas?

10. En una prueba clínica para comparar la efectividad de dos analgésicos, a una muestra de 100 pacientes se le suministró el medicamento A, y en una muestra independiente de 200 pacientes se les suministró el medicamento B. De los pacientes que recibieron el medicamento A, 76 notificaron un importante alivio, mientras que del otro grupo de pacientes, 128 informaron lo mismo. ¿Puede concluir que el medicamento A es más efectivo que el B?
11. Con el propósito de determinar si una nueva campaña de publicidad estaba dirigida más hacia hombres o a mujeres, un publicista proporcionó a cada pareja de una muestra aleato-

ria de 500 parejas casadas un nuevo tipo de control remoto de TV que, se supone, es más fácil de encontrarse cuando se le necesita. De los 500 esposos, 62% dijo que el nuevo control remoto era más fácil de encontrarse que el viejo. De las 500 esposas, solamente 54% dijo que el nuevo control remoto era más fácil de encontrarse. Sea p_1 la proporción poblacional de hombres casados que consideran que el nuevo control remoto es más fácil de encontrar, y sea p_2 la proporción correspondiente de mujeres casadas. ¿Puede utilizarse el estadístico $\hat{p}_1 - \hat{p}_2 = 0.62 - 0.54$ para probar $H_0: p_1 - p_2 = 0$ contra $H_1: p_1 - p_2 \neq 0$? Si es así, realice la prueba y calcule el P -valor. Si no, explique por qué.

12. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $p_1 - p_2$ entre dos proporciones poblacionales.

Test and CI for Two Proportions

Sample	X	N	Sample p
1	41	97	0.422680
2	37	61	0.606557

Difference = $p(1) - p(2)$

Estimate for difference: -0.183877

95% CI for difference: (-0.341016, -0.026738)

Test for difference = 0 (vs not = 0): $Z = -2.25$ P-Value = 0.024

- a) ¿Esta es una prueba de una o de dos colas?
- b) ¿Cuál es la hipótesis nula?
- c) ¿Se puede rechazar H_0 a un nivel de 5%? ¿Cómo justifica su respuesta?

13. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $p_1 - p_2$ entre dos proporciones poblacionales. Faltan algunos de los números. Complete los números del (a) al (d).

Test and CI for Two Proportions

Sample	X	N	Sample p
1	101	153	(a)
2	(b)	90	0.544444

Difference = $p(1) - p(2)$

Estimate for difference: 0.115686

95% CI for difference: (-0.0116695, 0.243042)

Test for difference = 0 (vs not = 0): $Z = (c)$ P-Value = (d)

6.7 Pruebas de hipótesis para la diferencia entre dos medias con muestras pequeñas

La prueba t se puede utilizar en algunos casos donde las muestras son pequeñas, por lo cual el teorema del límite central no es aplicable. Aquí se presenta un ejemplo.

El artículo “The Achondroplasia Paternal Age Effect Is Not Explained By an Increase in Mutant Frequency” (I. Tiemann-Boege, W. Navidi, y cols., en *Proceedings of the National Academy of Sciences*, 2002: 14952-14957) describe un experimento en el que se cuenta un número de moléculas de ADN, y se necesita determinar si éstas contienen cierta secuencia de nucleótidos. Lo anterior se hace al repetir el experimento con una enzima adicional que digiere la secuencia de interés. Si el número promedio es menor con la enzima presente, entonces se concluye que las moléculas que se están contando contienen la secuencia.

Suponga que en seis muestras preparadas de forma idéntica, sin la enzima presente, los números de moléculas contadas son 33, 30, 26, 22, 37 y 34. Suponga que en cuatro muestras preparadas de forma idéntica, sin la enzima presente, los números de moléculas fueron 22, 29, 25 y 23. ¿Se puede concluir que dicho número es menor cuando la enzima está presente?

Se tienen solamente algunas observaciones para cada proceso, así que el teorema del límite central no es aplicable. Si ambas poblaciones son aproximadamente normales, la distribución t de Student se puede utilizar para construir una prueba de hipótesis.

Sean $X_1, \dots, X_6$ los números de moléculas obtenidos sin la enzima, y $Y_1, \dots, Y_4$ los que tienen la enzima. Sean μ_X y μ_Y las medias poblacionales de las cuales se tomaron estas muestras; y n_X y n_Y los tamaños de muestra. Las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y \leq 0 \quad \text{contra} \quad H_1: \mu_X - \mu_Y > 0$$

Se supone que ambas poblaciones siguen distribuciones normales. Por tanto (como se analizó en la sección 5.6), la cantidad

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}} \quad (6.5)$$

tiene una aproximada distribución t de Student con ν grados de libertad, donde

$$\nu = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{(s_X^2/n_X)^2}{n_X - 1} + \frac{(s_Y^2/n_Y)^2}{n_Y - 1}} \quad \text{redondeando hacia abajo al entero más próximo.}$$

Los valores observados para las medias y desviaciones estándar muestrales son $\bar{X} = 30.333$, $\bar{Y} = 24.750$, $s_X = 5.538$, $s_Y = 3.096$. Los tamaños muestrales son $n_X = 6$ y $n_Y = 4$. Al sustituir los valores para las desviaciones estándar y los tamaños muestrales, se calcula que $\nu = 7.89$, que se redondea hacia abajo a 7. Bajo H_0 , $\mu_X - \mu_Y = 0$. Por tanto, el estadístico de prueba es

$$t = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

Bajo H_0 , el estadístico de prueba tiene la distribución t de Student con siete grados de libertad. Al sustituir los valores de $\bar{X}$, $\bar{Y}$, s_X , s_Y , n_X y n_Y se calcula que el estadístico de prueba es

$$t = \frac{5.583 - 0}{2.740} = 2.038$$

Al consultar la tabla t con siete grados de libertad, se encuentra que el valor que corresponde al 5% en la cola derecha es 1.895, y el que corresponde a 2.5% es 2.365. Por tanto, el P -valor se encuentra entre 0.025 y 0.05 (véase la figura 6.13). Se concluye que el número promedio de moléculas es menor cuando la enzima está presente.

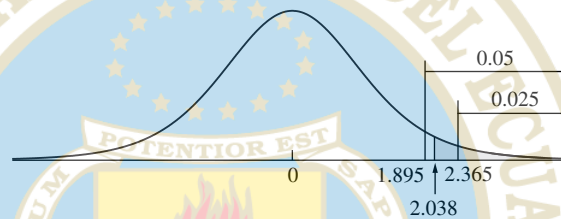


FIGURA 6.13 La distribución nula es la distribución t de Student con siete grados de libertad. El valor observado del estadístico de prueba es 2.038. Si H_0 es verdadera, la probabilidad de que t tenga un valor mayor o igual que el observado está entre 2.5 y 5%.

Ejemplo 6.12

Un buen diseño del sitio **web** puede hacer la navegación más fácil. El artículo “The Implications of Visualization Ability and Structure Preview Design for Web Information Search Tasks” (H. Zhang y G. Salvendy, en *International Journal of Human-Computer Interaction* 2001:75-95) presenta una comparación del reconocimiento de elementos entre dos diseños. Una muestra de diez usuarios que utilizan un diseño web convencional tuvo promedio de 32.3 elementos identificados, con desviación estándar de 8.56. Una muestra de diez usuarios que utilizan un nuevo diseño web estructurado tuvo promedio de 44.1 elementos identificados, con desviación estándar de 10.09. ¿Se puede concluir que el número promedio de elementos identificados es mayor con el nuevo diseño estructurado?

Solución

Sea $\bar{X} = 44.1$ la media muestral para el diseño web estructurado. Entonces $s_X = 10.09$ y $n_X = 10$. Sea $\bar{Y} = 32.3$ la media muestral para el diseño web convencional. Entonces $s_Y = 8.56$ y $n_Y = 10$. Sean μ_X y μ_Y las medias poblacionales de las mediciones hechas mediante los métodos estructurado y convencional, respectivamente. Las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y \leq 0 \quad \text{contra} \quad H_1: \mu_X - \mu_Y > 0$$

El estadístico de prueba es

$$t = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

Al sustituir valores para $\bar{X}$, $\bar{Y}$, s_X , s_Y , n_X y n_Y , se calcula que el valor del estadístico de prueba es de $t = 2.820$. Bajo H_0 , éste tiene una distribución aproximada t de Student, con el número de grados de libertad dado por

$$v = \frac{\left(\frac{10.09^2}{10} + \frac{8.56^2}{10}\right)^2}{\frac{(10.09^2/10)^2}{9} + \frac{(8.56^2/10)^2}{9}} = 17.53 \approx 17$$

Al consultar la tabla t con 17 grados de libertad, se encuentra que el valor correspondiente a 1% en la cola derecha es 2.567, y el de 0.5%, 2.898. Por tanto, el área en cola derecha correspondiente a los valores que son mayores o iguales que el valor observado de 2.820 se encuentra entre 0.005 y 0.010. En consecuencia, $0.005 < P < 0.01$ (véase la figura 6.14). Hay gran evidencia de que el número promedio de elementos identificados es mayor en el diseño nuevo.

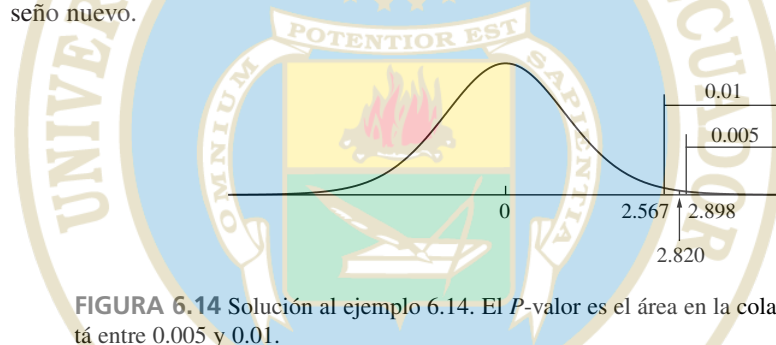


FIGURA 6.14 Solución al ejemplo 6.14. El P -valor es el área en la cola derecha, que está entre 0.005 y 0.01.

El siguiente resultado de computadora (de MINITAB) presenta los resultados del ejemplo 6.12.

Two-Sample T-Test and CI: Struct, Conven

Two-sample T for C1 vs C2

	N	Mean	StDev	SE Mean
Struct	10	44.10	10.09	3.19074
Conven	10	32.30	8.56	2.70691

Difference = mu (Struct) - mu (Conven)

Estimate for difference: 11.8000

95% lower bound for difference: 4.52100

T-Test of difference = 0 (vs >):

T-Value = 2.82 P-Value = 0.006 DF = 17

Observe que el límite inferior de confianza de 95% es consistente con la hipótesis alternativa. Esto último indica que el P -valor es menor a 5%.

Los métodos descritos en esta sección se pueden utilizar para probar la hipótesis de que dos medias poblacionales difieren por una constante especificada. El ejemplo 6.13 muestra cómo.

Ejemplo

6.13

Con referencia al ejemplo 6.12, ¿puede concluir que la media del número de elementos identificados con el nuevo diseño estructurado es mayor al diseño convencional por más de 2?

Solución

Las hipótesis nula y alternativa son

$$H_0: \mu_X - \mu_Y \leq 2 \quad \text{contra} \quad H_1: \mu_X - \mu_Y > 2$$

Se observa que $\bar{X} = 44.1$, $\bar{Y} = 32.3$, $s_X = 10.09$, $s_Y = 8.56$, $n_X = 10$ y $n_Y = 10$. Bajo H_0 , se tiene que $\mu_X - \mu_Y = 2$. El estadístico de prueba está dado por la expresión (6.5) como

$$t = \frac{(\bar{X} - \bar{Y}) - 2}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

Bajo H_0 , el estadístico de prueba tiene la distribución t de Student con 17 grados de libertad. Observe que el número de grados de libertad se calcula del mismo modo que en el ejemplo 6.12. El valor del estadístico de prueba es $t = 2.342$. Ésta es una prueba de una cola. El P -valor está entre 0.01 y 0.025. Se concluye que la media del número de elementos identificados con el nuevo diseño estructurado es mayor que con el diseño convencional por más de 2.

Resumen

Sean $X_1, \dots, X_{n_X}$ y $Y_1, \dots, Y_{n_Y}$ muestras que tienen poblaciones normales con medias μ_X y μ_Y y desviaciones estándar σ_X y σ_Y , respectivamente. Suponga que las muestras se extraen de manera independiente entre sí.

Si no se conoce que σ_X y σ_Y son iguales, entonces, para probar una hipótesis nula de la forma $H_0: \mu_X - \mu_Y \leq \Delta_0$, $H_0: \mu_X - \mu_Y \geq \Delta_0$, o $H_0: \mu_X - \mu_Y = \Delta_0$:

- Calcular $\nu = \frac{[(s_X^2/n_X) + (s_Y^2/n_Y)]^2}{[(s_X^2/n_X)^2/(n_X - 1)] + [(s_Y^2/n_Y)^2/(n_Y - 1)]}$, redondeado hacia abajo al entero más próximo.

- Calcular el estadístico de prueba $t = \frac{(\bar{X} - \bar{Y}) - \Delta_0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$.

- Calcular el P -valor. Éste es un área debajo la curva t de Student con ν grados de libertad, que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: \mu_X - \mu_Y > \Delta_0$$

$$H_1: \mu_X - \mu_Y < \Delta_0$$

$$H_1: \mu_X - \mu_Y \neq \Delta_0$$

P -valor

Área a la derecha de t

Área a la izquierda de t

Suma de las áreas en las colas correspondientes a t y $-t$

Cuando las poblaciones tienen varianzas iguales

Cuando se conoce que las varianzas poblacionales son aproximadamente iguales es posible utilizar la varianza combinada (véase la sección 5.6). Esta última la da

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}$$

El estadístico de prueba para probar cualesquiera de las hipótesis nulas $H_0: \mu_X - \mu_Y = 0$, $H_0: \mu_X - \mu_Y \leq 0$, o $H_0: \mu_X - \mu_Y \geq 0$ es

$$t = \frac{\bar{X} - \bar{Y}}{s_p \sqrt{1/n_X + 1/n_Y}}$$

Bajo H_0 , el estadístico de prueba tiene la distribución t de Student con $n_X + n_Y - 2$ grados de libertad.

Ejemplo 6.14

Se han desarrollado dos métodos para determinar el contenido de níquel del acero. En una muestra de cinco reproducciones del primer método sobre cierta clase de acero, la medición promedio (en porcentaje) fue $\bar{X} = 3.16$ y la desviación estándar $s_X = 0.042$. El promedio de las siete reproducciones del segundo método fue $\bar{Y} = 3.24$, y la desviación estándar, $s_Y = 0.048$. Suponga que se conoce que las varianzas poblacionales son casi iguales. ¿Se puede concluir que hay diferencia en las mediciones promedio entre los dos métodos?

Solución

Cuando se sustituyen los tamaños muestrales $n_X = 5$ y $n_Y = 7$ junto con las desviaciones estándar muestrales $s_X = 0.042$ y $s_Y = 0.048$, se calcula la desviación estándar combinada y se obtiene $s_p = 0.0457$.

Por tanto, el valor del estadístico de prueba es

$$t = \frac{3.16 - 3.24}{0.0457 \sqrt{1/5 + 1/7}} = -2.990$$

Bajo H_0 , el estadístico de prueba tiene la distribución t de Student con diez grados de libertad. Al consultar la tabla t de Student, se encuentra que el área debajo de la curva en cada cola se halla entre 0.01 y 0.005. Dado que la hipótesis nula establecía que las medias eran iguales, ésta es una prueba de dos colas, por lo que el P -valor representa la suma de las áreas en ambas colas. Se concluye que $0.01 < P < 0.02$ (véase la figura 6.15). Parece haber una diferencia en las medias de las mediciones entre los dos métodos.

No suponga que las varianzas poblacionales son iguales sólo porque las varianzas muestrales se aproximan

Es tentador suponer que las varianzas poblacionales son iguales siempre que las muestrales son aproximadamente iguales. Sin embargo, esta suposición no está justificada, ya que tal vez las varianzas muestrales sean casi iguales, incluso cuando las poblacionales son muy diferentes. La suposición de que éstas son iguales se hace sólo cuando se conocen los procesos que produjeron los datos que justificaban esta suposición. Véase el análisis en la p. 343.

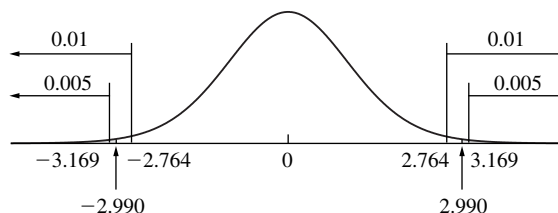


FIGURA 6.15 Solución al ejemplo 6.14. El P -valor es la suma de las áreas en ambas colas, que se encuentra entre 0.01 y 0.02.

Resumen

Sean $X_1, \dots, X_{n_X}$ y $Y_1, \dots, Y_{n_Y}$ muestras de poblaciones normales con medias μ_X y μ_Y y desviaciones estándar σ_X y σ_Y , respectivamente. Suponga que las muestras se extrajeron de manera independiente una de la otra.

Si se conoce que σ_X y σ_Y son iguales, entonces, para probar una hipótesis nula de la forma $H_0: \mu_X - \mu_Y \leq \Delta_0$, $H_0: \mu_X - \mu_Y \geq \Delta_0$, o $H_0: \mu_X - \mu_Y = \Delta_0$:

- Calcule $s_p = \sqrt{\frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}}$.
- Calcule el estadístico de prueba $t = \frac{(\bar{X} - \bar{Y}) - \Delta_0}{s_p \sqrt{1/n_X + 1/n_Y}}$.
- Calcule el P -valor. Éste es un área bajo la curva t de Student con $n_X + n_Y - 2$ grados de libertad, que depende de la hipótesis alternativa como se muestra a continuación:

Hipótesis alternativa

P -valor

$$H_1: \mu_X - \mu_Y > \Delta_0$$

Área a la derecha de t

$$H_1: \mu_X - \mu_Y < \Delta_0$$

Área a la izquierda de t

$$H_1: \mu_X - \mu_Y \neq \Delta_0$$

Suma de las áreas en las colas correspondientes a t y $-t$

Ejercicios para la sección 6.7

- Un fabricante de pinturas en tono pastel está comparando los efectos de dos clases de tintura amarilla para el brillo de las pinturas. La tintura B es más costosa que la A, pero se supone que puede producir una pintura con tono más fuerte. Se prueban cuatro pinturas en tonos pastel con cada tipo de tintura, y se mide para cada una la energía de impacto (en joules). Los resultados son:

Tintura A: 1.0 2.0 1.2 3.0

Tintura B: 3.0 3.2 2.6 3.4

- ¿Puede concluir que la media de la energía de las pinturas en tono pastel hechas con la tintura B es mayor que las hechas con la tintura A?

- b) ¿Puede concluir que la media de la energía de las pinturas en tono pastel hechas con la tintura B es mayor que las hechas con la tintura A por más de 1 J?

2. Se realiza un estudio para determinar si el chocolate semidulce se disuelve más rápidamente que el de leche. Ocho personas disolvieron una pieza de chocolate semidulce, mientras que siete disolvieron una pieza de chocolate de leche. Los tiempos de disolución, en segundos, fueron:

Semidulce: 30 55 50 22 46 45 30 44

De leche: 45 58 23 64 105 93 28

¿Puede concluir que la media de los tiempos de disolución del chocolate de leche es diferente a la del chocolate semidulce?

3. El artículo “Modeling Resilient Modulus and Temperature Correction for Saudi Roads” (H. Wahhab, I. Asi y R. Ramadhan, en *Journal of Materials in Civil Engineering*, 2001:298-305) describe un estudio diseñado para predecir el módulo elástico del pavimento a partir de sus propiedades físicas. Una de las cuestiones tratadas era si los módulos difieren cuando hay pavimento con y sin baches. A continuación se presentan mediciones del módulo elástico a 40°C (en 10⁶ kPa) de siete secciones de pavimento con baches y 12 sin baches.

Con baches: 1.48 1.88 1.90 1.29 3.53 2.43 1.00

Sin baches: 3.06 2.58 1.70 5.76 2.44 2.03 1.76
4.63 2.86 2.82 1.04 5.92

Realice una prueba de hipótesis para determinar si es factible que la media de los módulos elásticos sea la misma para pavimento con y sin baches. Calcule el *P*-valor. ¿Cuál es su conclusión?

4. El artículo “Time Series Analysis for Construction Productivity Experiments” (T. Abdelhamid y J. Everett, en *Journal of Construction Engineering and Management* 1999:87-95) presenta un estudio que compara la efectividad de un sistema de video que permite que un operador de grúa vea el punto de levantamiento con respecto a la operación del sistema antiguo, en el cual el operador se auxiliaba con las señales manuales de una persona. Se estudiaron tres levantamientos diferentes, A, B, y C. El levantamiento A fue de poca dificultad, el B, de moderada, y el C, de gran dificultad. Cada levantamiento se realizó varias veces, tanto con el nuevo sistema de video como con el sistema de señales manuales. Se registró el tiempo (en segundos) requerido para realizar cada levantamiento. Las siguientes tablas presentan las medias, las desviaciones estándar y los tamaños muestrales.

	Dificultad baja		
	Media	Desviación estándar	Tamaño muestral
Señales de una persona	47.79	2.19	14
Video	47.15	2.65	40

	Dificultad moderada		
	Media	Desviación estándar	Tamaño muestral
Señales de una persona	69.33	6.26	12
Video	58.50	5.59	24

	Gran dificultad		
	Media	Desviación estándar	Tamaño muestral
Señales de una persona	109.71	17.02	17
Video	84.52	13.51	29

- a) ¿Puede concluir que la media del tiempo para realizar un levantamiento de dificultad baja es menor cuando se utiliza el sistema de video que cuando se usa el sistema antiguo? Explique.
- b) ¿Puede concluir que la media del tiempo para realizar un levantamiento de dificultad moderada es menor cuando se utiliza el sistema moderno que cuando se emplea el sistema antiguo? Explique.
- c) ¿Puede concluir que la media del tiempo para realizar un levantamiento de gran dificultad es menor cuando se utiliza el sistema de video que cuando se emplea el sistema anterior? Explique.

5. El artículo “Calibration of an FTIR Spectrometer” (P. Pankrat, en *Statistical Case Studies for Industrial, and Process Improvement*, SIAM-ASA, 1997:19-38) describe el uso de un espectrómetro que hace cinco mediciones del contenido de carbono (en ppm) de cierta placa de silicio en cada uno de dos días sucesivos. Los resultados fueron:

Día 1: 2.1321 2.1385 2.0985 2.0941 2.0680

Día 2: 2.0853 2.1476 2.0733 2.1194 2.0717

¿Puede concluir que la calibración del espectrómetro ha cambiado del primero al segundo días?

6. Dos pesas, cada una etiquetada con 100 g, se pesan varias veces en la misma balanza. Los resultados, en unidades de μg arriba de 100 g, fueron:

Primera pesa: 53 88 89 62 39 66

Segunda pesa: 23 39 28 2 49

En virtud de que se utilizó la misma balanza para ambas pesas, y de que éstas son similares, es razonable suponer que la varianza en los pesos no depende del objeto que se está pesando. ¿Se puede concluir que las pesas son diferentes?

7. El artículo “Mechanical Grading of Oak Timbers” (D. Kretschmann y D. Green, en *Journal of Materials in Civil Engineering*, 1999:91-97) presenta las mediciones finales de la fuerza compresiva, en MPa, para el roble verde de siete por nueve maderas de roble mezcladas de Virginia Occidental y Pensilvania. Para 11 elementos de madera de grado 1, la fuerza compresiva promedio fue de 22.1 con desviación estándar de 4.09. Para siete elementos de madera de grado 2, la fuerza compresiva promedio fue 20.4 con desviación estándar de 3.08. ¿Puede concluir que la media de la fuerza compresiva para la madera de grado 1 es mayor que la de grado 2?
8. Se comparan dos métodos para medir el calor molar de fusión del agua. Diez mediciones realizadas con el método A tienen una media de 6.02 kilojoules por mole (kJ/mol) con desviación estándar de 0.02 (kJ/mol). Cinco mediciones

realizadas con el método B tienen una media de 6.00 kJ/mol y desviación estándar de 0.01 kJ/mol. ¿Puede concluir que la media de las mediciones difiere entre los dos métodos?

9. Con referencia al ejercicio 8 de la sección 5.6, ¿puede concluir que las capacidades caloríficas del carbón de las dos minas son diferentes?
10. Considerando el ejercicio 2 de la sección 5.6, ¿puede concluir que más de 0.1 μg se absorbe entre 30 y 60 minutos después de la exposición?

11. Considerando el ejercicio 4 de la sección 5.6:

- a) ¿Puede concluir que la media de la frecuencia de mutación para hombres de 60 años de edad es mayor que para hombres de 20?
- b) ¿Puede concluir que la media de la frecuencia de mutación para hombres de 60 años de edad es mayor a 25 secuencias por μg que para varones de 20?

12. Considerando el ejercicio 9 de la sección 5.6, ¿puede concluir que la media de la fuerza de ruptura es mayor para paños de hockey hechos con el compuesto B?

13. En relación con el ejercicio 10 de la sección 5.6, ¿puede concluir que la media del coeficiente de permeabilidad en 60°C es diferente que la de 61°C?

14. El siguiente resultado de MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales.

Two-sample T for X vs Y

	N	Mean	StDev	SE Mean
X	10	39.31	8.71	2.8
Y	10	29.12	4.79	1.5

Difference = $\mu(X) - \mu(Y)$

Estimate for difference: 10.1974

95% lower bound for difference: 4.6333

T-Test of difference = 0 (vs >): T-Value = 3.25 P-Value = 0.003 DF = 13

- a) ¿Esta es una prueba de una o de dos colas?
- b) ¿Cuál es la hipótesis nula?
- c) ¿Puede H_0 ser rechazada a un nivel de 1%? ¿Cómo justifica su respuesta?

15. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales. Faltan algunos de los números. Complete los números del (a) al (d).

Two-sample T for X vs Y				
	N	Mean	StDev	SE Mean
X	6	1.755	0.482	(a)
Y	13	3.239	(b)	0.094
Difference = mu (X) - mu (Y)				
Estimate for difference: (c)				
95% CI for difference: (-1.99996, -0.96791)				
T-Test of difference = 0 (vs not =): T-Value = (d) P-Value = 0.000 DF = 7				

6.8 Pruebas de hipótesis con datos apareados

En la sección 5.7 se observó que algunas veces es mejor diseñar un experimento de dos muestras, por lo que cada elemento en una muestra está apareado con un elemento en la otra. En esta sección se presenta un método para las pruebas de hipótesis que implica la diferencia entre dos medias poblacionales con base en datos apareados. Se inicia con un ejemplo.

Las emisiones de materia partícula (MP) de los automóviles son un problema ambiental serio. Se eligieron aleatoriamente ocho vehículos de una flota, y se midieron sus emisiones durante su recorrido en autopista y en condiciones de arranque y frenado; además, se calcularon las diferencias en ambas situaciones. Los resultados, en miligramos de partículas por galón de combustible, fueron los siguientes:

	Vehículo							
	1	2	3	4	5	6	7	8
Arranque y frenado	1 500	870	1 120	1 250	3 460	1 110	1 120	880
Recorrido en autopista	941	456	893	1 060	3 107	1 339	1 346	644
Diferencia	559	414	227	190	353	-229	-226	236

¿Se puede concluir que la media del nivel de emisiones es menor para el recorrido en autopista que para el arranque y frenado?

La idea básica detrás de la construcción de la prueba de hipótesis en este ejemplo es la misma que la idea detrás de la construcción de los intervalos de confianza para los datos apareados de la sección 5.7. Se trata de la colección de diferencias como una sola muestra aleatoria de una población de diferencias. La media poblacional se denota con μ_D y la desviación estándar con σ_D . Hay solamente ocho diferencias, de ellas una es muestra pequeña. Si se supone que la población de diferencias es aproximadamente normal, se puede utilizar la prueba t de Student, como se presentó en la sección 6.4.

El valor observado de la media muestral de diferencias es $\bar{D} = 190.5$. La desviación estándar muestral es $s_D = 284.1$. Las hipótesis nula y alternativa son

$$H_0: \mu_D \leq 0 \quad \text{contra} \quad H_1: \mu_D > 0$$

El estadístico de prueba es

$$t = \frac{\bar{D} - 0}{s_D/\sqrt{n}} = \frac{190.5 - 0}{284.1/\sqrt{8}} = 1.897$$

La distribución nula del estadístico de prueba es la t de Student con siete grados de libertad. La figura 6.16 presenta la distribución nula e indica la ubicación del estadístico de prueba. Ésta es una prueba de una cola. La tabla t indica que 5% del área en la cola corresponde a un valor t de 1.895, muy cerca del valor observado de 1.897. El P -valor es aproximadamente 0.05. El siguiente resultado de computadora (de MINITAB) presenta este resultado.

Paired T-Test and CI: StopGo, Highway

Paired T for StopGo - Highway

	N	Mean	StDev	SE Mean
StopGo	8	1413.75	850.780	300.796
Highway	8	1223.25	820.850	290.214
Difference	8	190.50	284.104	100.446

95% lower bound for mean difference: 0.197215

T-Test of mean difference = 0 (vs > 0):

T-Value = 1.90 P-Value = 0.050

Observe que el límite inferior de 95% es apenas consistente con la hipótesis alternativa. Esto último indica que el P -valor sólo es apenas menor que 0.05 (aunque está dado por 0.050 con dos dígitos significativos).

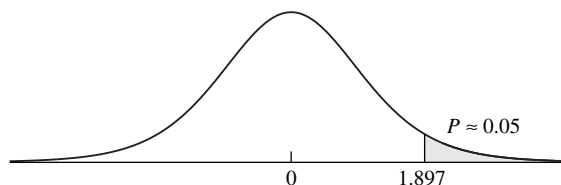


FIGURA 6.16 La distribución nula de $t = (\bar{D} - 0)/(s_D/\sqrt{8})$ es t_7 . El valor observado de t , correspondiente a los valores observados $\bar{D} = 190.5$ y $s_p = 284.1$, es 1.897. Si H_0 es verdadera, la probabilidad que t tome un valor tan extremo o más extremo que el observado está muy cercano a 0.05.

Resumen

Sea $(X_1, Y_1), \dots, (X_n, Y_n)$ una muestra de pares ordenados cuyas diferencias $D_1, \dots, D_n$ son muestra de una población *normal* con media μ_D .

Para probar una hipótesis nula de la forma $H_0: \mu_D \leq \mu_0$, $H_0: \mu_D \geq \mu_0$, o $H_0: \mu_D = \mu_0$:

■ Se calcula el estadístico de prueba $t = \frac{\bar{D} - \mu_0}{s_D/\sqrt{n}}$.

■ Se calcula el *P*-valor. Éste es un área debajo la curva *t* de Student con $n - 1$ grados de libertad, que depende de la hipótesis alternativa de la siguiente manera:

Hipótesis alternativa

$$H_1: \mu_D > \mu_0$$

$$H_1: \mu_D < \mu_0$$

$$H_1: \mu_D \neq \mu_0$$

P-valor

Área a la derecha de t

Área a la izquierda de t

Suma de las áreas en las colas correspondientes a t y $-t$

■ Si la muestra es grande, la D_i necesaria no está normalmente distribuida, el

estadístico de prueba es $z = \frac{\bar{D} - \mu_0}{s_D/\sqrt{n}}$, y se debe realizar la prueba *z*.

Ejercicios para la sección 6.8

- Los músculos se flexionan cuando se estimulan a través de impulsos eléctricos en puntos motores (puntos en el músculo) o en nervios. El artículo "Force Assessment of the Stimulated Arm Flexors: Quantification of Contractile Properties" (J. Hong y P. Iaizzo, en *Journal of Medical Engineering and Technology*, 2002:28-35) informa acerca de un estudio en el que ambos métodos se aplicaban a las regiones superiores del brazo de cada una de las partes. Se midió el tiempo de latencia (tiempo entre el estímulo y la contracción) (en milésimas de segundo) para cada parte. Los resultados para siete partes se presentan en la siguiente tabla (se ha eliminado un dato atípico).
- La maniobra de Valsalva implica soplar en un tubo cerrado con el fin de crear presión en las vías respiratorias. La impedancia cardiográfica se utiliza durante esta maniobra para evaluar la función cardíaca. El artículo "Impedance Cardiographic Measurement of the Physiological Response to the Valsalva Manoeuvre" (R. Patterson y J. Zhang, en *Medical and Biological Engineering and Computing*, 2003:40-43) presenta un estudio en el cual se midió el cociente de impedancia para cada una de las 11 partes, tanto en posición de pie como acostada. Los resultados de la presión de una vía aérea de 10 mmHg se presentan en la siguiente tabla.

	Parte						
	1	2	3	4	5	6	7
Nervio	59	57	58	38	53	47	51
Punto motriz	56	52	56	32	47	42	48
Diferencia	3	5	2	6	6	5	3

Parte	De pie	Acostada	Diferencia
1	1.45	0.98	0.47
2	1.71	1.42	0.29
3	1.81	0.70	1.11
4	1.01	1.10	-0.09
5	0.96	0.78	0.18
6	0.83	0.54	0.29
7	1.23	1.34	-0.11
8	1.00	0.72	0.28
9	0.80	0.75	0.05
10	1.03	0.82	0.21
11	1.39	0.60	0.79

¿Puede concluir que hay diferencia entre la media del coeficiente de la impedancia medida en las posiciones de pie y sentada?

3. Un proceso de grabado en seco se utiliza para grabar dióxido de silicio (SiO_2) fuera de placas de silicio. Un ingeniero desea estudiar la uniformidad del grabado a través de la superficie de la placa. Se muestrea un total de diez placas después del grabado, y se miden las tasas de grabado (en $\text{\AA}/\text{minuto}$) en dos sitios diferentes, uno cerca del centro de la placa y otro casi en el borde. Los resultados se presentan en la tabla siguiente.

Placa	Centro	Borde
1	586	582
2	568	569
3	587	587
4	550	543
5	543	540
6	552	548
7	562	563
8	577	572
9	558	559
10	571	566

¿Puede concluir que las tasas de grabado difieren entre el centro y el borde?

4. Se comparan dos microprocesadores en una muestra de seis códigos de puntos de referencia para determinar si hay una diferencia en la rapidez. Los tiempos (en segundos) utilizados para cada procesador en cada código están dados en la tabla siguiente.

	Código					
	1	2	3	4	5	6
Procesador A	27.2	18.1	27.2	19.7	24.5	22.1
Procesador B	24.1	19.3	26.8	20.1	27.6	29.8

¿Puede concluir que las medias de la rapidez de ambos procesadores son diferentes?

5. Se midió la fuerza compresiva, en kilopascales, para cada uno de cinco bloques de concreto, tanto tres como seis días después de verterlos. En la siguiente tabla se presentan los datos.

	Bloque				
	1	2	3	4	5
Después de tres días	1 341	1 316	1 352	1 355	1 327
Después de seis días	1 376	1 373	1 366	1 384	1 358

¿Puede concluir que la media de la fuerza después de seis días es mayor que la media de la fuerza después de tres días?

6. Considerando el ejercicio 3 de la sección 5.7, ¿puede concluir que la balanza 2 registra, en promedio, más peso que la balanza 1?

7. Teniendo en cuenta el ejercicio 4 de la sección 5.7, ¿puede concluir que hay una diferencia en la tensión entre los pernos 1 y 8?

8. Considerando el ejercicio 7 de la sección 5.7:

a) ¿Puede concluir que la media de los tiempos de vida de los patines de los frenos traseros es mayor que la de los delanteros?

b) ¿Puede concluir que la media de los tiempos de vida de los patines de los frenos traseros es mayor que la de los frenos delanteros en más de diez mil millas?

9. La directiva de una compañía de taxis está tratando de decidir si debe cambiar de neumáticos normales a neumáticos radiales para mejorar el ahorro de combustible. Se equiparon cada uno de los diez taxis con uno de los dos tipos de

neumáticos y se condujeron en una trayectoria de prueba. Sin cambiar de conductores, se seleccionó el tipo de neumáticos y se repitió la trayectoria de prueba. El ahorro de combustible (en milla/galón) para los diez automóviles es:

Automóvil	Radial	Normal
1	32.1	27.1
2	36.1	31.5
3	32.3	30.4
4	29.5	26.9
5	34.3	29.9
6	31.9	28.7
7	33.4	30.2
8	34.6	31.8
9	35.2	33.6
10	32.7	29.9

- Debido a que el cambio de neumáticos en la flota de taxis es caro, la directiva no quiere cambiar a menos que una prueba de hipótesis proporcione evidencias de que mejorará el millaje. Establezca las hipótesis nula y alternativa adecuadas, y encuentre el P -valor.
- Un análisis costo-beneficio muestra que será provechoso cambiar a neumáticos radiales si la media de la mejora del millaje es mayor a dos millas/galón. Establezca las hipótesis nula y alternativa adecuadas, y determine el P -valor, para una prueba de hipótesis diseñada como base de la decisión de cambiar.

10. El siguiente resultado MINITAB presenta los resultados de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales.

Paired T for X - Y

	N	Mean	StDev	SE Mean
X	12	134.233	68.376	19.739
Y	12	100.601	94.583	27.304
Difference	12	33.6316	59.5113	17.1794

95% lower bound for mean difference: 2.7793

T-Test of mean difference = 0 (vs > 0): T-Value = 1.96 P-Value = 0.038

- ¿Es ésta una prueba de una o de dos colas?
 - ¿Cuál es la hipótesis nula?
 - ¿Puede H_0 ser rechazada a un nivel de 1%? ¿Cómo justifica su respuesta?
 - Utilice el resultado y una tabla adecuada para calcular un intervalo de confianza de 98% para $\mu_X - \mu_Y$.
11. El siguiente resultado MINITAB presenta el desenlace de una prueba de hipótesis para la diferencia $\mu_X - \mu_Y$ entre dos medias poblacionales. Faltan algunos números. Complete los números del (a) al (d).

Paired T for X - Y

	N	Mean	StDev	SE Mean
X	7	12.4141	2.9235	(a)
Y	7	8.3476	(b)	1.0764
Difference	7	(c)	3.16758	1.19723

95% lower bound for mean difference: 1.74006

T-Test of mean difference = 0 (vs > 0): T-Value = (d) P-Value = 0.007

6.9 Pruebas de hipótesis con distribuciones libres

Las pruebas t de Student descritas en las secciones 6.4 y 6.7 requieren que las muestras provengan de poblaciones normales. Las pruebas con distribuciones libres obtienen su nombre del hecho de que no se requiere que las muestras provengan de ninguna distribución específica. Mientras que pruebas con distribuciones libres requieren de supuestos para su validez, éstos son un poco menos restrictivos que los necesarios para la prueba t . Las pruebas con distribuciones libres son en ocasiones llamadas **pruebas no paramétricas**.

En esta sección se analizan dos pruebas con distribuciones libres. La primera, **prueba del rango con signo de Wilcoxon**, es útil para una media poblacional, similar a la prueba t muestral analizada en la sección 6.4. La segunda, **prueba de la suma del rango de Wilcoxon**, o **prueba de Mann-Whitney**, es similar a la prueba t de dos muestras analizada en la sección 6.7.

La prueba del rango con signo de Wilcoxon

Esta prueba se muestra con un ejemplo. El contenido de níquel, en partes de miles por peso, se mide para seis soldaduras. Los resultados son 9.3, 0.9, 9.0, 21.7, 11.5 y 13.9. Sea μ la media del contenido de níquel para este tipo de soldadura. Se desea probar $H_0: \mu \geq 12$ contra $H_1: \mu < 12$. La prueba t de Student no es adecuada porque hay dos datos atípicos, 0.9 y 21.7, que indican que la población no es normal. La prueba del rango con signo de Wilcoxon se puede utilizar en esta situación. Esta prueba no exige que la población sea normal. Sin embargo, requiere que sea continua (en vez de discreta), y que la función de densidad de probabilidad sea simétrica. (La normal es un caso especial de una población simétrica continua.) La muestra dada proviene de una población continua, y la presencia de datos atípicos en cualquier lado hace razonable suponer que la población es casi simétrica. Por tanto, se procede como se muestra a continuación.

Bajo H_0 , la media poblacional es $\mu = 12$. Ya que se supone que la población es simétrica, también la mediana poblacional es 12. Para calcular el estadístico de la suma del rango se inicia restando 12 a cada observación de la muestra con el fin de obtener las diferencias. A la diferencia más cercana a 0, ignorando el signo, se le asigna un rango de 1. A la siguiente diferencia más cercana a 0, ignorando nuevamente el signo, se le asigna un rango de 2, y así sucesivamente. Por último, a los rangos que les corresponden diferencias negativas se les asignan signos negativos. La siguiente tabla muestra los resultados.

x	$x - 12$	Rango con signo
11.5	-0.5	-1
13.9	1.9	2
9.3	-2.7	-3
9.0	-3.0	-4
21.7	9.7	5
0.9	-11.1	-6

La suma de los rangos positivos se denota con S_+ , y la suma de los valores absolutos de los rangos negativos con S_- . Tanto S_+ como S_- se pueden utilizar como un estadístico de prueba; se utilizará S_+ . En este ejemplo $S_+ = 2 + 5 = 7$, y $S_- = 1 + 3 + 4 + 6 = 14$. Observe

que debido a que el tamaño de muestra es 6, necesariamente $S_+ + S_- = 1 + 2 + 3 + 4 + 5 + 6 = 21$. Para cualquier muestra, éste es el caso $S_+ + S_- = 1 + 2 + \dots + n = n(n+1)/2$. En algunos casos, donde hay muchos más rangos positivos que rangos negativos, es más fácil calcular primero a S_- al sumar los rangos negativos y después calcular $S_+ = n(n+1)/2 - S_-$.

Las figuras 6.17 y 6.18 muestran cómo se puede utilizar S_+ como un estadístico de prueba. En la figura 6.17, $\mu > 12$. Para esta distribución las diferencias positivas son más probables que las negativas y tienden a ser más grandes en magnitud. Por tanto, es probable que los rangos positivos sean más grandes tanto en número como en magnitud que los rangos negativos, por lo que es posible que S_+ invierta la situación. Aquí los rangos positivos son menos en número y menores en magnitud, por lo que S_+ es pequeño.



FIGURA 6.17 La mediana verdadera es mayor que 12. Es más probable que las observaciones de la muestra estén por arriba que por debajo de 12. Además, las observaciones arriba de 12 tenderán a tener diferencias mayores de 12 que las observaciones debajo de 12. Por tanto, S_+ probablemente es grande.

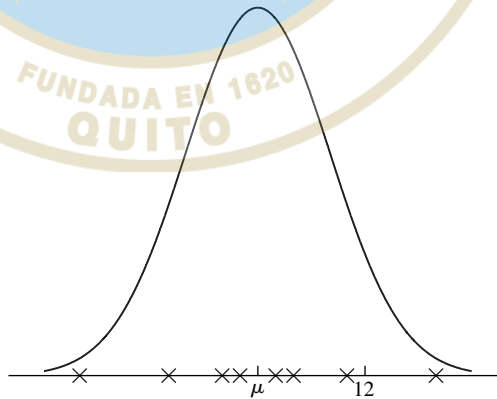


FIGURA 6.18 La mediana verdadera es menor de 12. Es más probable que las observaciones de la muestra estén por debajo de 12 que por arriba de 12. Además, las observaciones debajo de 12 tenderán a tener diferencias mayores de 12, que las observaciones arriba de 12. Por tanto, S_+ probablemente es pequeño.

Se observa que, en general, los valores grandes de S_+ proporcionan evidencia contra una hipótesis nula de la forma $H_0: \mu \leq \mu_0$, mientras que valores pequeños de S_+ proporcionarán evidencia en contra de una hipótesis nula de la forma $H_0: \mu \geq \mu_0$.

En este ejemplo, la hipótesis nula es $H_0: \mu \geq 12$, por eso un valor pequeño de S_+ proporcionará evidencia en contra de H_0 . Se observa que $S_+ = 7$. El P -valor es la probabilidad de observar un valor de S_+ que es menor o igual a 7 cuando H_0 es verdadera. La tabla A.4 (en el apéndice A) presenta ciertas probabilidades para la distribución nula de S_+ . Al consultar esta tabla con el tamaño de muestra $n = 6$, se encuentra la probabilidad de observar un valor de 4 o menor de 0.1094. La probabilidad de observar un valor de 7 o menor debe ser más grande que esto último, por lo que se concluye que $P > 0.1094$; por consiguiente, no se rechaza H_0 .

Ejemplo 6.15

En el ejemplo analizado previamente se midió el contenido de níquel para seis soldaduras de 9.3, 0.9, 9.0, 21.7, 11.5 y 13.9. Utilice estos datos para probar $H_0: \mu \leq 5$ contra $H_1: \mu > 5$.

Solución

La tabla de diferencias y los rangos con signos son:

x	$x - 5$	Rango con signo
9.0	4.0	1
0.9	-4.1	-2
9.3	4.3	3
11.5	6.5	4
13.9	8.9	5
21.7	16.7	6

El valor observado del estadístico de prueba es $S_+ = 19$. Ya que la hipótesis nula es de la forma $H_0: \mu \leq \mu_0$, valores grandes de S_+ proporcionan evidencia contra H_0 . Por tanto, el P -valor es el área en la cola derecha de la distribución nula, correspondiendo a valores igual o superiores a 19. Al consultar la tabla A.4 se encuentra que el P -valor es 0.0469.

Ejemplo 6.16

Utilice los datos del ejemplo 6.15 para probar $H_0: \mu = 16$ contra $H_1: \mu \neq 16$.

Solución

La tabla de diferencias y rangos con signos es:

x	$x - 16$	Rango con signo
13.9	-2.1	-1
11.5	-4.5	-2
21.7	5.7	3
9.3	-6.7	-4
9.0	-7.0	-5
0.9	-15.1	-6

Puesto que la hipótesis nula es de la forma $H_0: \mu = \mu_0$, ésta es una prueba de dos colas. El valor observado del estadístico de prueba es $S_+ = 3$. Al consultar la tabla A.4, se encuentra que el área en la cola izquierda, correspondiente a valores menores que o iguales a 3, es 0.0781. El P -valor es el doble de esta cantidad, ya que es la suma de áreas en dos colas iguales. Por tanto, el P -valor es $2(0.0781) = 0.1562$.

Empates

A veces dos o más de las cantidades a las cuales debe asignárseles rango tienen exactamente el mismo valor. Se dice que estas cantidades están empatadas. El método usual para tratar los empates es asignar a cada observación empatada el promedio de los rangos que habrían recibido si hubieran sido ligeramente diferentes. Por ejemplo, las cantidades 3, 4, 4, 5, 7 recibirían los rangos 1, 2.5, 2.5, 4, 5 y las cantidades 12, 15, 16, 16, 20 recibirían los rangos 1, 2, 4, 4, 6.

Diferencias de cero

Si la media bajo H_0 es μ_0 y una de las observaciones es igual a μ_0 , entonces su diferencia es 0, que no es positivo ni negativo. Una observación igual a μ_0 no puede recibir un rango con signo. El procedimiento adecuado consiste en eliminar estas observaciones de la muestra, y considerar que el tamaño de ésta se ha reducido como consecuencia del número de aquéllas. El ejemplo 6.17 sirve para ilustrar este punto.

Ejemplo

6.17

Utilice los datos del ejemplo 6.15 para probar $H_0: \mu = 9$ contra $H_0: \mu \neq 9$.

Solución

La tabla de diferencias y de rangos con signo es:

x	$x - 9$	Rango con signo
9.0	0.0	—
9.3	0.3	1
11.5	2.5	2
13.9	4.9	3
0.9	-8.1	-4
21.7	12.7	5

El valor del estadístico de prueba es $S_+ = 11$. El tamaño muestral para propósitos de la prueba es 5, ya que el valor 9.0 no tiene rango. Al ingresar a la tabla A.4 con tamaño muestral 5, se encuentra que $S_+ = 12$, el P -valor sería $2(0.1562) = 0.3124$. Se concluye que para $S_+ = 11$, $P > 0.3124$.

Aproximación para muestras grandes

Cuando el tamaño muestral n es grande, el estadístico de prueba S_+ está distribuido aproximadamente en forma normal. Una regla general es que la aproximación normal es buena si $n > 20$. Se puede demostrar mediante métodos avanzados que bajo H_0 S_+ tiene una media $n(n + 1)/4$ y varianza $n(n + 1)(2n + 1)/24$. La prueba del rango con signo de Wilcoxon se realiza cuando se calcula el puntaje z de S_+ , y después al usar la tabla normal para encontrar el P -valor. El puntaje z es

$$z = \frac{S_+ - n(n + 1)/4}{\sqrt{n(n + 1)(2n + 1)/24}}$$

El ejemplo 6.18 ilustra el método.

Ejemplo
6.18

El artículo “Exact Evaluation of Batch-Ordering Inventory Policies in Two-Echelon Supply Chains with Periodic Review” (G. Chacon, en *Operations Research*, 2001:79-98) presenta una evaluación de un reordenamiento de una política focal. Se estiman los costos para 32 escenarios. Sea μ la media del costo. Pruebe que $H_0: \mu \geq 70$ contra $H_1: \mu < 70$. Los datos, junto con las diferencias y los rangos con signo, se presentan en la tabla 6.1.

TABLA 6.1 Datos para el ejemplo 6.18

Rango			Rango			Rango		
x	$x - 70$	con signo	x	$x - 70$	con signo	x	$x - 70$	con signo
79.26	9.26	1	30.27	-39.73	-12	11.48	-58.52	-23
80.79	10.79	2	22.39	-47.61	-13	11.28	-58.72	-24
82.07	12.07	3	118.39	48.39	14	10.08	-59.92	-25
82.14	12.14	4	118.46	48.46	15	7.28	-62.72	-26
57.19	-12.81	-5	20.32	-49.68	-16	6.87	-63.13	-27
55.86	-14.14	-6	16.69	-53.31	-17	6.23	-63.77	-28
42.08	-27.92	-7	16.50	-53.50	-18	4.57	-65.43	-29
41.78	-28.22	-8	15.95	-54.05	-19	4.09	-65.91	-30
100.01	30.01	9	15.16	-54.84	-20	140.09	70.09	31
100.36	30.36	10	14.22	-55.78	-21	140.77	70.77	32
30.46	-39.54	-11	11.64	-58.36	-22			

Solución

El tamaño de muestra es $n = 32$, por lo que la media es $n(n + 1)/4 = 264$ y la varianza es $n(n + 1)(2n + 1)/24 = 2\,860$. La suma de los rangos positivos es $S_+ = 121$. Se calcula

$$z = \frac{121 - 264}{\sqrt{2\,860}} = -2.67$$

Debido a que la hipótesis nula es de la forma $H_0: \mu \geq \mu_0$, valores pequeños de S_+ proporcionan evidencias contra H_0 . Por tanto, el P -valor es el área debajo de la curva normal a la izquierda de $z = -2.67$. Esta área, y por consiguiente, el P -valor, es 0.0038.

La prueba de la suma del rango de Wilcoxon

La prueba de la suma del rango de Wilcoxon, también llamada de Mann-Whitney, es útil para probar la diferencia en las medias poblacionales en ciertos casos donde las poblaciones no son normales. Son necesarias dos suposiciones. Primero, las poblaciones deben ser continuas. Segundo, sus funciones de densidad de probabilidad deben ser idénticas en forma y tamaño; la única diferencia posible entre ellas será su posición. Para describir la prueba, sea $X_1, \dots, X_m$ una muestra aleatoria de una población y sea $Y_1, \dots, Y_n$ una muestra aleatoria de la otra. Se adopta la notación convencional que cuando los tamaños muestrales no son iguales, la muestra menor se denotará mediante $X_1, \dots, X_m$. Por tanto, los tamaños muestrales son m y n , con $m \leq n$. Las medias poblacionales se denotan con μ_X y μ_Y , respectivamente.

La prueba se realiza ordenando los $m + n$ valores obtenidos de la combinación de las dos muestras, y se asignan los rangos $1, 2, \dots, m + n$ a ellos. El estadístico de prueba, denotado con W , representa la suma de los rangos correspondientes a $X_1, \dots, X_m$. Debido a que las poblaciones son idénticas con la posible excepción de la posición, en consecuencia si $\mu_X < \mu_Y$, los valores en la muestra X tenderán a ser menores que aquellos en la muestra Y , por lo que la suma de los rangos W también tenderán a ser menores. Por un razonamiento similar, si $\mu_X > \mu_Y$, W tenderá a ser mayor. Se muestra la prueba en el ejemplo 6.19.

Ejemplo 6.19

Se miden resistencias, en $m\Omega$, para cinco alambres de un tipo y seis de otro. Los resultados son:

X:	36	28	29	20	38	
Y:	34	41	35	47	49	46

Utilice la prueba de la suma del rango de Wilcoxon para probar $H_0: \mu_X \geq \mu_Y$ contra $H_1: \mu_X < \mu_Y$.

Solución

Se ordenan los 11 valores y se les asignan rangos

Valor	Rango	Muestra	Valor	Rango	Muestra
20	1	X	38	7	X
28	2	X	41	8	Y
29	3	X	46	9	Y
34	4	Y	47	10	Y
35	5	Y	49	11	Y
36	6	X			

El estadístico de prueba W es la suma de los rangos que corresponden a los valores de X , por lo que $W = 1 + 2 + 3 + 6 + 7 = 19$. El P -valor se determina consultando la tabla A.5 (en el apéndice A). Se observa que valores pequeños de W proporcionan evidencia en contra de $H_0: \mu_X \geq \mu_Y$, por lo que el P -valor es el área en la cola izquierda de la distribución nula. Buscando en la tabla con $m = 5$ y $n = 6$ se encuentra que el área a la izquierda de $W = 19$ es 0.0260. Éste es el P -valor.

Aproximación para muestras grandes

Cuando ambos tamaños de muestra m y n son mayores que 8, se puede mostrar mediante métodos avanzados que la distribución nula del estadístico de prueba W es aproximadamente normal con media $m(m + n + 1)/2$ y la varianza $mn(m + n + 1)/12$. En estos casos la prueba se realiza al calcular el puntaje z de W , y después con la tabla normal para encontrar el P -valor. El puntaje z es

$$z = \frac{W - m(m + n + 1)/2}{\sqrt{mn(m + n + 1)/12}}$$

El ejemplo 6.20 ilustra el método.

Ejemplo

6.20

El artículo “Cost Analysis Between SABER and Design Bid Build Contracting Methods” (E. Henry y H. Brothers, en *Journal of Construction Engineering and Management*, 2001:359-366) presenta datos de costos de construcción para diez trabajos ofrecidos por el método tradicional (denotados por X) y 19 trabajos ofrecidos por un sistema experimental (denotados por Y). Los datos, en unidades de dólares por metro cuadrado, y sus rangos se presentan en la tabla 6.2. Pruebe que $H_0: \mu_X \leq \mu_Y$ contra $H_1: \mu_X > \mu_Y$.

TABLA 6.2 Datos para el ejemplo 6.20

Valor	Rango	Muestra	Valor	Rango	Muestra
57	1	X	613	16	X
95	2	Y	622	17	Y
101	3	Y	708	18	X
118	4	Y	726	19	Y
149	5	Y	843	20	Y
196	6	Y	908	21	Y
200	7	Y	926	22	X
233	8	Y	943	23	Y
243	9	Y	1 048	24	Y
341	10	Y	1 165	25	X
419	11	Y	1 293	26	X
457	12	X	1 593	27	X
584	13	X	1 952	28	X
592	14	Y	2 424	29	Y
594	15	Y			

Solución

La suma de los rangos X es $W = 1 + 12 + 13 + 16 + 18 + 22 + 25 + 26 + 27 + 28 = 188$. Los tamaños de muestra son $m = 10$ y $n = 19$. Se utiliza la aproximación normal y se calcula

$$\begin{aligned} z &= \frac{188 - 10(10 + 19 + 1)/2}{\sqrt{10(19)(10 + 19 + 1)/12}} \\ &= 1.74 \end{aligned}$$

Valores grandes de W proporcionan evidencias contra la hipótesis nula. Por tanto, el P -valor es el área debajo de la curva normal a la derecha de $z = 1.74$. De la tabla z se encuentra que el P -valor es 0.0409.

Los métodos con distribuciones libres no están libres de supuestos

Se debe señalar que los métodos con distribución libre presentados aquí requieren ciertos supuestos para su validez. Desafortunadamente, esto último a veces se olvida en la práctica. Es tentador usar un procedimiento con distribución libre automáticamente en cualquier situación en la que la prueba t de Student no parece estar justificada, y suponer que los resultados serán válidos siempre que éste no sea el caso. Las suposiciones necesarias de simetría para la prueba del rango con signo y de formas idénticas y extensiones para la prueba de la suma del rango son, en realidad, algo restrictivas. Mientras que estas pruebas funcionan razonablemente bien en violaciones moderadas de estos supuestos, no son aplicables de modo universal.

Ejercicios para la sección 6.9

- El artículo “Wastewater Treatment Sludge as a Raw Material for the Production of *Bacillus thuringiensis* Based Biopesticides” (M. Tirado Montiel, R. Tyagi y J. Valero, en *Water Research* 2001:3807-3816) presenta mediciones de sólidos totales, en g/l, para siete muestras de lodo. Los resultados (redondeados al gramo más cercano) son 20, 5, 25, 43, 24, 21 y 32. Suponga que la distribución de sólidos totales es casi simétrica.
 - ¿Puede concluir que la media de la concentración de sólidos totales es mayor que 14 g/l? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - ¿Puede concluir que la media de la concentración de sólidos totales es menor que 30 g/l? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - Un ingeniero ambiental afirma que la media de la concentración de sólidos totales es igual a 18 g/l. ¿Los datos refutan esta afirmación?
- Se mide el espesor de ocho almohadillas diseñadas para emplearlas en monturas de motores de aviones. Los resultados, en mm, son 41.83, 41.01, 42.68, 41.37, 41.83, 40.50, 41.70 y 41.42. Suponga que los espesores son una muestra de distribución casi simétrica.
 - ¿Puede concluir que la media del espesor es mayor a 41 mm? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - ¿Puede concluir que la media del espesor es menor a 41.8 mm? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - El espesor objetivo es 42 mm, ¿puede concluir que la media del espesor difiere del valor objetivo? Calcule el estadístico de prueba adecuado y encuentre el P -valor.
- El artículo “Reaction Modeling and Optimization Using Neural Networks and Genetic Algorithms: Case Study Involving TS-I-Catalyzed Hydroxylation of Benzene” (S. Nandi, P. Mukherjee y cols., en *Industrial and Engineering Chemistry Research*, 2002:2159-2169) presenta conversiones de benceno (en porcentaje molar) para 24 reacciones de hidroxilación de benceno diferentes. Los resultados son

52.3	41.1	28.8	67.8	78.6	72.3	9.1	19.0
30.3	41.0	63.0	80.8	26.8	37.3	38.1	33.6
14.3	30.1	33.4	36.2	34.6	40.0	81.2	59.4.

 - ¿Puede concluir que la media de conversión es menor de 45? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - ¿Puede concluir que la media de conversión es mayor de 30? Calcule el estadístico de prueba adecuado y determine el P -valor.
 - ¿Puede concluir que la media de conversión es diferente de 55? Calcule el estadístico de prueba adecuado y determine el P -valor.

4. El artículo “Abyssal Peridotites > 3 800 Ma from Southern West Greenland: Field Relationships, Petrography, Geochronology, Whole-Rock and Mineral Chemistry of Dunite and Harzburgite Inclusions in the Itsaq Gneiss Complex” (C. Friend, V. Bennett y A. Nutman, en *Contributions to Mineral Petrology*, 2002:71-92) presenta concentraciones de dióxido de silicio (SiO_2) (en porcentaje de peso) para diez peridotitas. Los resultados son:

40.57 41.48 40.76 39.68 43.68 43.53
43.76 44.86 43.06 46.14.

- a) ¿Puede concluir que la media de la concentración es mayor a 41? Calcule el estadístico de prueba adecuado y determine el P -valor.
- b) ¿Puede concluir que la media de la concentración es menor a 43? Calcule el estadístico de prueba adecuado y determine el P -valor.
- c) ¿Puede concluir que la media de la concentración difiere de 44? Calcule el estadístico de prueba adecuado y determine el P -valor.
5. Este ejercicio muestra que la prueba del rango con signo se puede utilizar con datos apareados. Se comparan dos calibradores que miden la profundidad de huella de los neumáticos. Se miden diez diferentes puntos en un neumático una vez con cada calibrador. Los resultados, en mm, se presentan en la tabla siguiente.

Punto	Calibrador 1	Calibrador 2	Diferencia
1	3.95	3.80	0.15
2	3.23	3.30	-0.07
3	3.60	3.59	0.01
4	3.48	3.61	-0.13
5	3.89	3.88	0.01
6	3.76	3.73	0.03
7	3.45	3.56	-0.11
8	3.01	3.02	-0.01
9	3.82	3.77	0.05
10	3.44	3.49	-0.05

Suponga que las diferencias son una muestra de una población casi simétrica con media μ . Utilice la prueba del rango con signo de Wilcoxon para probar $H_0: \mu = 0$ contra $H_1: \mu \neq 0$.

6. El artículo “ n -Nonane Hydroconversion on Ni and Pt Containing HMFI, HMOR and HBEA” (G. Kinger y H. Vinek, en *Applied Catalysis A: General*, 2002:139-149) presenta tasas de hidroconversión (en $\mu\text{mol/g}\cdot\text{s}$) de n nonanos en ambos catalizadores HMFI y HBEA. Los resultados son:

HMFI: 0.43 0.93 1.91 2.56 3.72 6.19 11.00
HBEA: 0.73 1.12 1.24 2.93

¿Puede concluir que la media de la tasa es diferente entre los dos catalizadores?

7. Se está comparando un nuevo tratamiento posquirúrgico con un tratamiento usual. Siete personas reciben el nuevo tratamiento, mientras que otras siete (los testigos) reciben el tratamiento usual. Los tiempos de recuperación, en días, fueron:

Tratamiento (X): 12 13 15 19 20 21 27
Testigo (Y): 18 23 24 30 32 35 40

¿Puede concluir que la media de la tasa es diferente entre el tratamiento y el testigo?

8. En un experimento para determinar el efecto del tiempo de cura en la fuerza compresiva de bloques de concreto, dos muestras de 15 bloques estaban preparadas en forma idéntica excepto por el tiempo de cura. Los bloques en una muestra se curaron durante dos días, mientras que los bloques en la otra lo fueron durante seis días. Las fuerzas compresivas de los bloques, en MPa, son las siguientes:

Curados dos días (X): 1 326 1 302 1 314 1 270
1 287 1 328 1 318 1 296
1 306 1 329 1 255 1 310
1 255 1 291 1 280

Curados seis días (Y): 1 387 1 301 1 376 1 397
1 399 1 378 1 343 1 349
1 321 1 364 1 332 1 396
1 372 1 341 1 374

¿Puede concluir que la media de la fuerza es mayor para los bloques curados durante seis días?

9. En una comparación de la eficacia del material aprendido en clases tradicionales, 12 estudiantes tomaron un curso de administración de empresas en línea, mientras que 14 lo tomaron en el salón de clases. Las calificaciones del examen final fueron:

En línea: 66 75 85 64 88 77 74
 91 72 69 77 83

En el salón de clases: 80 83 64 81 75 80 86
 81 51 64 59 85 74 77

¿Puede concluir que la media de la calificación difiere entre los dos tipos de cursos?

10. Una mujer que se ha cambiado a una nueva casa está tratando de determinar cuál de dos rutas a su trabajo tiene el tiempo de conducción promedio más corto. Los tiempos en minutos para seis viajes en la ruta A y cinco en la ruta B son:

A: 16.0 15.7 16.4 15.9 16.2 16.3
B: 17.2 16.9 16.1 19.8 16.7

¿Puede concluir que la media del tiempo es menor para la ruta A?

6.10 La prueba Ji cuadrada

En la sección 4.1 se estudió el ensayo de Bernoulli, que representa un proceso con dos resultados posibles, indicados como “éxito” y “fracaso”. Si se realiza un número de ensayos de Bernoulli, y se cuenta el número de éxitos, se puede probar la hipótesis nula de que la probabilidad p de éxito sea igual a un valor preespecificado p_0 . Esto último se trató en la sección 6.3. Si se hacen dos conjuntos de ensayos de Bernoulli, con probabilidad de éxito p_1 para el primer conjunto y p_2 para el segundo, se puede probar la hipótesis nula que $p_1 = p_2$. Lo anterior se trató en la sección 6.6.

Una generalización del ensayo de Bernoulli es el **ensayo multinomial** (véase la sección 4.4): un experimento que puede dar como resultado cualquiera de los k resultados, donde $k \geq 2$. Las probabilidades de los k resultados se denotan mediante $p_1, \dots, p_k$. Por ejemplo, el lanzamiento de un dado constituye un ensayo multinomial con seis resultados 1, 2, 3, 4, 5, 6; y probabilidades $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 1/6$. En esta sección se generalizan las pruebas para una probabilidad de Bernoulli a los ensayos multinomiales. Se inicia con un ejemplo en el que se prueba la hipótesis nula de que las probabilidades multinomiales $p_1, p_2, \dots, p_k$ son iguales a un conjunto preespecificado de valores, $p_{01}, p_{02}, \dots, p_{0k}$, por lo que la hipótesis nula tiene la forma $H_0: p_1 = p_{01}, p_2 = p_{02}, \dots, p_k = p_{0k}$.

Imagine que un jugador quiere probar un dado para ver si se desvía de su honestidad. Sea p_i la probabilidad de que salga el número i . La hipótesis nula establecerá que el dado es honesto, por lo que las probabilidades especificadas bajo la hipótesis nula son $p_{01} = p_{02} = p_{03} = p_{04} = p_{05} = p_{06} = 1/6$. La hipótesis nula es $H_0: p_1 = \dots = p_6 = 1/6$.

El jugador lanza el dado 600 veces y obtiene los resultados que se muestran en la tabla 6.3, en la columna “Observados”. Los resultados obtenidos se denominan **valores observados**. Para probar la hipótesis nula se construye una segunda columna, llamada “Esperados”. Esta columna contiene los **valores esperados**. El valor esperado para un resultado dado es la media del número de ensayos que resultarían en ese resultado si H_0 fuera verdadera. Para calcular los valores esperados, sea N el número total de ensayos. (En el ejemplo del dado, $N = 600$.) Cuando H_0 es verdadera, la probabilidad de que en un ensayo salga el resultado i es p_{0i} , por lo que el número esperado de ensayos en los que sale el resultado i es Np_{0i} . En el ejemplo del dado, el número esperado de ensayos para cada resultado es 100.

TABLA 6.3 Valores observados y esperados para 600 lanzamientos de un dado

Categoría	Observados	Esperados
1	115	100
2	97	100
3	91	100
4	101	100
5	110	100
6	86	100
Total	600	600

La idea detrás de la prueba de hipótesis es que si H_0 es verdadera, entonces los valores observados y esperados están probablemente cercanos entre sí. Por tanto, se construirá un estadístico de prueba que mida la cercanía entre los valores observados y los esperados. El estadístico se llama **estadístico Ji cuadrada**. Para definirlo, sea k el número de resultados ($k = 6$ en el ejemplo del dado), y sean O_i y E_i los números observados y esperados de los ensayos, respectivamente, que salen en el resultado i . El estadístico Ji cuadrada es

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (6.6)$$

Entre mayor sea el valor χ^2 , más fuerte es la evidencia contra H_0 . Para determinar el P -valor para la prueba se debe conocer la distribución nula de este estadístico de prueba. En general, no se puede determinar exactamente la distribución nula. Sin embargo, cuando los valores esperados son todos grandes, una buena aproximación está disponible. Ésta se llama **distribución Ji cuadrada** con $k - 1$ grados de libertad, que se denota mediante χ_{k-1}^2 . Observe que el número de grados de libertad es uno menos que el número de categorías. El uso de la distribución Ji cuadrada es adecuado siempre que todos los valores esperados sean mayores que o iguales a 5.

En el apéndice A se proporciona una tabla para la distribución Ji cuadrada (tabla A.6). La tabla proporciona valores para ciertos cuantiles, o puntos de porcentajes superiores, para muchas elecciones de grados de libertad. Como ejemplo, la figura 6.19 presenta la función de densidad de probabilidad de la distribución χ_{10}^2 . El 5% superior de la distribución está sombreado. Para encontrar el punto 5% superior en la tabla, busque abajo de $\alpha = 0.05$ y en grados de libertad $\nu = 10$. El valor es 18.307.

Ahora se calcula el valor de la estadística Ji cuadrada para los datos en la tabla 6.3. El número de grados de libertad es 5 (uno menos que el número de resultados). Utilizando la ecuación (6.6), el valor del estadístico es

$$\begin{aligned} \chi^2 &= \frac{(115 - 100)^2}{100} + \cdots + \frac{(86 - 100)^2}{100} \\ &= 2.25 + \cdots + 1.96 \\ &= 6.12 \end{aligned}$$

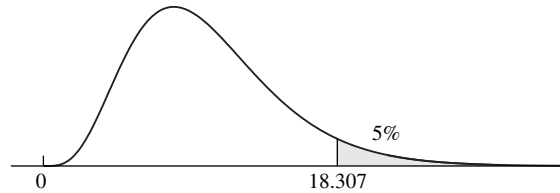


FIGURA 6.19 Función de densidad de probabilidad de la distribución χ^2_{10} . El punto superior de 5% es 18.307. [Véase la tabla Ji cuadrada (Tabla A.6) en el Apéndice A.]

Determinar el P -valor para el estadístico de prueba requiere observar que todos los valores esperados son iguales o mayores que 5, por lo que es adecuado utilizar la distribución Ji cuadrada. Se consulta la tabla Ji cuadrada con cinco grados de libertad. El punto 10% superior es 9.236. Se concluye que $P > 0.10$. (Véase la figura 6.20.) No hay evidencias que sugerir si el dado no es honesto.

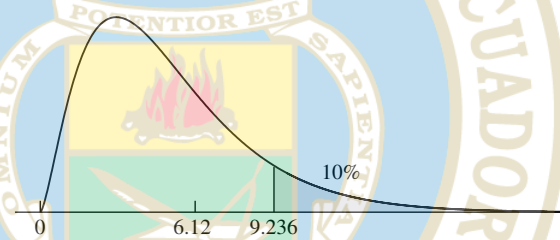


FIGURA 6.20 Función de densidad de probabilidad de la distribución χ^2_5 . El valor observado del estadístico de prueba es 6.12. El punto 10% superior es 9.236. Por tanto, el P -valor es mayor que 0.10.

La prueba Ji cuadrada para la homogeneidad

En el ejemplo anterior se probó la hipótesis nula de que las probabilidades de los resultados para un ensayo multinomial eran iguales al conjunto de valores preespecificados. A veces algunos ensayos multinomiales están dirigidos, cada uno con el mismo conjunto de resultados posibles. La hipótesis nula es que las probabilidades de los resultados son las mismas para cada experimento. Se presenta un ejemplo.

Cuatro máquinas fabrican alfileres cilíndricos de acero. Los alfileres están sujetos a una especificación de diámetro. Un alfiler podría satisfacer la especificación, o ser demasiado angosto o demasiado grueso. Se muestrean alfileres de cada máquina y se cuenta su número en cada categoría. La tabla 6.4 presenta los resultados.

La tabla 6.4 representa un ejemplo de una **tabla de contingencia**. Cada renglón especifica una categoría respecto de un criterio (máquina, en este caso), y cada columna especifica una categoría respecto de otro (el espesor, en este caso). Cada intersección de renglón y columna se llama **celda**, por lo que hay 12 celdas en la tabla 6.4.

TABLA 6.4 Números de alfileres observados en diferentes categorías, considerando una especificación del diámetro

	Muy angosto	Está bien	Muy grueso	Total
Máquina 1	10	102	8	120
Máquina 2	34	161	5	200
Máquina 3	12	79	9	100
Máquina 4	10	60	10	80
Total	66	402	32	500

El número en la celda en la intersección del renglón i con la columna j constituye el número de ensayos cuyo resultado estaba en la categoría del renglón i y en la categoría de columna j . Este número se llama **valor observado** para la celda ij . Observe que se han incluido los totales de los valores observados para cada renglón y columna. Éstos se llaman **marginales totales**.

La hipótesis nula es que la proporción de alfileres muy angostos, está bien o muy gruesos es la misma para todas las máquinas. De manera general, la hipótesis nula dice que no importa qué renglón se elija, las probabilidades de los resultados asociadas con las columnas son las mismas. Se desarrollará cierta notación para expresar H_0 y definir el estadístico de prueba.

Sea I el número de renglones en la tabla, y J el de columnas. Sea p_{ij} la probabilidad de que el resultado de un ensayo esté en la columna j en virtud de que está en el renglón i . Entonces la hipótesis nula es

$$H_0: \text{Para cada columna } j, p_{1j} = \cdots = p_{Ij} \quad (6.7)$$

Sea O_{ij} el valor observado en la celda ij . Sea O_i la suma de los valores observados en el renglón i , sea O_j la suma de los valores observados en la columna j , y sea $O_{..}$ la suma de los valores observados en todas las celdas (véase la tabla 6.5).

TABLA 6.5 Notación para los valores observados

	Columna 1	Columna 2	...	Columna J	Total
Renglón 1	O_{11}	O_{12}	...	O_{1J}	$O_{1.}$
Renglón 2	O_{21}	O_{22}	...	O_{2J}	$O_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Renglón I	O_{I1}	O_{I2}	...	O_{IJ}	$O_{I.}$
Total	$O_{.1}$	$O_{.2}$	...	$O_{.J}$	$O_{..}$

Definir un estadístico de prueba requiere calcular la tabla. Bajo H_0 la probabilidad de que el resultado de un ensayo esté en la columna j es la misma para cada renglón i . La mejor estimación de esta probabilidad es la proporción de ensayos cuyo resultado está en la colum-

na j . Esta proporción es $O_{j}/O_{..}$. Se necesita calcular el *número* esperado de ensayos cuyo resultado está en la celda ij . Se denota este valor esperado por E_{ij} . Es igual a la proporción de ensayos cuyo resultado está en la columna j , multiplicado por el número O_{i} de ensayos en el renglón i . Esto es,

$$E_{ij} = \frac{O_{i.} O_{.j}}{O_{..}} \quad (6.8)$$

El estadístico de prueba está basado en las diferencias entre los valores observados y esperados:

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (6.9)$$

Bajo H_0 este estadístico de prueba tiene una distribución Ji cuadrada con $(I - 1)(J - 1)$ grados de libertad. El uso de la distribución Ji cuadrada es adecuado siempre que los valores esperados sean todos iguales a o mayores que 5.

Ejemplo

6.21

Use los datos de la tabla 6.4 para **probar** la hipótesis nula de que las proporciones de alfileres que son muy angostos, están bien o muy gruesos son las mismas para todas las máquinas.

Solución

Se inicia mediante la ecuación (6.8) para calcular los valores esperados E_{ij} . Se muestran los cálculos de E_{11} y E_{23} en detalle:

$$E_{11} = \frac{(120)(66)}{500} = 15.84$$

$$E_{23} = \frac{(200)(32)}{500} = 12.80$$

La tabla completa de los valores esperados es:

Valores esperados para la tabla 6.4

	Muy angosto	Está bien	Muy grueso	Total
Máquina 1	15.84	96.48	7.68	120.00
Máquina 2	26.40	160.80	12.80	200.00
Máquina 3	13.20	80.40	6.40	100.00
Máquina 4	10.56	64.32	5.12	80.00
Total	66.00	402.00	32.00	500.00

Se observa que todos los valores esperados son mayores que 5. Por tanto, la prueba Ji cuadrada es adecuada. Se utiliza la ecuación (6.9) para calcular el valor del estadístico Ji cuadrada:

$$\begin{aligned}\chi^2 &= \frac{(10 - 15.84)^2}{15.84} + \cdots + \frac{(10 - 5.12)^2}{5.12} \\ &= \frac{34.1056}{15.84} + \cdots + \frac{23.8144}{5.12} \\ &= 15.5844\end{aligned}$$

Puesto que hay cuatro renglones y tres columnas, el número de grados de libertad es $(4 - 1)(3 - 1) = 6$. Para obtener el P -valor se consulta la tabla Ji cuadrada (tabla A.6). Buscando abajo de seis grados de libertad se encuentra que el punto 2.5% superior es 14.449, y el punto 1% superior es 16.812. Por tanto, $0.01 < P < 0.025$. Es razonable concluir que las máquinas difieren en las proporciones de alfileres que son muy angostos, están bien o muy gruesos.

Se observa que el renglón observado y la columna total son idénticos al renglón esperado y la columna total. Éste es siempre el caso.

El siguiente resultado de computadora (de MINITAB) presenta los resultados de esta prueba de hipótesis.

Chi-Square Test: Thin, OK, Thick

Expected counts are printed below observed counts

Chi-Square contributions are printed below expected counts

	Thin	OK	Thick	Total
1	10 15.84 2.153	102 96.48 0.316	8 7.68 0.013	120
2	34 26.40 2.188	161 160.80 0.000	5 12.80 4.753	200
3	12 13.20 0.109	79 80.40 0.024	9 6.40 1.056	100
4	10 10.56 0.030	60 64.32 0.290	10 5.12 4.651	80
Total	66	402	32	500

Chi-Sq = 15.584 DF = 6, P-Value = 0.016

En el resultado, cada celda (la intersección de renglón y columna) contiene tres números. El número superior es el valor observado, el de enmedio, el valor esperado, y el de abajo, la contribución $(O_{ij} - E_{ij})^2/E_{ij}$ hechos por el estadístico Ji cuadrada de cada celda.

La prueba Ji cuadrada para la independencia

En el ejemplo 6.21 las columnas totales eran aleatorias, mientras que los renglones totales estaban presumiblemente fijos con anticipación, como consecuencia de que representaban los números de elementos probados de varias máquinas. En algunos casos, tanto el renglón como la columna total son aleatorios. En cualquiera de los dos casos se puede probar la hipótesis nula de que las probabilidades de los resultados columna son las mismas para cada resultado renglón, y la prueba es exactamente la misma para ambos casos. Se presenta un ejemplo donde tanto el renglón como la columna totales son aleatorios.

Ejemplo 6.22

Los alfileres cilíndricos de acero en el ejemplo 6.21 están sujetos a una especificación de longitud, así como a una de diámetro. Con respecto a la longitud, un alfiler podría satisfacer la especificación, o puede ser demasiado corto o demasiado largo. Se hace un muestreo y se categorizan mil alfileres en relación con las especificaciones de longitud y diámetro. En la tabla siguiente se presentan los resultados. Probar la hipótesis nula de que las proporciones de alfileres que son muy angostos, están bien o muy gruesos respecto de la especificación del diámetro no dependen de la clasificación con respecto a la especificación de longitud.

Valores observados para mil alfileres de acero

Longitud	Diámetro			Total
	Muy angosto	Está bien	Muy grueso	
Demasiado corto	13	117	4	134
Está bien	62	664	80	806
Demasiado largo	5	68	8	81
Total	80	849	92	1 021

Solución

Se inicia mediante la ecuación (6.8) para calcular los valores esperados. Estos últimos los da la siguiente tabla.

Valores observados para mil alfileres de acero

Longitud	Diámetro			Total
	Muy angosto	Está bien	Muy grueso	
Demasiado corto	10.50	111.43	12.07	134.0
Está bien	63.15	670.22	72.63	806.0
Demasiado largo	6.35	67.36	7.30	81.0
Total	80.0	849.0	92.0	1 021.0

Se observa que todos los valores esperados son mayores que o iguales a 5. (Uno de los valores observados no lo es; esto es irrelevante.) Por tanto, la prueba Ji cuadrada es adecuada. Se utiliza la ecuación (6.9) para calcular el valor del estadístico Ji cuadrada:

$$\begin{aligned}\chi^2 &= \frac{(13 - 10.50)^2}{10.50} + \cdots + \frac{(8 - 7.30)^2}{7.30} \\ &= \frac{6.25}{10.50} + \cdots + \frac{0.49}{7.30} \\ &= 7.46\end{aligned}$$

Debido a que hay tres renglones y tres columnas, el número de grados de libertad es $(3 - 1)(3 - 1) = 4$. Para obtener el P -valor se consulta la tabla Ji cuadrada (tabla A.6). Buscando debajo de cuatro grados de libertad, se encuentra que el punto 10% superior es 7.779. Se concluye que $P > 0.10$. No hay evidencia de que la longitud y el espesor estén relacionados.

Ejercicios para la sección 6.10

- Se fabrican remaches para un propósito determinado. La especificación de longitud es 1.20-1.30 cm. Se piensa que 90% de los remaches fabricados satisfacen la especificación, mientras que 5% son demasiado cortos, y 5% demasiado largos. En una muestra aleatoria de mil remaches, 860 satisfacían la especificación, 60 eran demasiado cortos, y 80 demasiado, largos. ¿Puede concluir que los porcentajes verdaderos son diferentes de 90, 5 y 5%?
 - Establezca la hipótesis nula adecuada.
 - Calcule los valores esperados bajo la hipótesis nula.
 - Calcule el valor de la estadística Ji cuadrada.
 - Determine el P -valor. ¿Cuál es su conclusión?
- Las especificaciones para el diámetro de un rodillo son 2.10-2.11 cm. Los rodillos muy gruesos pueden ser rectificadas, mientras que los muy angostos se deben desechar. Tres maquinistas pulen estos rodillos. Se coleccionaron muestras de los rodillos de cada maquinista y se midieron sus diámetros. Los resultados fueron:

¿Puede concluir que las proporciones de los rodillos en las tres categorías son diferentes entre los maquinistas?

 - Establezca la hipótesis nula adecuada.
 - Calcule los valores esperados bajo la hipótesis nula.
 - Calcule el valor de la estadística Ji cuadrada.
 - Determine el P -valor. ¿Cuál es su conclusión?
- El artículo "An Investment Tax Credit for Investing in New Technology: A Survey of California Firms" (R. Pope, en *The Engineering Economist*, 1997:269-287) examina el impacto potencial de un impuesto sobre el capital invertido. Se categorizaron varias compañías por tamaño (> 100 empleados contra ≤ 100 empleados) y la capacidad excesiva neta. Los números de las compañías en cada una de las categorías se presentan en la tabla siguiente:

Maquinista	Bien	Volver a pulir	Desechar	Total
A	328	58	14	400
B	231	48	21	300
C	409	73	18	500
Total	968	179	53	1 200

Capacidad excesiva neta	Pequeña	Grande
$< 0\%$	66	115
0-10%	52	47
11-20%	13	18
21-30%	6	5
$> 30\%$	36	25

¿Puede concluir que la distribución de la capacidad excesiva neta es diferente entre compañías pequeñas y grandes? Calcule el estadístico de prueba importante y el P -valor.

4. El artículo citado en el ejercicio 3 clasificó a las compañías por tamaño y porcentaje de la fuerza laboral actualmente empleada operando por completo. Los números de las compañías en cada una de las categorías se presentan en la siguiente tabla.

Porcentaje de fuerza laboral actualmente empleada operando por completo	Valores observados	
	Pequeña	Grande
> 100%	6	8
95–100%	29	45
90–94%	12	28
85–89%	20	21
80–84%	17	22
75–79%	15	21
70–74%	33	29
< 70%	39	34

¿Puede concluir que la distribución de fuerza laboral actualmente empleada es diferente entre compañías pequeñas y grandes? Calcule el estadístico de prueba relevante y el *P*-valor.

5. Para la tabla dada de los valores observados.

- a) Elabore la tabla correspondiente de valores esperados.
- b) Si es adecuado, realice la prueba Ji cuadrada para la hipótesis nula de que los resultados en los renglones y columnas son independientes. Si no es adecuado, explique por qué.

	Valores observados		
	1	2	3
A	15	10	12
B	3	11	11
C	9	14	12

6. Para la tabla dada de valores observados,

- a) Elabore la tabla correspondiente de valores esperados.
- b) Si es adecuado, realice la prueba Ji cuadrada para la hipótesis nula de que los resultados en los renglones y en las columnas son independientes. Si no lo es, explique por qué.

	Valores observados		
	1	2	3
A	25	4	11
B	3	3	4
C	42	3	5

7. Complete el enunciado: Para valores observados y esperados, _____.

- i) Los totales de los renglones en la tabla observada deben ser iguales que los totales de los renglones en la tabla esperada, pero los totales de las columnas no necesariamente son iguales.
- ii) Los totales de las columnas en la tabla observada deben ser iguales que los de las columnas en la tabla esperada, pero los totales de los renglones no necesariamente son iguales.
- iii) Tanto los totales de los renglones como los de las columnas en la tabla observada deben ser iguales que los totales de los renglones y los de las columnas, respectivamente, en la tabla esperada.
- iv) Ni los totales de los renglones ni los de las columnas en la tabla observada necesitan ser iguales a los totales de los renglones o a los de la columna en la tabla esperada.

8. Debido a una falla en la impresora, ninguno de los valores observados en la siguiente tabla se imprimieron, pero algunos de los totales marginales sí lo hicieron. ¿Es posible construir la tabla correspondiente de valores esperados a partir de la información dada? Si es así, constrúyala. Si no, describa la información adicional que necesitaría.

	Valores observados			Total
	1	2	3	
A	—	—	—	25
B	—	—	—	—
C	—	—	—	40
D	—	—	—	75
Total	50	20	—	150

9. Se supone que un generador de números aleatorios produce los dígitos 0 al 9 con probabilidad igual. Se generó una muestra de 200 dígitos, con la siguiente frecuencia generada para cada uno de los dígitos.

Dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia	21	17	20	18	25	16	28	19	22	14

¿Estos datos evidencian que el generador de números aleatorios no está funcionando adecuadamente? Explique.

10. En una planta de ensamblaje de camiones ligeros la supervisión rutinaria de la calidad de las soldaduras produce la tabla siguiente:

	Número de soldaduras		
	Alta calidad	Moderada calidad	Baja calidad
Turno de día	470	191	42
Turno de tarde	445	171	28
Turno de noche	257	129	17

¿Puede concluir que la calidad varía entre los turnos?

11. El artículo “Analysis of Unwanted Fire Alarm: Case Study” (W. Chow, N. Fong y C. Ho, en *Journal of Architectural Engineering*, 1999:62-65) presenta el número de falsas alarmas en varios sitios. Los números de falsas alarmas cada mes, divididos entre aquellos cuyas causas se conocen y los que se desconocen, están dados en la tabla siguiente. ¿Puede concluir que la proporción de falsas alarmas cuya causa se conoce es diferente mes a mes?

	Mes											
	1	2	3	4	5	6	7	8	9	10	11	12
Se conocen	20	13	21	26	23	18	14	10	20	18	14	
Se desconocen	12	2	16	12	22	30	32	32	14	16	10	12

12. En un cierto punto genético de un cromosoma cada individuo tiene una de tres secuencias de ADN diferentes (alelos). Los tres alelos se denotan mediante A, B, C. En otro punto genético en el mismo cromosoma cada organismo tenía uno de los tres alelos, denotados por 1, 2, 3. Por tanto, cada individuo tiene uno de los nueve posibles pares de alelos: A1, A2, A3, B1, B2, B3, C1, C2, o C3. Estos pares de alelos se

llaman *haplotipos*. Se dice que los puntos tienen un *enlace de equilibrio* si los dos alelos en un haplotipo particular son independientes. Se determinaron los haplotipos para 316 individuos. El siguiente resultado MINITAB presenta los resultados de una prueba Ji cuadrada de independencia.

Chi-Square Test: A, B, C

Expected counts are printed below
observed counts

Chi-Square contributions are printed
below expected counts

	A	B	C	Total
1	66	44	34	144
	61.06	47.39	35.54	
	0.399	0.243	0.067	
2	36	38	20	94
	39.86	30.94	23.20	
	0.374	1.613	0.442	
3	32	22	24	78
	33.08	25.67	19.25	
	0.035	0.525	1.170	
Total	134	104	78	316

Chi-Sq = 4.868, DF = 4,
P-Value = 0.301

- a) ¿Cuántos individuos que tenían el haplotipo B3 se observaron?
- b) ¿Cuál es el número esperado de individuos con el haplotipo A2?
- c) ¿Cuál de los nueve haplotipos se observó con menor frecuencia?
- d) ¿Cuál de los nueve haplotipos tiene el menor número esperado?
- e) ¿Puede concluir que los puntos no tienen un enlace de equilibrio (es decir, no independientes)? Explique.
- f) ¿Puede concluir que los puntos tienen una conexión de equilibrio (es decir, independientes)? Explique.

6.11 La prueba F para la igualdad de varianzas

Las pruebas que se han estudiado hasta ahora han implicado medias o proporciones. A veces se desea probar una hipótesis nula en la cual las dos poblaciones tienen varianzas iguales. En general, no hay ningún buen método para hacerlo. Sin embargo, para el caso especial donde ambas poblaciones son normales, hay un método disponible.

Sea $X_1, \dots, X_m$ una muestra aleatoria simple de una población $N(\mu_1, \sigma_1^2)$, y $Y_1, \dots, Y_n$ una muestra aleatoria simple de una población $N(\mu_2, \sigma_2^2)$. Suponga que las muestras se eligen de manera independiente. Los valores de las medias, μ_1 y μ_2 , aquí son irrelevantes; se trata sólo con las varianzas σ_1^2 y σ_2^2 . Observe que los tamaños muestrales, m y n , podrían ser diferentes. Sean s_1^2 y s_2^2 las varianzas muestrales. Esto es,

$$s_1^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2 \quad s_2^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

Cualesquiera de las tres hipótesis nulas puede probarse. Éstas son

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \quad \text{o de manera equivalente,} \quad \sigma_1^2 \leq \sigma_2^2$$

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \geq 1 \quad \text{o de manera equivalente,} \quad \sigma_1^2 \geq \sigma_2^2$$

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \quad \text{o de manera equivalente,} \quad \sigma_1^2 = \sigma_2^2$$

Los procedimientos para probar estas hipótesis son similares, pero no idénticos. Se describirá el procedimiento para probar la hipótesis nula $H_0: \sigma_1^2/\sigma_2^2 \leq 1$ contra $H_1: \sigma_1^2/\sigma_2^2 > 1$, y después se analizará cómo se puede modificar el procedimiento para probar las otras dos hipótesis.

El estadístico de prueba representa el cociente de las dos varianzas muestrales:

$$F = \frac{s_1^2}{s_2^2} \tag{6.10}$$

Cuando H_0 es verdadera se supone que $\sigma_1^2/\sigma_2^2 = 1$ (el valor más cercano a H_1), o de manera equivalente, que $\sigma_1^2 = \sigma_2^2$. Cuando H_0 es verdadera, s_1^2 y s_2^2 son, en promedio, del mismo tamaño, por lo que F es posiblemente cercano a 1. Cuando H_0 es falsa, $\sigma_1^2 > \sigma_2^2$, por lo que s_1^2 es posiblemente mayor que s_2^2 , y F es quizá mayor a 1. Para utilizar F como un estadístico de prueba se debe conocer su distribución nula. Esta última se denomina distribución F , enseguida se le describirá.

La distribución F

Los estadísticos que tienen una distribución F son cocientes de cantidades, tal como el de las dos varianzas muestrales en la ecuación (6.10). Por tanto, la distribución F tiene dos valores para los grados de libertad: uno asociado con el numerador, y el otro con el denominador. Los grados de libertad se indican con subíndices bajo la letra F . Por ejemplo, el símbolo $F_{3,16}$ de-

nota la distribución F con 3 grados de libertad para el numerador y 16 grados de libertad para el denominador. Observe que los grados de libertad para el numerador siempre se listan primero.

Se proporciona una tabla para la distribución F (tabla A.7 en el apéndice A). La tabla proporciona valores para ciertos cuantiles, o puntos de porcentajes superiores, para gran número de elecciones de grados de libertad. Como ejemplo, la figura 6.21 presenta la función de densidad de probabilidad de la distribución de $F_{3,16}$. El 5% superior de la distribución está sombreado. Para encontrar el punto 5% superior en la tabla, busque debajo de $\alpha = 0.050$, y en grados de libertad $\nu_1 = 3$, $\nu_2 = 16$. El valor es 3.24.

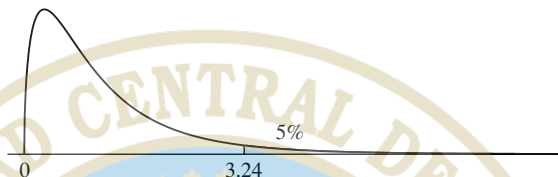


FIGURA 6.21 Función de densidad de probabilidad de la distribución $F_{3,16}$. El punto 5% superior es 3.24. [Véase la tabla F (tabla A.7) del apéndice A.]

El estadístico F para probar la igualdad de varianzas

La distribución nula del estadístico de prueba $F = s_1^2/s_2^2$ es $F_{m-1, n-1}$. El número de grados de libertad para el numerador es uno menos que el tamaño muestral usado para calcular s_1^2 , y la cantidad de grados de libertad para el denominador es uno menos que el tamaño muestral utilizado para calcular s_2^2 . Se muestra la prueba F con un ejemplo.

Ejemplo 6.23

En una serie de experimentos para determinar la tasa de absorción de ciertos pesticidas en la piel se aplicaron cantidades medidas de dos pesticidas a algunos especímenes de piel. Después de un tiempo se midieron las cantidades absorbidas (en μg). Para el pesticida A la varianza de las cantidades absorbidas en seis muestras fue de 2.3, mientras que para el B, la varianza de las cantidades absorbidas en diez especímenes fue de 0.6. Suponga que para cada pesticida las cantidades absorbidas constituyen una muestra aleatoria simple de una población normal. ¿Se puede concluir que la varianza en la cantidad absorbida es mayor para el pesticida A que para el B?

Solución

Sea σ_1^2 la varianza poblacional para el pesticida A, y σ_2^2 la varianza poblacional para el B. La hipótesis nula es

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1$$

Las varianzas muestrales son $s_1^2 = 2.3$ y $s_2^2 = 0.6$. El valor del estadístico de prueba es

$$F = \frac{2.3}{0.6} = 3.83$$

La distribución nula del estadístico de prueba es $F_{5,9}$. Si H_0 es verdadera, entonces s_1^2 será en promedio menor que s_2^2 . En consecuencia, entre mayor sea el valor de F , más fuerte es la evidencia contra H_0 . Al consultar la tabla F con cinco y nueve grados de libertad, se encuentra que el punto 5% superior es 3.48, mientras que el punto 1% superior es 6.06. Se concluye que $0.01 < P < 0.05$. En este sentido, hay fuertes evidencias contra la hipótesis nula. Véase la figura 6.22.

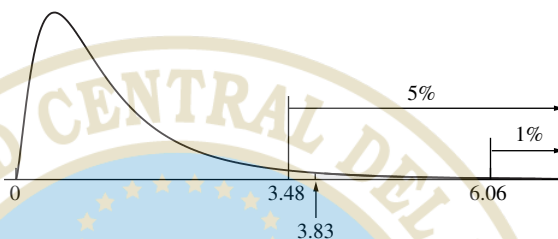


FIGURA 6.22 El valor observado del estadístico de prueba es 3.83. El punto 5% superior de la distribución $F_{5,9}$ es 3.48; el punto 1% superior es 6.06. Por tanto, el P -valor está entre 0.01 y 0.05.

Ahora se describen las modificaciones para el procedimiento mostrado en el ejemplo 6.23, que son necesarias para probar las otras hipótesis nulas. Para probar

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \geq 1$$

se podría, en principio, utilizar el estadístico de prueba s_1^2/s_2^2 , con valores *pequeños* del estadístico proporcionando evidencia contra H_0 . Sin embargo, puesto que la tabla F contiene solamente valores grandes (es decir, mayores que 1) para la estadística F , es más fácil utilizar el estadístico s_2^2/s_1^2 . Bajo H_0 , la distribución de s_2^2/s_1^2 es $F_{n-1,m-1}$.

Por último, se describe el método para probar la hipótesis de dos colas

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

Para esta hipótesis, tanto los valores grandes como los pequeños del estadístico s_1^2/s_2^2 proporcionan evidencia contra H_0 . El procedimiento es utilizar ya sea s_1^2/s_2^2 o s_2^2/s_1^2 , cualquiera que sea mayor que 1. El P -valor para la prueba de dos colas es dos veces el P -valor para la prueba de una cola. En otras palabras, el P -valor de la prueba de dos colas constituye dos veces el área de la cola superior de la distribución de F . Se muestra con un ejemplo.

Ejemplo

6.24

En el ejemplo 6.23, $s_1^2 = 2.3$ con tamaño de muestra de 6, y $s_2^2 = 0.6$ con tamaño de muestra de 10. Pruebe la hipótesis nula

$$H_0: \sigma_1^2 = \sigma_2^2$$

Solución

La hipótesis nula $\sigma_1^2 = \sigma_2^2$ es equivalente a $\sigma_1^2/\sigma_2^2 = 1$. Puesto que $s_1^2 > s_2^2$ se utiliza el estadístico de prueba s_1^2/s_2^2 . En el ejemplo 6.23 se encuentra que para la prueba de una cola, $0.01 < P < 0.05$. Por tanto, para la prueba de dos colas, $0.02 < P < 0.10$.

El siguiente resultado de computadora (de MINITAB) presenta la solución al ejemplo 6.24.

Test for Equal Variances

F-Test (normal distribution)

Test statistic = 3.83, p-value = 0.078

La prueba F es sensible a las desviaciones de la normalidad

La prueba F , de la misma manera que la prueba t , requiere que las muestras provengan de poblaciones normales. A diferencia de la prueba t , la prueba F para comparar varianzas es bastante sensible a esta suposición. Si las formas de las poblaciones son muy diferentes de la curva normal, la prueba F puede dar resultados falsos. Por esta razón, la prueba F para comparar varianzas se debe utilizar con precaución.

En los capítulos 8 y 9 se utilizará la distribución F para realizar ciertas pruebas de hipótesis en el contexto de la regresión lineal y del análisis de varianza. En estos ajustes, la prueba F es menos sensible a las violaciones de la suposición de normalidad.

La prueba F no puede probar que dos varianzas son iguales

En la sección 6.7 se presentan dos versiones de la prueba t para la diferencia entre dos medias. Una versión es, en general, aplicable, mientras que la segunda versión, que usa la varianza combinada, es adecuada solamente cuando las varianzas poblacionales son iguales. Cuando se decide si es adecuado suponer que las varianzas poblacionales son iguales, es tentador realizar una prueba F y suponer que las varianzas son iguales si la hipótesis nula de la igualdad no es rechazada. Desgraciadamente, este procedimiento es poco fiable, por la razón básica de que falla al rechazar la hipótesis nula que no justifica la suposición de que aquélla es verdadera. En general, una suposición de que las varianzas poblacionales sean iguales no se puede justificar con una prueba de hipótesis.

Ejercicios para la sección 6.11

- Determine el punto 5% superior de $F_{7,20}$.
- Determine el punto 1% superior de $F_{2,5}$.
- Una prueba F con cinco grados de libertad en el numerador y siete en el denominador produjo un estadístico de prueba cuyo valor fue de 7.46.
 - ¿Cuál es el P -valor si la prueba es de una cola?
 - ¿Cuál es el P -valor si la prueba es de dos colas?
- El contenido de azúcar, en mg/ml, de un caldo utilizado para fabricar un producto farmacéutico fue medido varias veces en cada uno de tres días sucesivos.
- Considerando el ejercicio 9 de la sección 5.6, ¿puede concluir que la varianza de las fuerzas de ruptura es diferente entre los dos compuestos?
- Respecto al ejercicio 11 de la sección 5.6, ¿puede concluir que el tiempo en que se paraliza la computadora es más variable en el séptimo mes que en el primero después de la instalación?

Día 1:	5.0	4.8	5.1	5.1	4.8	5.1	4.8
	4.8	5.0	5.2	4.9	4.9	5.0	
Día 2:	5.8	4.7	4.7	4.9	5.1	4.9	5.4
	5.3	5.3	4.8	5.7	5.1	5.7	
Día 3:	6.3	4.7	5.1	5.9	5.1	5.9	4.7
	6.0	5.3	4.9	5.7	5.3	5.6	

6.12 Prueba de nivel fijo

Puntos críticos y regiones de rechazo

Una prueba de hipótesis mide la factibilidad de la hipótesis nula produciendo un P -valor. Entre menor sea éste menos factible será la nula. Se ha indicado que no hay línea divisoria científicamente legítima entre la factibilidad e infactibilidad, por lo que es imposible especificar un P -valor “correcto” debajo del cual se deba rechazar H_0 . Cuando sea posible, es mejor sólo notificar el P -valor y no tomar una decisión firme de si se rechaza o no. Sin embargo, a veces se tiene que tomar una decisión. Por ejemplo, si se ha hecho un muestreo de los elementos de una línea de montaje para probar si la media del diámetro está dentro de la tolerancia, se debe tomar la decisión de si se recalibra el proceso. Si se extrae una muestra de piezas de un embarque y se examina en busca de defectos, se debe decidir si aceptar o no la remesa. Si se tomará una decisión con base en una prueba de hipótesis, no hay ninguna elección pero se elige un punto límite para el P -valor. Cuando se hace esto último la prueba se denomina de **nivel fijo**.

La prueba de nivel fijo es muy parecida a la de hipótesis que se ha analizado hasta ahora, excepto que se ha establecido una regla firme antes de tiempo para rechazar la hipótesis nula. Se elige un valor α , donde $0 < \alpha < 1$. Entonces se calcula el P -valor. Si $P \leq \alpha$, la hipótesis nula se rechaza y la hipótesis alternativa se toma como verdadera. Si $P > \alpha$, entonces la hipótesis nula se considera factible. El valor de α se llama **nivel de significancia**, o de manera más sencilla, **nivel**, de la prueba. Recuerde, de la sección 6.2, que si en una prueba se

obtiene un P -valor menor que o igual a α , se dice que la hipótesis nula se rechaza a un nivel de α (o $100\alpha\%$), o que el resultado es estadísticamente significativo a un nivel de α (o $100\alpha\%$). Como se ha mencionado, una elección común para α es 0.05.

Resumen

Para realizar una prueba de nivel fijo:

- Se elige un número α , donde $0 < \alpha < 1$. Éste se llama nivel de significancia, o nivel, de la prueba.
- Se calcula el P -valor en la manera acostumbrada.
- Si $P \leq \alpha$, rechaza H_0 . Si $P > \alpha$, no se rechaza H_0 .

Ejemplo 6.25

En relación con el ejemplo 6.1 de la sección 6.1, la media del desgaste en una muestra de 45 bolas de acero fue $\bar{X} = 673.2 \mu\text{m}$, y la desviación estándar $s = 14.9 \mu\text{m}$. Sea μ la media poblacional de desgaste. Una prueba de $H_0: \mu \geq 675$ contra $H_1: \mu < 675$ produjo un P -valor de 0.209. ¿Se puede rechazar H_0 a un nivel de 25%? ¿Se puede rechazar H_0 a un nivel de 5%?

Solución

El P -valor de 0.209 es menor que 0.25, por lo que si se hubiera elegido un nivel de significancia de $\alpha = 0.25$, se rechazaría H_0 . Por tanto, se rechaza H_0 a un nivel de 25%. Dado que $0.209 > 0.05$, no se rechaza H_0 a un nivel de 5%.

En una prueba de nivel fijo, un **punto crítico** constituye un valor del estadístico de prueba que produce un P -valor exactamente igual a α . Un punto crítico representa una línea divisoria para el estadístico de prueba igual que el nivel de significancia representa una línea divisoria para el P -valor. Si el estadístico de prueba está en un lado del punto crítico, el P -valor será menor que α y se rechazará H_0 . Si el estadístico de prueba está al otro lado del punto crítico, el P -valor será mayor que α , y H_0 no será rechazada. La región en el lado del punto crítico que conduce al rechazo se llama región de rechazo. El punto crítico también está en la **región de rechazo**.

Ejemplo 6.26

Se está evaluando una nueva mezcla de concreto. El plan es hacer un muestreo de 100 bloques de concreto hechos con la nueva mezcla; calcule la media muestral de la fuerza compresiva $\bar{X}$, después pruebe $H_0: \mu \leq 1\,350$ contra $H_1: \mu > 1\,350$, donde las unidades son MPa. Se

ha supuesto en las pruebas previas que en esta clase la desviación estándar poblacional σ será cercana a 70 MPa. Determine el punto crítico y la región de rechazo si la prueba se hace a un nivel de significancia de 5%.

Solución

Se rechazará H_0 si el P -valor es menor o igual a 0.05. El P -valor para esta prueba será el área a la derecha del valor de $\bar{X}$. Por tanto, el P -valor será menor que 0.05, y H_0 será rechazada, si el valor de $\bar{X}$ está en el 5% superior de la distribución nula (véase la figura 6.23). Por tanto, la región de rechazo consta de 5% superior de la distribución nula. El punto crítico es el límite de 5% superior. La distribución nula es normal, y de la tabla z se encuentra que el puntaje z del punto que corresponde a 5% superior de la curva normal es $z_{.05} = 1.645$. Por tanto, se puede expresar el punto crítico como $z = 1.645$, y la región de rechazo como $z \geq 1.645$. Con frecuencia es más conveniente expresar el punto crítico y la región de rechazo en función de $\bar{X}$, al convertir el puntaje z en las unidades originales. La distribución nula tiene una media $\mu = 1\,350$ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n} \approx 70/\sqrt{100} = 7$. Por tanto, el punto crítico se puede expresar como $\bar{X} = 1\,350 + (1.645)(7) = 1\,361.5$. La región de rechazo consta de los valores de $\bar{X}$ mayores que o iguales a 1 361.5.

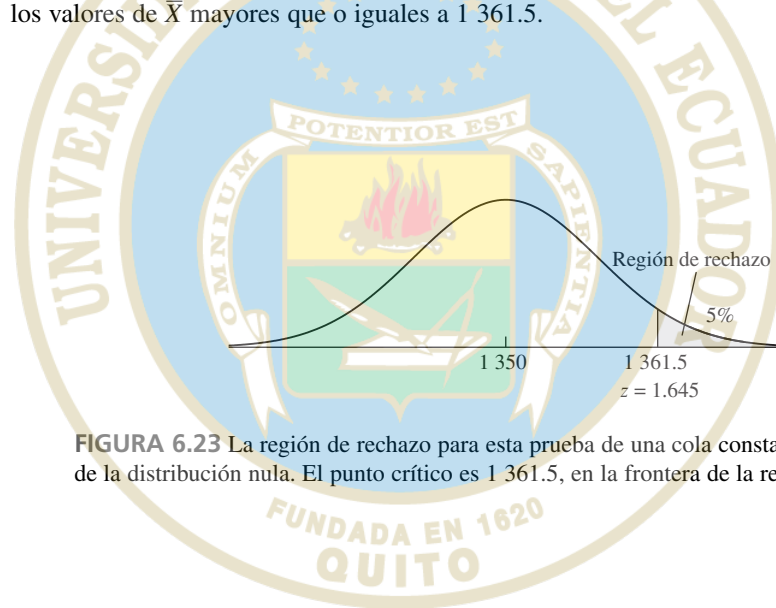


FIGURA 6.23 La región de rechazo para esta prueba de una cola consta del 5% superior de la distribución nula. El punto crítico es 1 361.5, en la frontera de la región de rechazo.

Ejemplo

6.27

En una prueba de hipótesis para determinar si una balanza está calibrada, la hipótesis nula es $H_0: \mu = 1\,000$ y la distribución nula de $\bar{X}$ es $N(1\,000, 0.26^2)$. (Esta situación se presentó en el ejemplo 6.2 de la sección 6.1.) Determine la región de rechazo si la prueba se hará en un nivel de significancia de 5%.

Solución

Dado que ésta es una prueba de dos colas, la región de rechazo está contenida en ambas colas de la distribución nula. Específicamente, H_0 será rechazada si $\bar{X}$ está en el 2.5% superior o inferior de la distribución nula (véase la figura 6.24). Los puntajes z que corresponden al

2.5% superior e inferior de la distribución son $z = \pm 1.96$. Por tanto, la región de rechazo consta de todos valores de $\bar{X}$ mayores que o iguales a $1\,000 + (1.96)(0.26) = 1\,000.51$, junto con los valores menores que o iguales a $1\,000 - (1.96)(0.26) = 999.49$. Observe que hay dos puntos críticos, 999.49 y 1 000.51.

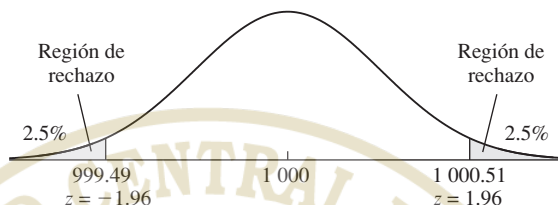


FIGURA 6.24 La región de rechazo para la prueba de dos colas consta tanto de 2.5% inferior como superior de la distribución nula. Hay dos puntos críticos, 999.49 y 1 000.51.

Errores de tipos I y II

Debido a que una prueba de nivel fijo da como resultado una decisión firme, hay una posibilidad de que la decisión podría estar equivocada. Hay dos formas en las cuales la decisión puede estar equivocada. Se puede rechazar H_0 cuando en realidad es verdadera. Esto último se conoce como error de tipo I. O se puede aceptar H_0 cuando es falsa. Ello se conoce como error de tipo II.

Cuando se diseñan experimentos cuyos datos serán analizados con una prueba de nivel fijo es importante tratar de hacer las probabilidades de los errores de tipos I y II razonablemente pequeñas. No tiene caso realizar un experimento que tiene grandes probabilidades de dar como resultado una decisión incorrecta. Es fácil controlar la probabilidad de un error de tipo I, como lo muestra el resultado siguiente.

Si α representa el nivel de significancia que se ha elegido para la prueba, entonces la probabilidad de un error de tipo I nunca es mayor que α .

Este hecho se muestra con el siguiente ejemplo. Sea $X_1, \dots, X_n$ una muestra aleatoria grande de una población con media μ y varianza σ^2 . Entonces $\bar{X}$ está normalmente distribuida con media μ y varianza σ^2/n . Suponga que se quiere probar $H_0: \mu \leq 0$ contra $H_1: \mu > 0$ a un nivel fijo $\alpha = 0.05$. Esto es, se rechazará H_0 si $P \leq 0.05$. La distribución nula, que se muestra en la figura 6.25, es normal con media 0 y varianza $\sigma_{\bar{X}}^2 = \sigma^2/n$. Suponga que la hipótesis nula es verdadera. Se calculará la probabilidad de un error de tipo I y se mostrará que no es mayor de 0.05.

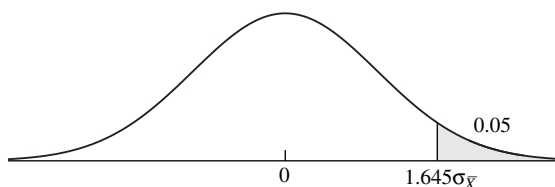


FIGURA 6.25 Distribución nula con región de rechazo para $H_0: \mu \leq 0$.

Un error de tipo I ocurrirá si se rechaza H_0 , ello sucederá si $P \leq 0.05$, además si $\bar{X} \geq 1.645\sigma_{\bar{X}}$. Por tanto, la región de rechazo es la región $\bar{X} \geq 1.645\sigma_{\bar{X}}$. Ahora, puesto que H_0 es verdadera, $\mu \leq 0$. Primero se considerará el caso donde $\mu = 0$. Entonces la distribución de $\bar{X}$ la da la figura 6.25. En este caso, $P(\bar{X} \geq 1.645\sigma_{\bar{X}}) = 0.05$, por lo que la probabilidad de rechazar H_0 y cometer un error de tipo I es igual a 0.05. Después se considera el caso donde $\mu < 0$. Entonces la distribución de $\bar{X}$ se obtiene al correr la curva en la figura 6.25 a la izquierda, por lo que $P(\bar{X} \geq 1.645\sigma_{\bar{X}}) < 0.05$, y la probabilidad de un error de tipo I es menor de 0.05. Se podría repetir esta ilustración con cualquier número α en lugar de 0.05. Se concluye que si H_0 es verdadera, la probabilidad de un error de tipo I nunca es mayor que α . Además, se observa que si μ está en el límite de H_0 ($\mu = 0$ en este caso), entonces la probabilidad de un error de tipo I es igual a α .

Por tanto, se puede hacer la probabilidad de un error de tipo I tan pequeña como se quiera, porque nunca es mayor que el nivel de significancia α que se eligió. Desgraciadamente, como se verá en la sección 6.13, entre menor se haga la probabilidad de un error de tipo I, mayor será la probabilidad de propiciar un error de tipo II. La estrategia usual es elegir un valor para α con el propósito de que la probabilidad del error de tipo I sea razonablemente pequeña. Como se ha mencionado, una elección convencional para α es 0.05. Entonces se calcula la probabilidad de un error de tipo II y se espera que no sea demasiado grande. Si lo fuera, se podría reducir solamente al rediseñar el experimento; por ejemplo, si se incrementa el tamaño muestral. Calcular y controlar el tamaño del error de tipo II es un poco más difícil que hacer lo mismo con el tamaño del error de tipo I. Esto último se analizará en la sección 6.13.

Resumen

Cuando se realiza una prueba de nivel fijo a un nivel de significancia α existen dos tipos de errores que se pueden cometer. Éstos son

- Error de tipo I: rechazar H_0 cuando es verdadera.
- Error de tipo II: aceptar H_0 cuando es falsa.

La probabilidad de un error de tipo I nunca es mayor que α .

Ejercicios para la sección 6.12

- Se realiza una prueba de hipótesis y el P -valor es 0.03. Verdadero o falso:
 - H_0 se rechaza a un nivel de 5%.
 - H_0 se rechaza a un nivel de 2%.
 - H_0 no se rechaza a un nivel de 10%.
- Se diseña un programa de tratamiento de aguas residuales para producir agua tratada con pH de 7. Sea μ la media del pH del agua tratada mediante dicho proceso. Se medirá el pH de 60 muestras de agua y se realizará una prueba de hipótesis $H_0: \mu = 7$ contra $H_1: \mu \neq 7$. Suponga que se sabe, con base en experimentos previos, que la desviación estándar del pH de las muestras de agua es aproximadamente 0.5.
 - Si la prueba se hace a un nivel de 5%, ¿cuál es la región de rechazo?
 - Si la media muestral del pH es 6.87, ¿se rechaza H_0 a un nivel de 10%?
 - Si la media muestral del pH es 6.87, ¿se rechaza H_0 a un nivel de 1%?
 - Si el valor 7.2 representa un punto crítico, ¿cuál es el nivel de la prueba?
- Se está evaluando un nuevo sistema de frenado en cierto tipo de automóvil; aquél se instalará si se puede demostrar que la distancia en que se detiene bajo ciertas condiciones controladas a velocidad de 30 mi/h es menor a 90 pies. Se sabe que bajo estas condiciones la desviación estándar de la distancia en que se detiene es de casi 5 pies. Se hará una muestra de 150 frenadas a 30 mi/h. Sea μ la media de la distancia en que el automóvil se detiene con el nuevo sistema de frenado.
 - Establezca las hipótesis nula y alternativa adecuadas.
 - Determine la región de rechazo si la prueba se realiza a un nivel de 5%.
 - Alguien sugiere rechazar H_0 si $\bar{X} \geq 89.4$ pies. ¿Esta es una región de rechazo adecuada, o algo está equivocado? Si fuera lo primero, determine el nivel de la prueba. De otro modo, explique la falla.
 - Alguien sugiere rechazar H_0 si $\bar{X} \leq 89.4$ pies. ¿Esta es una región de rechazo adecuada, o algo falla? Si fuera el primer caso, determine el nivel de la prueba. De otro modo, explique el error.
- Alguien sugiere rechazar H_0 si $\bar{X} \leq 89.4$ pies o si $\bar{X} \geq 90.6$ pies. ¿Esta es una región de rechazo adecuada, o algo está equivocado? Si fuera lo primero, determine el nivel de la prueba. De otro modo explique la falla.
- Se hace una prueba de las hipótesis $H_0: \mu \leq 10$ contra $H_1: \mu > 10$. Para cada una de las situaciones siguientes, determine si la decisión fue correcta u ocurrieron errores de tipo I o II.
 - $\mu = 8$, H_0 es rechazada.
 - $\mu = 10$, H_0 no es rechazada.
 - $\mu = 14$, H_0 no es rechazada.
 - $\mu = 12$, H_0 es rechazada.
- Un distribuidor afirma que no más de 10% de las partes que proporciona están defectuosas, sea p la proporción real de piezas con defecto. Se hace una prueba de hipótesis $H_0: p \leq 0.10$ contra $H_1: p > 0.10$. Para cada una de las siguientes situaciones, determine si la decisión fue correcta u ocurrieron errores de tipo I o II.
 - La afirmación es verdadera, y H_0 es rechazada.
 - La afirmación es falsa, y H_0 es rechazada.
 - La afirmación es verdadera, y H_0 no es rechazada.
 - La afirmación es falsa, y H_0 no es rechazada.
- Se realiza una prueba de hipótesis y se decide rechazar la hipótesis nula si $P < 0.10$. Si H_0 es, en efecto, verdadera, ¿cuál es la probabilidad máxima de que sea rechazada?
- Se está considerando un nuevo proceso para la licuefacción de carbón. El anterior proceso producía una media de 15 kg de combustible destilado por kilogramo de hidrógeno consumido en el proceso. Sea μ la media del nuevo proceso. Se realizará una prueba de $H_0: \mu \leq 15$ contra $H_1: \mu > 15$. El nuevo proceso se realizará si H_0 es rechazada. Implantar en la producción el nuevo proceso es muy costoso. Por tanto, sería un error muy oneroso instaurar en la producción el nuevo proceso si efectivamente no es mejor que el anterior. ¿Qué procedimiento proporciona una probabilidad menor para este error, probar a un nivel de 5% o bien un nivel de 1%?

6.13 Potencia

Una prueba de hipótesis da como resultado un error de tipo II si H_0 no es rechazada cuando es falsa. La **potencia** de una prueba es la probabilidad de *rechazar* H_0 cuando ésta es falsa. Por tanto,

$$\text{Potencia} = 1 - P(\text{error tipo II})$$

Para ser de utilidad, una prueba debe tener probabilidades razonablemente pequeñas tanto como errores de tipos I y II. El error de tipo I se conserva pequeño al elegir un valor pequeño de α como nivel de **significancia**. Entonces se calcula la potencia de la prueba. Si aquélla es grande, entonces la probabilidad de un error de tipo II es también pequeña, y la prueba es útil. Observe que los cálculos de potencia generalmente se hacen antes de recopilar los datos. El propósito de un cálculo de potencia es determinar, cuando se realiza una prueba de hipótesis, si posiblemente *rechaza o no* H_0 en el evento en que H_0 es falsa.

Como ejemplo de un cálculo de potencia, suponga que se ha desarrollado un nuevo proceso químico que puede aumentar la producción del proceso actual. Se sabe que éste tiene una media de producción de 80 y desviación estándar de 5, donde las unidades representan los porcentajes de un máximo teórico. Si la media de la producción del nuevo proceso se muestra mayor a 80, el nuevo proceso será realizado. Sea μ la media de la producción del nuevo proceso. Se propone operar el nuevo proceso 50 veces y después probar la hipótesis

$$H_0: \mu \leq 80 \quad \text{contra} \quad H_1: \mu > 80$$

a un nivel de significancia de 5%. Si H_0 es rechazada, se concluirá que $\mu > 80$, y el nuevo proceso se realizará. Se supone que si el nuevo proceso tuviera una media de la producción de 81, entonces sería muy benéfico instaurarlo en la producción. Si éste es, en efecto, el caso que $\mu = 81$, que es la potencia de la prueba, ¿cuál es la probabilidad de que H_0 se rechazará?

Antes de presentar la solución, se observa que para calcular la potencia es necesario especificar un valor particular de μ , en este caso $\mu = 81$, para la hipótesis alternativa. La razón para esto último consiste en que la potencia es diferente para valores diferentes de μ . Se verá que si μ está cerca de H_0 , la potencia será pequeña, mientras que si μ está lejos de H_0 la potencia será grande.

Calcular la potencia implica dos pasos:

1. Se calcula la región de rechazo.
2. Se calcula la probabilidad de que el estadístico de prueba esté en la región de rechazo, si la hipótesis alternativa es verdadera. Ésta es la potencia.

Se iniciará por encontrar la potencia de la prueba al calcular la región de rechazo mediante el método que se muestra en el ejemplo 6.26 de la sección 6.12. Primero se debe encontrar la distribución nula. Se sabe que el estadístico $\bar{X}$ tiene una distribución normal con media μ , y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, donde $n = 50$ es el tamaño muestral. Bajo H_0 , se toma $\mu = 80$. Ahora se debe encontrar una aproximación para σ . En la práctica éste puede ser un problema difícil, porque la muestra aún no se ha extraído, por lo que no hay una desviación estándar muestral s . Hay algunas maneras en las que podría ser posible aproximar σ .

A veces se ha extraído una muestra preliminar pequeña; por ejemplo, en un estudio de viabilidad, y la desviación estándar de esta muestra podría ser una aproximación satisfactoria para σ . En otros casos, podría existir una muestra de una población similar, cuya desviación estándar podría utilizarse. En este ejemplo hay una larga historia de un proceso actualmente utilizado, cuya desviación estándar es 5. Se dice que es razonable suponer que la desviación estándar del nuevo proceso sea similar a la del proceso actual. Por tanto, se desea suponer que la desviación estándar poblacional para el nuevo proceso es $\sigma = 5$ y que $\sigma_{\bar{X}} = 5/\sqrt{50} = 0.707$.

La figura 6.26 presenta la distribución nula de $\bar{X}$. Dado que H_0 especifica que $\mu \leq 80$, valores grandes de $\bar{X}$ no concuerdan con H_0 , por lo que el P -valor será el área a la derecha del valor observado de $\bar{X}$. El P -valor será menor o igual a 0.05 si $\bar{X}$ está en el 5% superior de la distribución nula. Este 5% superior está en la región de rechazo. El punto crítico tiene un puntaje z de 1.645, por lo que su valor es $80 + (1.645)(0.707) = 81.16$. Se rechazará H_0 si $\bar{X} \geq 81.16$. Ésta es la región de rechazo.

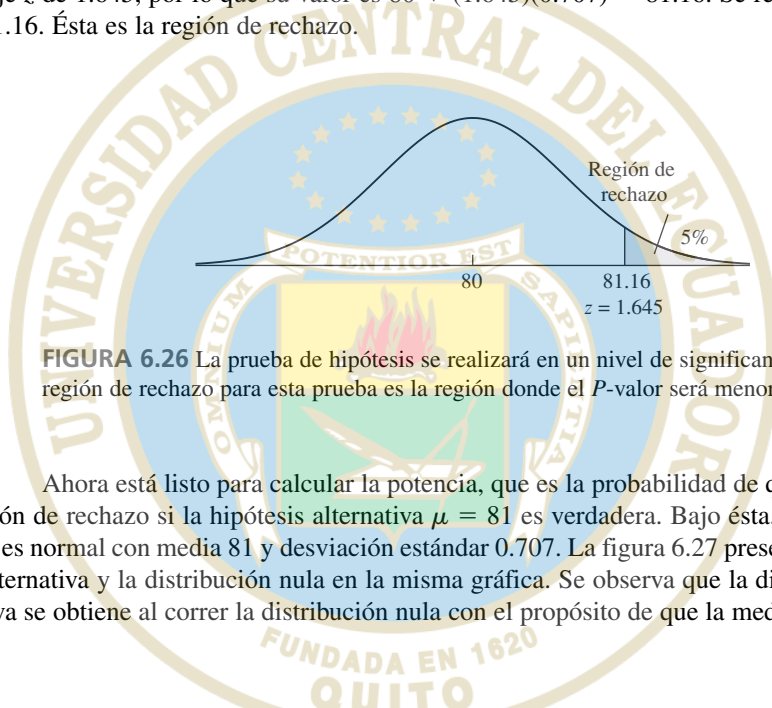


FIGURA 6.26 La prueba de hipótesis se realizará en un nivel de significancia de 5%. La región de rechazo para esta prueba es la región donde el P -valor será menor de 0.05.

Ahora está listo para calcular la potencia, que es la probabilidad de que $\bar{X}$ esté en la región de rechazo si la hipótesis alternativa $\mu = 81$ es verdadera. Bajo ésta, la distribución de $\bar{X}$ es normal con media 81 y desviación estándar 0.707. La figura 6.27 presenta la distribución alternativa y la distribución nula en la misma gráfica. Se observa que la distribución alternativa se obtiene al correr la distribución nula con el propósito de que la media sea la media al-

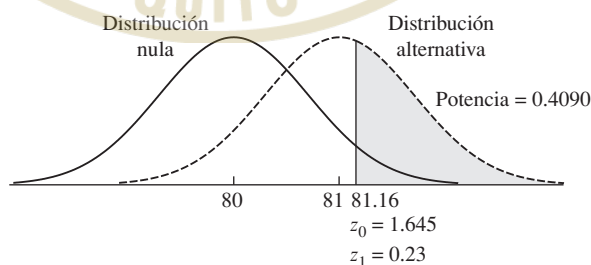


FIGURA 6.27 La región de rechazo, que consta de 5% superior de la distribución nula, está sombreada. El puntaje z del punto crítico es $z_0 = 1.645$ bajo la distribución nula y $z_1 = 0.23$ bajo la alternativa. La potencia constituye el área de la región de rechazo bajo la distribución alternativa, que es 0.4090.

ternativa de 81, en vez de la media nula de 80. Debido a que la distribución alternativa es corrida hacia delante, la probabilidad de que el estadístico de prueba esté en la región de rechazo es mayor a que esté bajo H_0 . De manera específica, el puntaje z bajo H_1 para el punto crítico 81.16 es $z = (81.16 - 81)/0.707 = 0.23$. El área a la derecha de $z = 0.23$ es 0.4090. Ésta es la potencia de la prueba.

Una potencia de 0.4090 es muy baja. Esto último significa que si la media de la producción del nuevo proceso es, en realidad, igual a 81, hay solamente una probabilidad de 41% de que el experimento propuesto detecte la mejoría sobre el proceso antiguo y permita que el nuevo proceso se realice. Sería poco sabio invertir tiempo y dinero en aplicar este experimento, ya que tiene gran un riesgo de fallar.

Es natural preguntarse qué tan grande debe ser la potencia para que sea útil realizar una prueba. Como con los P -valores, no hay una línea divisoria científicamente válida entre potencia suficiente e insuficiente. En general, las pruebas con potencias mayores a 0.80 o quizás 0.90 se consideran aceptables, pero no hay reglas prácticas bien establecidas.

Se ha mencionado que la potencia depende del valor elegido de μ para representar la hipótesis alternativa y que es mayor cuando el valor está lejos de la media de la nula. El ejemplo 6.28 muestra esto último.

Ejemplo

6.28

Determine la potencia de la prueba a un nivel de 5% de $H_0: \mu \leq 80$ contra $H_1: \mu > 80$ para la media de la producción del nuevo proceso bajo la alternativa $\mu = 82$, suponiendo que $n = 50$ y $\sigma = 5$.

Solución

Ya se ha terminado el primer paso de la solución, que es calcular la región de rechazo. Se rechaza H_0 si $\bar{X} \geq 81.16$. La figura 6.28 presenta las distribuciones alternativa y nula en la misma gráfica. El puntaje z para el punto crítico de 81.16 bajo la hipótesis alternativa es $z = (81.16 - 82)/0.707 = -1.19$. El área a la derecha de $z = -1.19$ es 0.8830. Ésta es la potencia.

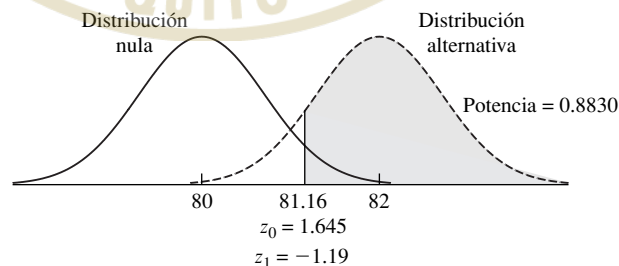


FIGURA 6.28 La región de rechazo, que consta del 5% superior de la distribución nula, está sombreada. El puntaje z del punto crítico es $z_0 = 1.645$ bajo la distribución nula y $z_1 = -1.19$ bajo la hipótesis alternativa. La potencia es el área de la región de rechazo bajo la distribución alternativa, que es 0.8830.

Debido a que la distribución alternativa se obtiene al correr a la distribución nula, la potencia depende de qué tipo de valor alternativo se elija para μ , y tiene un rango desde apenas ser mayor del nivel de significancia α hasta arriba de 1. Si la media alternativa se elige muy cerca de la media de la nula, la curva alternativa será casi idéntica a la nula, y la potencia será muy cercana a α . Si la media alternativa está lejos de la nula, casi toda el área bajo la curva alternativa estará en la región de rechazo, y la potencia será cercana a 1.

Cuando la potencia no es lo suficientemente grande, se puede aumentar al incrementar el tamaño muestral. Cuando se planea un experimento se puede determinar el tamaño muestral necesario para lograr la potencia deseada. El ejemplo 6.29 ilustra esto último.

Ejemplo

6.29

Con el propósito de evaluar la hipótesis $H_0: \mu \leq 80$ contra $H_1: \mu > 80$, considerando la media de la producción del nuevo proceso, ¿cuántas veces se debe operar el nuevo proceso con el fin de que una prueba con un nivel de significancia de 5% tenga una potencia de 0.90 contra la alternativa $\mu = 81$, si se supone que $\sigma = 5$?

Solución

Sea n el tamaño de muestra necesario. Primero se utiliza la distribución nula para expresar el punto crítico de la prueba en función de n . La distribución nula de $\bar{X}$ es normal con media 80 y desviación estándar $5/\sqrt{n}$. Por tanto, el punto crítico es $80 + 1.645(5/\sqrt{n})$. Ahora se utiliza la distribución alternativa con el fin de obtener una expresión diferente para el punto crítico en función de n . Consulte la figura 6.29. La potencia de la prueba es el área de la región de rechazo bajo la curva alternativa. Esta área debe ser 0.90. Por tanto, el puntaje z para el punto crítico, bajo la hipótesis alternativa, es $z = -1.28$. El punto crítico es, por tanto, $81 - 1.28(5/\sqrt{n})$. Ahora se tienen dos expresiones diferentes para el punto crítico. Debido a que sólo hay un punto crítico, estas dos expresiones son iguales. En consecuencia, se hacen iguales y se despeja n .

$$80 + 1.645 \left(\frac{5}{\sqrt{n}} \right) = 81 - 1.28 \left(\frac{5}{\sqrt{n}} \right)$$

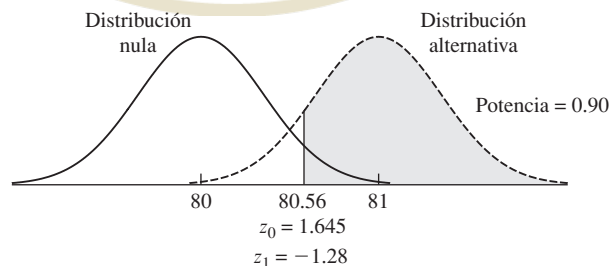


FIGURA 6.29 Para lograr una potencia de 0.90, con nivel de significancia de 0.05, el puntaje z para el punto crítico debe ser $z_0 = 1.645$ bajo la distribución nula y de $z_1 = -1.28$. Bajo la distribución alternativa.

Despejando para n produce $n \approx 214$. El punto crítico se puede calcular cuando se sustituye este valor para n en cualquier lado de la ecuación anterior. El punto crítico es 80.56.

Utilización de una computadora para calcular la potencia

Se ha presentado un método con el cual calcular la potencia y el tamaño muestral necesario con el fin de obtener una potencia específica para una prueba de una cola con muestras grandes de una media poblacional. Es razonablemente directo extender este método para calcular la potencia y los tamaños muestrales necesarios para pruebas de dos colas y de proporciones. Es más difícil calcular la potencia para una prueba t , prueba F , o prueba Ji cuadrada. Sin embargo, los paquetes computacionales pueden calcular la potencia y los tamaños muestrales necesarios para todas estas pruebas. Se presentan algunos ejemplos.

Ejemplo

6.30

Un individuo realizará una encuesta a una muestra aleatoria de votantes en una comunidad para calcular la proporción de los que aprueban una medida respecto de bonos escolares. Sea p la proporción poblacional que aprueba la medida. El encuestador probará $H_0: p = 0.50$ contra $H_1: p \neq 0.50$ a un nivel de 5%. Si se hace un muestreo de 200 votantes, ¿cuál es la potencia de la prueba si el valor verdadero de p es 0.55?

Solución

El siguiente resultado de computadora (de MINITAB) presenta la solución:

```
Power and Sample Size
Test for One Proportion
Testing proportion = 0.5 (versus not = 0.5)
Alpha = 0.05

Alternative      Sample      Power
Proportion      Size      Power
0.55            200      0.292022
```

Los primeros dos renglones establecen que éste es un cálculo de potencia para una prueba de una sola proporción poblacional p . Los siguientes dos renglones establecen las hipótesis nula y alternativa, así como el nivel de significancia de la prueba. Observe que se ha especificado una prueba de dos colas con un nivel de significancia $\alpha = 0.05$. El renglón siguiente representa la proporción alternativa, el valor de p (0.55) que se está suponiendo como verdadero cuando se calcula la potencia. El tamaño muestral se ha especificado en 200, y la potencia se calcula en 0.292.

Ejemplo

6.31

En relación con el ejemplo 6.30, ¿cuántos votantes se deben muestrear con el propósito de que la potencia sea de 0.8 cuando el valor verdadero es $p = 0.55$?

Solución

El siguiente resultado de computadora (de MINITAB) presenta la solución:

Power and Sample Size			
Test for One Proportion			
Testing proportion = 0.5 (versus not = 0.5)			
Alpha = 0.05			
Alternative Proportion	Sample Size	Target Power	Actual Power
0.55	783	0.8	0.800239

El tamaño muestral necesario es de 783. Observe que la potencia verdadera es ligeramente mayor a 0.80. Debido a que el tamaño muestral es discreto, no es posible encontrar un tamaño muestral que proporcione exactamente la potencia pedida (la potencia objetivo). Como consecuencia de lo anterior, MINITAB calcula el tamaño muestral menor para la potencia mayor que la que se pidió originalmente.

Ejemplo

6.32

Embarques de granos de café se examinan en busca de humedad. Un alto índice de humedad indica posible contaminación de agua, lo que origina el rechazo del embarque. Sea μ la media del contenido de humedad (en porcentaje por peso) en un embarque. Se hicieron cinco mediciones de humedad a los granos elegidos aleatoriamente del embarque. Se hará una prueba de hipótesis $H_0: \mu \leq 10$ contra $H_1: \mu > 10$ a un nivel de 5%, utilizando la prueba t de Student. ¿Cuál es la potencia de la prueba si el contenido de humedad verdadero es de 12%, y la desviación estándar, $\sigma = 1.5\%$?

Solución

El siguiente resultado de computadora (de MINITAB) presenta la solución:

Power and Sample Size		
1-Sample t Test		
Testing mean = null (versus > null)		
Calculating power for mean = null + difference		
Alpha = 0.05 Assumed standard deviation = 1.5		
Difference	Sample Size	Power
2	5	0.786485

La potencia depende solamente de la diferencia entre la media verdadera y la media de la nula que es $12 - 10 = 2$, y no sobre las medias mismas. La potencia es 0.786. Observe que el resultado especifica que ésta es la potencia para una prueba de una cola.

Ejemplo

6.33

Con referencia al ejemplo 6.32, encuentre el tamaño muestral necesario con el propósito de que la potencia sea al menos 0.9.

Solución

El siguiente resultado de computadora (de MINITAB) presenta la solución:

Power and Sample Size

1-Sample t Test

Testing mean = null (versus > null)
 Calculating power for mean = null + difference
 Alpha = 0.05 Assumed standard deviation = 1.5

Difference	Sample Size	Target Power	Actual Power
2	7	0.9	0.926750

El valor menor de la muestra para la cual la potencia es 0.9 o más es 7. La potencia verdadera es 0.927.

Resumiendo, los cálculos de potencia son importantes para asegurar que los experimentos tengan el potencial de proporcionar conclusiones útiles. Muchos organismos que proporcionan financiamiento para investigación científica requieren que con cada propuesta se presenten cálculos de potencia de las pruebas de hipótesis que realizarán.

Ejercicios para la sección 6.13

- Una prueba tiene potencia de 0.90 cuando $\mu = 15$. Verdadero o falso:
 - La probabilidad de rechazar H_0 cuando $\mu = 15$ es 0.90.
 - La probabilidad de tomar una decisión correcta cuando $\mu = 15$ es 0.90.
 - La probabilidad de tomar una decisión correcta cuando $\mu = 15$ es 0.10.
 - La probabilidad de que H_0 es verdadera cuando $\mu = 15$ es 0.10.
 - La probabilidad de cometer un error de tipo I cuando $\mu = 3.5$ es 0.20.
 - La probabilidad de cometer un error de tipo II cuando $\mu = 3.5$ es 0.80.
 - La probabilidad de cometer un error de tipo II cuando $\mu = 3.5$ es 0.20.
 - La probabilidad de que H_0 es falsa cuando $\mu = 3.5$ es 0.80.
- Una prueba tiene una potencia de 0.80 cuando $\mu = 3.5$. Verdadero o falso:
 - La probabilidad de rechazar H_0 cuando $\mu = 3.5$ es 0.80.
 - La probabilidad de cometer un error de tipo I cuando $\mu = 3.5$ es 0.80.
- Si el tamaño de muestra permanece constante, y el nivel α aumenta, entonces la potencia será _____. Opciones: aumenta, disminuye.
- Si el nivel α permanece constante, y el tamaño muestral aumenta, entonces la potencia _____. Opciones: aumentará, disminuirá.

5. Una compañía productora de neumáticos afirma que los tiempos de vida de sus llantas tienen un promedio de cincuenta mil millas. Se sabe que la desviación estándar de los tiempos de vida de los neumáticos es de cinco mil millas. Se muestrean 100 neumáticos y se probará la hipótesis de que la media de los tiempos de vida de los neumáticos es, al menos, de cincuenta mil millas contra la alternativa que es menor. Suponga, que en efecto, la media verdadera de los tiempos de vida es de 49 500 millas.
- Establezca las hipótesis nula y alternativa. ¿Cuál es verdadera?
 - Se ha decidido rechazar H_0 si la media de muestra es menor a 49 400. Determine el nivel y la potencia de esta prueba.
 - Si la prueba se realiza a un nivel de 5%, ¿cuál es la potencia?
 - ¿En qué nivel se debe realizar la prueba con el propósito de que la potencia sea de 0.80?
 - Le están dando la oportunidad de probar más neumáticos. ¿Qué tantas llantas se deben probar en total con el propósito de que la potencia sea de 0.80 si la prueba se hace a un nivel de 5%?
6. La media del tiempo de secado de cierta pintura en determinada aplicación es de 12 minutos. Se probará un nuevo aditivo para probar si reduce el tiempo de secado. Se pintarán cientos de muestras y se calculará la media del tiempo de secado $\bar{X}$. Suponga que la desviación estándar poblacional del tiempo de secado es $\sigma = 2$ minutos. Sea μ la media del tiempo de secado para la nueva pintura. La hipótesis nula $H_0: \mu \geq 12$ será probada contra la alternativa $H_1: \mu < 12$. Suponga que los investigadores desconocen que la media verdadera del tiempo de secado de la nueva pintura es de 11.5 minutos.
- Se decide rechazar H_0 si $\bar{X} \leq 11.7$. Determine el nivel y la potencia de esta prueba.
 - ¿Para qué valores de $\bar{X}$ se debe rechazar H_0 con el propósito de que la potencia de la prueba sea de 0.90? ¿Cuál será entonces el nivel?
 - ¿Para qué valores de $\bar{X}$ se debe rechazar H_0 con el propósito de que el nivel de la prueba sea de 5%? ¿Cuál será la potencia?
 - ¿Qué tan grande se necesita que sea la muestra con el propósito de que una prueba de nivel 5% tenga potencia de 0.90?
7. Un cálculo de potencia ha mostrado que si $\mu = 8$, la potencia de una prueba de $H_0: \mu \geq 10$ contra $H_1: \mu < 10$ es 0.90. Si en lugar de $\mu = 7$, ¿cuál de los siguientes enunciados es verdadero?
- La potencia de la prueba será menor de 0.90.
 - La potencia de la prueba será mayor a 0.90.
 - No se puede determinar la potencia de la prueba sin conocer la desviación estándar poblacional σ .
8. Se supone que un nuevo proceso para producir placas de silicio de circuitos integrados reduce la proporción de defectos en 10%. Se probará una muestra de 250 placas. Sea X el número de defectos en la muestra. Sea p la proporción poblacional de defectos producidos por el nuevo proceso. Se hará una prueba de $H_0: p \geq 0.10$ contra $H_1: p < 0.10$. Suponga que el valor verdadero de p es realmente 0.06.
- Se decide rechazar H_0 si $X \leq 18$. Determine el nivel de esta prueba.
 - Se decide rechazar H_0 si $X \leq 18$. Determine la potencia de esta prueba.
 - ¿Se debe utilizar la misma desviación estándar para X con el fin de calcular tanto la potencia como el nivel? Explique.
 - ¿Cuántas placas se deben muestrear con el propósito de que la potencia sea de 0.90 si la prueba se hace a un nivel de 5%?
9. El siguiente resultado de MINITAB presenta los resultados de un cálculo de potencia para una prueba respecto de una proporción poblacional p .
-
- Power and Sample Size
Test for One Proportion
Testing proportion = 0.5
(versus not = 0.5)
Alpha = 0.05
- | Alternative Proportion | Sample Size | Power |
|------------------------|-------------|----------|
| 0.4 | 150 | 0.691332 |
-
- ¿La potencia calculada para una prueba es de una o de dos colas?
 - ¿Cuál es la hipótesis nula para la cual se calcula la potencia?
 - ¿Para qué valor alterno de p es la potencia calculada?
 - Si el tamaño muestral fuera de 100, ¿sería la potencia menor que 0.7, mayor a 0.7, o es imposible deducirlo del resultado? Explique.
 - Si el tamaño muestral fuera 200, ¿la potencia sería menor que 0.6, mayor a 0.6, o es imposible deducirlo del resultado? Explique.

f) Para un tamaño de muestra de 150, ¿la potencia contra la alternativa $p = 0.3$ es menor que 0.65, mayor que 0.65, o es imposible deducirlo del resultado? Explique.

g) Para un tamaño de muestra de 150, ¿la potencia contra la alternativa $p = 0.45$ es menor que 0.65, mayor que 0.65, o es imposible deducirlo del resultado? Explique.

10. El siguiente resultado de MINITAB presenta los resultados de un cálculo de potencia con respecto a una media poblacional μ .

Power and Sample Size

1-Sample t Test

Testing mean = null (versus $>$ null)

Calculating power for mean = null + difference

Alpha = 0.05 Assumed standard deviation = 1.5

Difference	Sample Size	Target Power	Actual Power
1	18	0.85	0.857299

- ¿La potencia calculada para una prueba es de una o de dos colas?
- Suponga que el valor de μ utilizado para la hipótesis nula es $\mu = 3$. ¿Para qué valor alterno de μ se calcula la potencia?
- Si el tamaño muestral fuera 25, ¿la potencia sería menor que 0.85, mayor que 0.85, o es imposible deducirlo del resultado? Explique.
- Si la diferencia fuera 0.5, ¿la potencia sería menor que 0.90, mayor que 0.90, o es imposible deducirlo del resultado? Explique.
- Si el tamaño muestral era de 17, ¿la potencia sería menor que 0.85, mayor que 0.85, o es imposible deducirlo del resultado? Explique.

11. El siguiente resultado de MINITAB presenta los resultados de un cálculo de potencia para una prueba de la diferencia entre dos medias $\mu_1 - \mu_2$.

Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus not =)

Calculating power for mean 1 = mean 2 + difference

Alpha = 0.05 Assumed standard deviation = 5

Difference	Sample Size	Target Power	Actual Power
3	60	0.9	0.903115

The sample size is for each group.

- ¿La potencia calculada para una prueba es de una o de dos colas?
- Si los tamaños de muestra fueran 50 en cada grupo, ¿la potencia sería menor que 0.9, mayor que 0.9, o es imposible deducirlo del producto? Explique.
- Si la diferencia fuera 4, ¿la potencia sería menor que 0.9, mayor que 0.9, o es imposible deducirlo del resultado? Explique.

6.14 Pruebas múltiples

A veces se presenta una situación en la que es necesario realizar muchas pruebas de hipótesis. La regla básica de esta situación es que entre más pruebas se realicen, la confianza que se puede tener en los resultados disminuye. En esta sección se presenta un ejemplo para ilustrar dicho enfoque.

En cierta aplicación industrial, se supone que la aplicación de una capa dura que contiene partículas muy pequeñas de carburo de tungsteno puede reducir el desgaste en engranes de levas. Hay muchas formulaciones posibles para la capa, variando el tamaño y la concentración de las partículas de carburo de tungsteno. Se elaboraron 20 formulaciones diferentes. Cada una se probó aplicándola a muchos engranes, y después se midió el desgaste de éstos después de transcurrido cierto lapso. Se sabe, con base en la experiencia, que la media del desgaste para engranes no cubiertos en dicho periodo es de $100 \mu\text{m}$. Para cada formulación se hizo una prueba de la hipótesis nula $H_0: \mu \geq 100 \mu\text{m}$. H_0 muestra que la formulación no reduce el desgaste. Para 19 de las 20 formulaciones, el P -valor era mayor a 0.05, por lo que H_0 no fue rechazada. Para una formulación, H_0 fue rechazada. Podría parecer natural llegar a la conclusión de que esta formulación reduce realmente el desgaste. Los ejemplos 6.34 a 6.37 mostrarán que esta conclusión es prematura.

Ejemplo

6.34

Si sólo se probó una formulación, y no tenía ningún efecto sobre el desgaste, ¿cuál es la probabilidad de que H_0 fuera rechazada, dando como resultado una conclusión equivocada?

Solución

Si la formulación no tiene ningún efecto sobre el desgaste, entonces $\mu = 100 \mu\text{m}$, por lo que H_0 es verdadera. Rechazar H_0 representa un error de tipo I. Por tanto, la cuestión es preguntarse por la probabilidad de un error de tipo I. En general, esta probabilidad es siempre menor que o igual al nivel de significancia de la prueba, que en este caso es de 5%. Debido a que $\mu = 100$ está en la frontera de H_0 , la probabilidad de un error de tipo I es igual al nivel de significancia. La probabilidad de que H_0 se rechazará es 0.05.

Ejemplo

6.35

Como consecuencia de que H_0 fue rechazada para una de las 20 formulaciones, ¿es factible que realmente no tenga ningún efecto sobre el desgaste?

Solución

Sí es factible que ninguna de las formulaciones, incluyendo la que rechazó H_0 , tenga algún efecto sobre el desgaste. Se hicieron 20 pruebas de hipótesis. Para cada prueba había una probabilidad de 5% (es decir, una oportunidad entre 20) de un error del tipo I. Por tanto, se espera que, en promedio, de las 20 hipótesis nulas verdaderas, una será rechazada. Por lo que rechazar H_0 es una, exactamente, de las 20 pruebas que se esperaría para el caso de que en ninguna de las formulaciones hubiera diferencia.

Ejemplo 6.36

Si, en efecto, ninguna de las 20 formulaciones tiene efecto en el desgaste, ¿cuál es la probabilidad de que H_0 sea rechazada en una o más de éstas?

Solución

Se encuentra primero la probabilidad de que la conclusión correcta (no rechazar H_0) se toma para todas las formulaciones. Para cada formulación, la probabilidad de que H_0 no sea rechazada es $1 - 0.05 = 0.95$, por lo que la probabilidad de que H_0 no sea rechazada para ninguna de las 20 formulaciones es $(0.95)^{20} = 0.36$. Por tanto, la probabilidad es $1 - 0.36 = 0.64$ de que es incorrecto rechazar H_0 para una o más de las formulaciones.

Ejemplo 6.37

El experimento se repite. Esta vez el operador se olvida de aplicar las capas, así que cada una de las 20 mediciones de desgaste realmente se hacen en engranes no cubiertos. ¿Es probable que una o más de las formulaciones parecerán reducir el desgaste, en tal caso H_0 probablemente será rechazada?

Solución

Sí. El ejemplo 6.36 muestra que la probabilidad es 0.64 de que una o más de las capas parecerá reducir el desgaste, aun si no se aplica realmente la capa.

Los ejemplos 6.34 a 6.37 ilustran un fenómeno conocido como problema de la **prueba múltiple**. En otras palabras, el problema de prueba múltiple es éste: cuando se rechaza H_0 , se tienen fuertes evidencias de que es falsa, pero ello no significa que haya certeza en dicha conclusión. En ocasiones una hipótesis nula verdadera será rechazada. Cuando se realizan muchas pruebas, es más probable que algunas hipótesis nulas verdaderas sean rechazadas. Por tanto, cuando se realizan muchas pruebas, es difícil decir cuáles de las hipótesis nulas rechazadas son realmente falsas y a cuáles les corresponden errores de tipo I.

El método de Bonferroni

Este método proporciona una manera de ajustar P -valores hacia arriba cuando se realizan algunas pruebas de hipótesis. Si un P -valor permanece pequeño después del ajuste, la hipótesis nula puede ser rechazada. Para hacer el ajuste de Bonferroni, sólo multiplique el P -valor por el número de pruebas realizadas. He aquí dos ejemplos.

Ejemplo 6.38

Se prueban cuatro diferentes formulaciones de capa para verificar si reducen el desgaste sobre engranes de levas a un valor debajo de los $100 \mu\text{m}$. La hipótesis nula $H_0: \mu \geq 100 \mu\text{m}$ es probada para cada formulación, y los resultados son

- Formulación A: $P = 0.37$
- Formulación B: $P = 0.41$
- Formulación C: $P = 0.005$
- Formulación D: $P = 0.21$

El operador sospecha que la formulación C podría ser eficaz, pero sabe que el P -valor de 0.005 es poco fiable porque se han realizado algunas pruebas. Utilice el ajuste de Bonferroni para producir un P -valor confiable.

Solución

Se realizaron cuatro pruebas, por lo que el ajuste de Bonferroni producía $P = (4)(0.005) = 0.02$ para la formulación C, por lo que la evidencia es razonablemente fuerte de que la formulación C es eficaz.

Ejemplo

6.39

Se supone que en el ejemplo 6.38 el P -valor de la formulación C era 0.03 en lugar de 0.005. Entonces ¿a qué conclusión llegaría?

Solución

El ajuste de Bonferroni produciría $P = (4)(0.03) = 0.12$. Esto no es probablemente una evidencia fuerte para concluir que la formulación C es eficaz. Debido a que el P -valor original era pequeño, es probable que sería deseable no querer abandonar aún más la formulación C.

El ajuste de Bonferroni es conservador; en otras palabras, el P -valor que produce nunca es menor que el P -valor verdadero. En este sentido, cuando el P -valor con el ajuste de Bonferroni es pequeño, con seguridad la hipótesis nula puede rechazarse. Desafortunadamente, como se muestra en el ejemplo 6.39, hay muchas ocasiones en las que el P -valor original es bastante pequeño para suscitar una fuerte sospecha de que una hipótesis nula podría ser falsa, pero el ajuste de Bonferroni no permite rechazar la hipótesis.

Cuando el P -valor con el ajuste de Bonferroni es demasiado grande para rechazar una hipótesis nula, aun cuando el P -valor original sea bastante pequeño para sospechar que la hipótesis es efectivamente falsa, con frecuencia lo mejor es reexaminar la hipótesis que parece falsa mediante datos de un nuevo experimento. Si el P -valor es de nuevo pequeño, esta vez sin pruebas múltiples, ello proporciona evidencias reales contra la hipótesis nula.

A los procesos industriales reales se les da seguimiento mediante muestreo y con la comprobación del resultado del proceso para verificar si se satisfacen las especificaciones. Con frecuencia, el resultado parece alejado de las especificaciones. Pero en estos casos, ¿cómo se sabe si el proceso realmente funciona mal (fuera de control) o si el resultado es un error de tipo I? Ésta es una versión del problema de pruebas múltiples que ha recibido mucha atención. El tema de control de calidad estadístico (véase el capítulo 10) está dedicado, en gran parte, a encontrar las maneras para superar el problema de pruebas múltiples.

Ejercicios para la sección 6.14

1. Se prueban seis ajustes diferentes en una máquina con el fin de comprobar si ninguno reducirá la proporción de piezas defectuosas. Para cada ajuste se prueba una adecuada hipótesis nula para verificar si la proporción de piezas defectuosas se ha reducido. Los seis P -valores son 0.34, 0.27, 0.002, 0.45, 0.03 y 0.19.
 - a) Determine el P -valor con el ajuste de Bonferroni para el ajuste cuyo P -valor es 0.002. ¿Puede concluir que este ajuste reduce la proporción de piezas defectuosas? Explique.
 - b) Determine el P -valor con el ajuste de Bonferroni para el ajuste cuyo P -valor es 0.03. ¿Puede concluir que este

ajuste reduce la proporción de partes defectuosas? Explique.

2. Se operan cinco variaciones diferentes de un proceso para fabricar pernos con el propósito de verificar si alguna puede aumentar la media de la fuerza de ruptura sobre el proceso actual. Los P -valores son 0.13, 0.34, 0.03, 0.28 y 0.38. De las siguientes elecciones, ¿cuál será la siguiente opción que se aplicará?
 - i) Instrumentar el proceso cuyo P -valor fue 0.03, ya que se realiza mejor.
 - ii) Debido a que ninguno de los procesos tenía P -valores con el ajuste de Bonferroni menores de 0.05, se debe continuar con el proceso actual.
 - iii) Repetir el proceso cuyo P -valor era 0.03 para ver si permanece pequeño en ausencia de pruebas múltiples.
 - iv) Repetir las cinco variaciones otra vez, para ver si ninguna de ellas produce un P -valor pequeño la segunda vez.
3. Se están probando veinte formulaciones de una capa con el fin de comprobar si alguna de ellas reduce el desgaste de los engranes. Con la idea de que el P -valor con el ajuste de Bonferroni para una formulación sea 0.05, ¿cuál debe ser el P -valor original?
4. Se han estado probando cinco nuevos aditivos de pintura con el fin de verificar si alguno puede reducir la media del tiempo de secado del valor actual de 12 minutos. Se pintaron diez muestras con cada una de los nuevos tipos de pintura y se midieron los tiempos de secado (en minutos). Los resultados fueron:

	Aditivo				
	A	B	C	D	E
1	14.573	10.393	15.497	10.350	11.263
2	12.012	10.435	9.162	7.324	10.848
3	13.449	11.440	11.394	10.338	11.499
4	13.928	9.719	10.766	11.600	10.493
5	13.123	11.045	11.025	10.725	13.409
6	13.254	11.707	10.636	12.240	10.219
7	12.772	11.141	15.066	10.249	10.997
8	10.948	9.852	11.991	9.326	13.196
9	13.702	13.694	13.395	10.774	12.259
10	11.616	9.474	8.276	11.803	11.056

Para cada aditivo realice una prueba de hipótesis de la hipótesis nula $H_0: \mu \geq 12$ contra la alternativa $H_1: \mu < 12$. Puede suponer que cada población es aproximadamente normal.

- a) ¿Cuáles son los P -valores para las cinco pruebas?
 - b) Con base en los resultados, ¿cuál de las tres conclusiones siguientes parece más adecuada? Explique su respuesta.
 - i) Al menos uno de los nuevos aditivos mejora el resultado.
 - ii) Ninguno de los nuevos aditivos mejora el resultado.
 - iii) Algunos de los nuevos aditivos mejoran el resultado, pero las pruebas no son concluyentes.
5. A diario, durante 200 días, un ingeniero responsable de la calidad muestrea 144 fusibles tasados en 15 A y mide el amperaje con el que se queman. Realiza una prueba de hipótesis de $H_0: \mu = 15$ contra $H_1: \mu \neq 15$, donde μ es la media del amperaje con que se queman los fusibles fabricados ese día.
- a) Durante diez de los 200 días, H_0 es rechazada a un nivel de 5%. ¿Esto último proporciona la evidencia concluyente de que la media del amperaje de quemado era diferente de 15 A en al menos uno de los 200 días? Explique.
 - b) ¿Sería la respuesta en el inciso (a) diferente si H_0 hubiera sido rechazada en 20 de los 200 días? Explique.

6.15 Uso de la simulación para realizar pruebas de hipótesis

Si $X_1, \dots, X_n$ son variables aleatorias independientes, con desviaciones estándar conocidas $\sigma_1, \dots, \sigma_n$, y $U = U(X_1, \dots, X_n)$ es una función de $X_1, \dots, X_n$, entonces el método de propagación del error (véase el capítulo 3) se puede utilizar para calcular la desviación estándar, o la incertidumbre, en U . Si además las variables aleatorias $X_1, \dots, X_n$ están distribuidas normalmente, con frecuencia, pero no siempre, se da el caso de que U también está distribuida aproximadamente en forma normal. En estos casos se pueden realizar pruebas de hipótesis sobre la media μ_U de U . Para determinar si U está distribuida normalmente se puede utilizar la simulación. El método es similar al descrito en la sección 5.8.

Se muestra con un ejemplo. Sea R la medición del radio de un cilindro, y H la medición de la altura. Suponga que ambas mediciones no tienen sesgos y están normalmente distribuidas. Sea $V = \pi R^2 H$ la medida del volumen del cilindro que se calcula de R y H . Ahora suponga que $R = 4.8$ cm, $H = 10.1$ cm, y las incertidumbres (desviaciones estándar) son $\sigma_R = 0.1$ cm y $\sigma_H = 0.2$ cm. El volumen medido es $V = \pi(4.8^2)(10.1) = 731.06$ cm³. Suponga que se desea determinar si se puede concluir que el volumen verdadero del cilindro es mayor a 700 cm³. Sea μ_V la media de V . Dado que R y H no tienen sesgos, con incertidumbres bastante pequeñas, V está cercanamente sin sesgos (véase el análisis en la p. 173), por lo que μ_V está cerca del volumen verdadero del cilindro. Por tanto, se puede tratar la pregunta respecto del volumen verdadero realizando una prueba de hipótesis $H_0: \mu_V \leq 700$ contra $H_1: \mu_V > 700$.

Se inicia mediante el método de la propagación del error para calcular la incertidumbre en V :

$$\begin{aligned}\sigma_V &= \sqrt{\left(\frac{\partial V}{\partial R}\right)^2 \sigma_R^2 + \left(\frac{\partial V}{\partial H}\right)^2 \sigma_H^2} \\ &= \sqrt{(2\pi RH)^2 \sigma_R^2 + (\pi R^2)^2 \sigma_H^2} \\ &= \sqrt{[2\pi(4.8)(10.1)]^2 (0.1)^2 + [\pi(4.8^2)]^2 (0.2)^2} \\ &= 33.73\end{aligned}$$

Ahora si el volumen medido V está distribuido normalmente, se puede proceder de la siguiente manera: bajo H_0 , $V \sim N(700, 33.73^2)$. Se observa el valor $V = 731.06$. El P -valor para $H_0: \mu_V \leq 700$ es $P(V \geq 731.06)$ donde la probabilidad se calcula bajo la suposición de que $V \sim N(700, 33.73^2)$. El puntaje z es $(731.06 - 700)/33.73 = 0.92$, y el P -valor es 0.18.

La validez de la prueba apenas realizada depende de la suposición de que V está normalmente distribuido. Se comprueba esta suposición con una simulación. Se genera, primero, un gran número N de valores $R_1^*, \dots, R_N^*$ para la medición de los radios. Se sabe que las mediciones de los radios están normalmente distribuidas con desviación estándar $\sigma_R = 0.1$. No se conoce la media de las mediciones del radio, que es igual al radio verdadero, pero se puede aproximar a ésta con el valor observado 4.8. Por tanto, se genera $R_1^*, \dots, R_N^*$ de la distribución $N(4.8, 0.1^2)$. Similarmente, se genera $H_1^*, \dots, H_N^*$ de la distribución $N(10.1, 0.2^2)$. Después se calculan mediciones de volumen simuladas $V_i^* = \pi(R_i^*)^2 H_i^*$. Una gráfica de probabilidad normal de V_i^* se puede, entonces, utilizar para determinar si V es aproximadamente normal.

La figura 6.30 presenta una gráfica de probabilidad normal para una muestra de mil valores de V_i^* . Se satisface la suposición de normalidad. El P -valor de 0.18 es válido.

Por ultimo, se hace la observación de que si se satisface la suposición de normalidad, la desviación estándar muestral V_i^* se puede utilizar en lugar del valor 33.73 que se calculó con la propagación del error. En la muestra de mil que se generó, la desviación estándar de muestra era de 31.67, que estaba razonablemente cerca del valor calculado con la propagación del error.

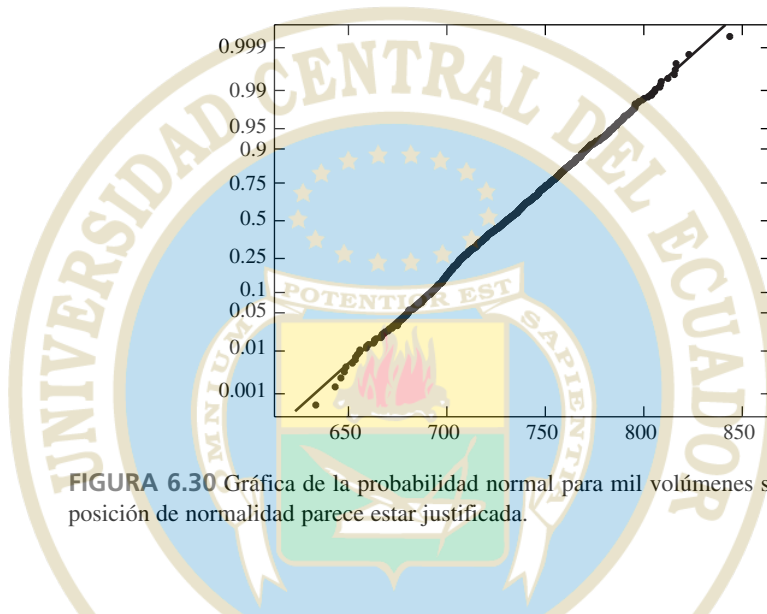


FIGURA 6.30 Gráfica de la probabilidad normal para mil volúmenes simulados. La suposición de normalidad parece estar justificada.

Pruebas de hipótesis con intervalos de confianza bootstrap

Dado un intervalo de confianza de nivel $100(1 - \alpha)\%$ para un parámetro, tal como una media poblacional μ , se puede rechazar la hipótesis nula a un nivel de $100\alpha\%$ de que el parámetro sea igual para cualquier valor dado fuera del intervalo, y no se puede rechazar la hipótesis nula a un nivel de $100\alpha\%$ de que el parámetro sea igual para cualquier valor dado dentro del intervalo (véase el análisis al inicio de la p. 381). Esta idea se puede aplicar a un intervalo de confianza bootstrap para construir una prueba de hipótesis de nivel fijo. Se presenta un ejemplo.

Ejemplo 6.40

En la sección 5.8 un intervalo de confianza de 95% aproximado para la media del millaje, en millas/galón, de una población de camiones se encontró, con un método bootstrap, que era de (4.7643, 6.4757). ¿Se puede concluir a un nivel de 5% que la media poblacional del millaje difiere de 5 mi/galón? ¿De 7 mi/galón?

Solución

Un intervalo de confianza de 95%, si se calculó mediante bootstrap u otros medios, contiene los valores que no son rechazados a un nivel de 5%. Por tanto, se concluye que a un nivel de 5% la media poblacional difiere de 7 mi/galón, pero no se puede concluir que en ese nivel difiere de 5 mi/galón.

Pruebas aleatorias

Las pruebas aleatorias también se llaman pruebas de permutación; estaban entre los primeros métodos desarrollados para probar la diferencia entre dos medias poblacionales. Mientras no requieran ninguna suposición acerca de la distribución de los datos, implican muchos cálculos y no eran realmente viables hasta hace más o menos poco tiempo. Se presenta un ejemplo.

Una científica quiere determinar si la producción de lechuga se incrementará con el uso de un fertilizante con mayor contenido de nitrógeno. Dirige un experimento que implica 20 parcelas. Se eligen aleatoriamente diez de éstas para tratarse con el fertilizante A, que tiene un bajo contenido de nitrógeno. Las otras diez parcelas se tratan con el fertilizante B, que tiene un mayor contenido de nitrógeno.

La siguiente tabla presenta el tratamiento aplicado (A o B) para cada parcela y la producción, según la cantidad de plantas de lechuga cosechadas.

	Número de parcela																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Tratamiento	A	A	B	B	A	A	A	B	B	A	A	B	A	A	B	B	B	B	A	B
Cosecha	145	142	144	141	142	155	143	157	152	143	103	151	150	148	150	162	149	158	144	151

La hipótesis nula es que no hay ninguna diferencia entre los fertilizantes con respecto a la cosecha; en otras palabras, la cosecha para cada parcela habría sido la misma sin que importara el tipo de fertilizante que había recibido. Por ejemplo, la cosecha en la parcela 1 habría sido de 145, ya sea que se hubieran usado los fertilizantes tipos A o B. Si H_0 es verdadera, entonces las 20 producciones observadas son constantes y las producciones relacionadas con el fertilizante B son una muestra aleatoria simple de diez de estas 20 cosechas constantes. Denote la media de las diez cosechas asociadas con el fertilizante A por $\bar{A}$, y la media de las diez cosechas asociadas con el fertilizante B por $\bar{B}$. Debido a que el interés principal en el experimento es determinar si el fertilizante B aumenta la cosecha, un estadístico de prueba razonable es la diferenciar $\bar{B} - \bar{A}$. El valor observado de este estadístico es $151.5 - 141.5 = 10.0$. Entre mayor sea el valor del estadístico de prueba, más fuerte es la evidencia contra H_0 . La fuerza de la evidencia se mide por medio del P -valor. Ahora se analiza cómo calcular el P -valor.

El experimento implica una elección aleatoria de diez parcelas de las 20 que reciben el fertilizante A. En general, el número de elecciones diferentes de k artículos que serán seleccionados de un grupo de n elementos se denota por $\binom{n}{k}$ y está dado por (véase la ecuación 2.12 de la sección 2.2 para su deducción)

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Por tanto, el número de elecciones posibles para estas diez parcelas es $\binom{20}{10} = 184\,756$. Esto último significa que hay tal cantidad de maneras de que el experimento pudiera tener un resultado; el experimento verdadero consiste en observar una de ellas elegida aleatoriamente. La elección que se tomó en realidad proporcionó un valor de $\bar{B} - \bar{A} = 10$ para el estadístico de prueba. Debido a que, bajo H_0 , las producciones no dependen de qué tipo de fertilizante se utilizó, se podría, en principio, calcular el valor del estadístico de prueba $\bar{B} - \bar{A}$ para cada uno de los 184 756 resultados posibles del experimento. El P -valor es la probabilidad, bajo H_0 , de que el estadístico de prueba tenga un valor igual o mayor a diez. Esta probabilidad es igual a la proporción de los 184 756 resultados posibles del experimento para los cuales $\bar{B} - \bar{A} \geq 10$. La tabla 6.6 presenta un listado parcial de los resultados posibles del experimento.

TABLA 6.6 Resultados posibles del experimento aleatorio

Resultado	Cosechas asignadas a A	Cosechas asignadas a B	$\bar{A}$	$\bar{B}$	$\bar{B} - \bar{A}$
1	103 141 142 142 143 143 144 144 145 148	149 150 150 151 151 152 155 157 158 162	139.5	153.5	14.0
2	103 141 142 142 143 143 144 144 145 149	148 150 150 151 151 152 155 157 158 162	139.6	153.4	13.8
⋮	⋮	⋮	⋮	⋮	⋮
184 755	148 150 150 151 151 152 155 157 158 162	103 141 142 142 143 143 144 144 145 149	153.4	139.6	-13.8
184 756	149 150 150 151 151 152 155 157 158 162	103 141 142 142 143 143 144 144 145 148	153.5	139.5	-14.0

El P -valor exacto se puede encontrar cuando se completa la tabla 6.6 y después al determinar la proporción de resultados para los cuales $\bar{B} - \bar{A} \geq 10$. Este procedimiento se llama **prueba aleatoria**, o **prueba de permutación**. Calcular el P -valor exacto es una tarea intensiva, aun para una computadora. El método más fácil, bueno en la práctica, es trabajar con un conjunto de resultados generados aleatoriamente. Esto último se hace cuando se genera un número grande (mil o más) de subconjuntos elegidos aleatoriamente entre diez producciones asignadas al tratamiento A. Cada subconjunto elegido corresponde a uno de los resultados posibles del experimento, y para cada subconjunto se calcula el valor del estadístico de prueba. El P -valor se aproxima con la proporción de resultados elegidos aleatoriamente para los cuales el valor del estadístico de prueba es igual o mayor al valor observado de diez.

La tabla 6.7 presenta los primeros cinco y los últimos de mil resultados elegidos aleatoriamente para el experimento. De los primeros cinco resultados, ninguno de ellos tiene valores de $\bar{B} - \bar{A}$ mayores que o iguales a diez, por lo que la estimación del P -valor basado en estos cinco es $0/5 = 0$. Por supuesto, cinco resultados no son suficientes para obtener una conclusión confiable. De todo el conjunto de los mil resultados, solamente nueve tenían valores de $\bar{B} - \bar{A}$ mayores o iguales a diez. Por tanto, se calcula que el P -valor es de 0.009, bastante pequeño para rechazar de manera tajante la hipótesis nula de que no hay ninguna diferencia entre los fertilizantes. Parece razonable concluir que el fertilizante tipo B tiende a aumentar la producción.

TABLA 6.7 Mil resultados simulados del experimento aleatorio

Resultado	Cosechas asignadas a A	Cosechas asignadas a B	$\bar{A}$	$\bar{B}$	$\bar{B} - \bar{A}$
1	157 151 144 150 142 150 155 144 143 141	145 148 142 143 103 152 158 149 162 151	147.70	145.30	-2.40
2	143 103 158 151 142 151 155 150 148 141	142 144 149 144 143 162 157 150 152 145	144.20	148.80	4.60
3	162 158 144 141 148 155 103 143 144 157	143 150 142 152 145 150 142 149 151 151	145.50	147.50	2.00
4	145 151 143 141 150 142 162 148 149 158	144 155 157 103 152 150 144 151 143 142	148.90	144.10	-4.80
5	152 148 144 142 157 155 162 103 150 151	145 150 158 149 144 143 141 143 151 142	146.40	146.60	0.20
⋮	⋮	⋮	⋮	⋮	⋮
1 000	144 152 143 155 142 148 143 145 158 151	144 103 149 142 150 162 150 141 151 157	148.10	144.90	-3.20

Los experimentos aleatorios, como el recién descrito, tienen un papel muy importante en las investigaciones científicas y se analizan con más profundidad en el capítulo 9. Cuando ningún dato atípico está presente, se ha mostrado que la prueba t de Student para la diferencia entre medias (véase la sección 6.7) proporciona una buena aproximación a la prueba aleatoria cuando se están comparando dos tratamientos. Los datos de los experimentos aleatorios se pueden tratar, en general, como si consistieran de muestras aleatorias de poblaciones diferentes; éste es el enfoque que se tomará en el capítulo 9. El trabajo de Freedman, Pisani y Purves (1998) contiene un buen análisis de este tema. A veces también las pruebas de rango (véase la sección 6.9) se usan para estos experimentos.

Las pruebas aleatorias se pueden utilizar en algunos casos cuando los datos consisten de dos muestras de dos poblaciones, que es el caso que se analiza en la sección 6.7. Por tanto, las pruebas aleatorias pueden ser una alternativa a la prueba t para la diferencia entre medias cuando hay datos atípicos.

Se puede encontrar más información sobre pruebas aleatorias en Efron y Tibshirani (1993).

Uso de la simulación para calcular la potencia

Para algunas pruebas es difícil calcular la potencia con una fórmula; en estos casos, con frecuencia se puede utilizar la simulación para calcularla. A continuación se muestra un ejemplo.

Ejemplo

6.41

Se está desarrollando un nuevo tipo de soldadura. Si se demuestra concluyentemente que la media de la dureza de fractura de la soldadura es mayor de 20 pies-lb, la soldadura se usará en cierta aplicación. Suponga que la dureza está distribuida normalmente con desviación es-

tándar igual a 4 pies·lb. Se harán seis soldaduras, y se medirá la dureza de fractura de cada una. Se hará una prueba t de Student de la hipótesis nula $H_0: \mu \leq 20$ contra $H_1: \mu > 20$. Si la prueba se realiza en un nivel de significancia de 5%, ¿cuál es la potencia de la prueba si la media verdadera de la dureza es 25 pies·lb?

Solución

Sean $X_1, \dots, X_6$ las seis durezas de la muestra, y s la desviación estándar muestral. Ésta es una muestra de una distribución $N(25, 16)$. El estadístico de prueba es $T = (\bar{X} - 20)/(s/\sqrt{6})$. Bajo H_0 , esta estadística tiene una distribución t de Student con cinco grados de libertad. La hipótesis nula será rechazada si el valor del estadístico de prueba es mayor que $t_{5,0.05} = 2.015$. Por tanto, la potencia, es igual a $P(T > 2.015)$. No es fácil calcular esta probabilidad directamente, ya que la hipótesis nula es falsa, por lo que T no tiene la distribución t de Student. Sin embargo, se puede estimar con un experimento de simulación.

Se generarán diez mil muestras $X_{1i}^*, \dots, X_{6i}^*$, cada uno de la distribución $N(25, 16)$. Para cada muestra se calculará la media muestral $\bar{X}_i^*$, la desviación estándar muestral s_i^* , y el estadístico de prueba $T_i^* = (\bar{X}_i^* - 20)/(s_i^*/\sqrt{6})$. Debido a que cada muestra simulada es extraída de la misma distribución, como las durezas verdaderas de las soldaduras, cada muestra simulada es estadísticamente equivalente a una muestra de soldadura verdadera. Por tanto, sólo se puede estimar la potencia cuando se calcula la proporción de muestras simuladas para las que se rechaza la hipótesis nula, esto es, para la cual el valor del estadístico de prueba es mayor a 2.015. La tabla 6.8 presenta los resultados para las diez primeras muestras y las últimas. La columna derecha contiene un “1” si el valor del estadístico de prueba es mayor a 2.015, y un “0” si la situación es otra.

La hipótesis nula se rechaza para nueve de las diez primeras muestras. Si se basaron los resultados en éstas, se estimaría la potencia de 0.9. Por supuesto, diez muestras no son suficientes. De las diez mil muestras, la hipótesis nula fue rechazada para 8 366 de ellas. La estimación de la potencia es, por tanto, 0.8366.

TABLA 6.8 Datos simulados para el ejemplo 6.41

i	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*	X_6^*	$\bar{X}^*$	s^*	T^*	$T^* > 2.015$
1	23.24	23.78	15.65	25.67	24.08	25.88	23.05	3.776	1.978	0
2	26.51	19.89	20.53	25.03	28.35	28.01	24.72	3.693	3.131	1
3	28.61	28.19	29.48	20.06	30.00	21.19	26.26	4.423	3.465	1
4	22.84	28.69	23.93	27.37	19.51	30.28	25.44	4.046	3.291	1
5	22.36	21.26	26.37	23.61	34.45	29.97	26.34	5.061	3.067	1
6	26.54	28.63	24.79	20.63	25.44	26.69	25.45	2.703	4.940	1
7	24.05	24.42	20.32	23.74	24.14	24.66	23.56	1.615	5.394	1
8	28.38	29.51	23.80	29.05	26.39	23.76	26.81	2.579	6.472	1
9	23.55	21.73	19.57	25.04	22.34	29.71	23.66	3.484	2.570	1
10	29.98	34.65	21.17	28.43	23.43	34.44	28.68	5.559	3.825	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
10 000	30.75	19.99	26.20	22.41	31.53	21.78	25.45	4.862	2.744	1

Ejercicios para la sección 6.15

- Este ejercicio es una continuación del 9 de los ejercicios adicionales del capítulo 3. El artículo "Insights into Present-Day Crustal Motion in the Central Mediterranean Area from GPS Surveys" (M. Anzidei, P. Baldi y cols., en *Geophysical Journal International*, 2001:98-100) notifica mediciones de la velocidad de la corteza terrestre en Zimmerwald, Suiza. La componente de la velocidad en dirección norte se midió de $X = 22.10$, y la componente en dirección este se midió de $Y = 14.30$, donde las unidades son mm/año. Las incertidumbres en las mediciones están dadas como $\sigma_X = 0.34$ y $\sigma_Y = 0.32$.
 - Calcule la velocidad estimada V de la corteza terrestre basado en estas mediciones. Utilice el método de propagación del error para calcular su incertidumbre.
 - Suponiendo que la velocidad esté normalmente distribuida, encuentre el P -valor para la hipótesis $H_0: \mu_V \leq 25$.
 - Suponiendo que las componentes de la velocidad en las direcciones norte y este son independientes y están distribuidas normalmente, genere una adecuada muestra simulada de los valores V^* . ¿Es razonable suponer que V se distribuye aproximadamente normal?
- Un genetista de población está estudiando los genes encontrados en dos ubicaciones diferentes sobre el genoma. Él calcula que la proporción p_1 de organismos que tienen un gen A en la primera ubicación es $\hat{p}_1 = 0.42$, con incertidumbre de $\sigma_1 = 0.049$. Asimismo, calcula la proporción de organismos que tienen un gen B en la segunda posición de $\hat{p}_2 = 0.23$, con incertidumbre $\sigma_2 = 0.043$. Bajo las suposiciones que usualmente se hacen en poblaciones genéticas (equilibrio de Hardy-Weinberg), $\hat{p}_1$ y $\hat{p}_2$ son independientes y están normalmente distribuidos, y la proporción p de organismos que tienen tanto genes A como B se estima con $\hat{p} = \hat{p}_1 \hat{p}_2$.
 - Calcule $\hat{p}$ y utilice la propagación del error para calcular su incertidumbre.
 - Suponga que $\hat{p}$ está normalmente distribuido, encuentre el P -valor para probar $H_0: p \geq 0.10$.
 - Genere una muestra simulada adecuada de valores $\hat{p}^*$. ¿Es razonable suponer que $\hat{p}$ está normalmente distribuido?
- Con referencia al ejercicio 6 de la sección 5.8, sea μ la media poblacional de la fuerza compresiva, en MPa. Considere las siguientes hipótesis nulas:
 - $H_0: \mu = 38.53$
 - $H_0: \mu = 38.35$
 - $H_0: \mu = 38.45$
 - $H_0: \mu = 38.55$
 - Con los datos bootstrap presentados en el ejercicio 6 de la sección 5.8, ¿cuál de estas hipótesis nulas se puede rechazar a un nivel de 5%, si se construye un intervalo de confianza con el método 1 de la p. 358?
 - Utilizando los datos bootstrap presentados en el ejercicio 6 de la sección 5.8, ¿cuál de estas hipótesis nulas se puede rechazar a un nivel de 10%, si se construye un intervalo de confianza mediante el método 1 de la p. 358?
- Con referencia al ejercicio 6 de la sección 5.8, sea μ la media poblacional de la fuerza compresiva, en MPa. Genere mil muestras bootstrap.
 - Utilizando los datos bootstrap que usted generó, ¿cuál de estas hipótesis nulas se puede rechazar a un nivel de 5%, utilizando el método 1 de la p. 358?
 - Utilizando los datos bootstrap que usted generó, ¿cuál de estas hipótesis nulas se puede rechazar a un nivel de 10%, con el método 1 de la p. 358?
 - Si se realiza un experimento bootstrap dos veces con los mismos datos, ¿es necesario que los resultados concuerden? Explique.
- En el ejemplo de la cosecha de lechugas presentado en la p. 464, ¿sería una buena idea utilizar la prueba t descrita en la sección 6.7 para determinar si los fertilizantes difieren en sus efectos sobre la cosecha? ¿Por qué sí o no?
- Se sospecha que utilizar gasolina premium en lugar de regular aumentará el millaje para automóviles con un diseño de motor especial. Se usan 16 automóviles en un experimento aleatorio. De la misma manera se eligen ocho para probarlos con gasolina regular, mientras que los otros ocho se prueban con gasolina premium. Los resultados, en mi/galón, son los siguientes:

Regular:	29.1	27.1	30.8	17.3	27.6	16.3
	28.4	30.2				
Premium:	28.3	32.0	27.4	35.3	29.9	35.6
	30.9	29.7				

 - Bajo la hipótesis nula de que cada automóvil obtendrá el mismo millaje con cualquier tipo de gasolina, ¿cuántos resultados diferentes son posibles para este experimento?
 - Sean $\bar{R}$ y $\bar{P}$ la media de los millajes de muestra para los grupos de gasolina regular y premium, respectivamente. Calcule $\bar{R}$ y $\bar{P}$.

- c) Realice una prueba aleatoria para determinar si se puede concluir que la gasolina premium tiende a incrementar el millaje. Utilice el estadístico de prueba $\bar{P} - \bar{R}$. Genere al menos mil resultados aleatorios y calcule el P -valor.
- d) Utilice la prueba t de Student descrita en la sección 6.7 para probar la hipótesis nula de que la media del millaje utilizando gasolina regular es mayor que o igual a la media del millaje utilizando gasolina premium. ¿Este resultado es confiable? Explique
7. Para los datos de la cosecha de lechuga (p. 464) se supone que las cosechas con el fertilizante tipo A podrían tener una varianza mayor que la varianza con el B.
- a) Calcule las varianzas muestrales s_A^2 y s_B^2 de las cosechas atribuibles a A y B, respectivamente, y el cociente s_A^2/s_B^2 .
- b) Alguien sugiere utilizar la prueba F para este problema en la sección 6.11. ¿Esta es una buena idea? ¿Por qué sí o no?
- c) Realice una prueba aleatoria de $H_0: s_A^2 \leq s_B^2$ contra $H_1: s_A^2 > s_B^2$, utilizando el estadístico de prueba s_A^2/s_B^2 , y un mínimo de mil resultados aleatorios.
(Sugerencia: Proceda de la misma manera que en el ejemplo del texto, pero para cada resultado calcule s_A^2 , s_B^2 , y s_A^2/s_B^2 en lugar de $\bar{A}$, $\bar{B}$, y $\bar{B} - \bar{A}$.)
8. Con referencia al ejercicio 6, realice una prueba aleatoria para determinar si el millaje de gasolina regular tiene una varianza mayor que la varianza del millaje utilizando gasolina premium. Genere al menos mil resultados aleatorios.
9. Se supone que cierto método de tratamiento de agua residual se aplica con el propósito de que la media de su pH sea 7. Se realizarán mediciones de pH en siete muestras de agua residual tratada, y se hará una prueba de hipótesis $H_0: \mu = 7$ contra $H_1: \mu \neq 7$ utilizando la prueba t de Student (sección 6.4). Suponga que la media verdadera es $\mu = 6.5$, las mediciones de pH están normalmente distribuidas con media μ y desviación estándar 0.5, y la prueba se efectúa a un nivel de 5%.
- a) Sean $X_1, \dots, X_7$ las mediciones de pH, sea que $\bar{X}$ denote su media, y sea que s denote la desviación estándar muestral. ¿Para cuáles valores del estadístico de prueba H_0 será rechazada?
- b) Genere diez mil pruebas $X_1^*, \dots, X_7^*$ de la distribución verdadera de las mediciones de pH. Para cada muestra, calcule el estadístico de prueba y determine si H_0 es rechazada. Calcule la potencia de la prueba.
10. Este ejercicio requiere de las ideas de la sección 2.6. En un experimento de dos muestras, cuando cada elemento en una muestra es apareado con un elemento en la otra, la prueba t apareada (sección 6.8) se puede utilizar para probar hipótesis respecto de la diferencia entre dos medias poblacionales. Si se desprecia el hecho de que los datos están apareados, también se puede utilizar la prueba t de dos muestras (sección 6.7). La pregunta que surge es cuál prueba tiene la potencia mayor. El siguiente experimento de simulación se ha diseñado para abordar esta pregunta.
- Sea $(X_1, Y_1), \dots, (X_8, Y_8)$ una muestra aleatoria de ocho pares, con $X_1, \dots, X_8$ extraídos de una población $N(0, 1)$ y $Y_1, \dots, Y_8$ extraídos de una población $N(1, 1)$. Se desea probar $H_0: \mu_X - \mu_Y = 0$ contra $H_1: \mu_X - \mu_Y \neq 0$. Observe que $\mu_X = 0$ y $\mu_Y = 1$, observe que las varianzas de población son iguales, por lo que la diferencia verdadera entre las medias es 1. También observe que las varianzas poblacionales son iguales. Si se hace una prueba a un nivel de significancia de 5%, ¿qué prueba tiene la mayor potencia?
- Sea $D_i = X_i - Y_i$ para $i = 1, \dots, 10$. El estadístico de prueba para la prueba t apareada es $\bar{D}/(s_D/\sqrt{8})$, donde s_D representa la desviación estándar de los D_i (véase la sección 6.8). Su distribución nula es la t de Student con siete grados de libertad. Por tanto, la prueba t apareada rechazará H_0 si $|\bar{D}/(s_D/\sqrt{8})| > t_{7, .025} = 2.365$, por lo que la potencia es $P(|\bar{D}/(s_D/\sqrt{8})| > 2.365)$.
- Para la prueba t de dos muestras cuando las varianzas poblacionales son iguales, el estadístico de prueba es $\bar{D}/(s_p\sqrt{1/8 + 1/8}) = \bar{D}/(s_p/2)$, donde s_p constituye la desviación estándar combinada, que es igual en este caso a $\sqrt{(s_X^2 + s_Y^2)/2}$. (Véase la p. 411. Observe que $\bar{D} = \bar{X} - \bar{Y}$.) La distribución nula es la t de Student con 14 grados de libertad. Por tanto, la prueba t de dos muestras rechazará H_0 si $|\bar{D}/(s_p\sqrt{1/8 + 1/8})| > t_{14, .025} = 2.145$, y la potencia es $P(|\bar{D}/(s_p\sqrt{1/8 + 1/8})| > 2.145)$.
- La potencia de estas pruebas depende de la correlación entre X_i y Y_i .
- a) Genere diez mil muestras $X_{1i}^*, \dots, X_{8i}^*$ de una población $N(0, 1)$ y diez mil muestras $Y_{1i}^*, \dots, Y_{8i}^*$ de una población de $N(1, 1)$. Las variables aleatorias X_{ki}^* y Y_{ki}^* son independientes en este experimento, por lo que su correlación es 0. Para cada muestra, calcule los estadísticos de prueba $\bar{D}^*/(s_D^*/\sqrt{8})$ y $\bar{D}^*/(s_p^*/2)$. Calcule la potencia de cada prueba mediante la proporción de muestras para las cuales los estadísticos de prueba son mayores de su punto crítico (2.365 para la prueba apareada, 2.145 para la prueba de dos muestras). ¿Qué prueba tiene mayor potencia?
- b) Como en el inciso a), genere diez mil muestras $X_{1i}^*, \dots, X_{8i}^*$ de una población $N(0, 1)$. Esta vez, en lugar de generar los valores Y^* de manera independiente, génelos para que la correlación entre X_{ki}^* y Y_{ki}^* sea 0.8. Esto últi-

mo se puede hacer de la siguiente manera: genere diez mil muestras $Z_{1i}^*, \dots, Z_{8i}^*$ de una población $N(0, 1)$, independientemente de los valores de X^* . Después calcule $Y_{ki}^* = 1 + 0.8 X_{ki}^* + 0.6 Z_{ki}^*$. La muestra $Y_{1i}^*, \dots, Y_{8i}^*$ proviene de una población de $N(1, 1)$, y la correlación

entre X_{ki}^* y Y_{ki}^* será de 0.8, lo que significa que valores grandes de X_{ki}^* tienden a estar apareados con valores grandes de Y_{ki}^* , y viceversa. Calcule los estadísticos de prueba y estime la potencia de ambas pruebas, como en el inciso a). ¿Qué prueba tiene una potencia mayor?

Ejercicios adicionales para el capítulo 6

Los ejercicios 1 a 4 describen experimentos que requieren una prueba de hipótesis. Para cada experimento describa la prueba adecuada. Establezca las hipótesis nula y alternativa adecuadas, describa el estadístico de prueba y especifique qué tabla se debe utilizar para determinar el P -valor. Si es pertinente, establezca el número de grados de libertad para el estadístico de prueba.

- Una flota de 100 taxis se divide en dos grupos de 50 automóviles cada uno para averiguar si la gasolina premium reduce gastos de mantenimiento. La gasolina sin plomo premium se utiliza en el grupo A, mientras que la gasolina sin plomo regular se emplea en el B. Se registra el costo total del mantenimiento para cada automóvil durante un año. Se usará la gasolina premium si se demuestra que reduce gastos de mantenimiento.
- Se elige un grupo de 15 nadadores para que participe en un experimento con el fin de comprobar si un nuevo estilo de respiración mejora su resistencia. Se mide que cada nadador recupera su ritmo del pulso después de 20 minutos mediante el viejo estilo de respiración. Los nadadores practican el nuevo estilo durante dos semanas, y luego miden el ritmo del pulso después de 20 minutos con el nuevo estilo. Continuarán utilizando el nuevo estilo de respiración si se demuestra que reduce el tiempo de recuperación del pulso.
- Se está probando un nuevo programa de control de calidad para ver si reducirá la proporción de piezas enviadas defectuosas. Con el antiguo programa, la proporción de piezas defectuosas era de 0.10. Se probarán 200 piezas que superaron la inspección bajo el nuevo programa, y se contará el número de defectos. Se instrumentará el nuevo programa si se muestra que la proporción de defectos es menor que 0.10.
- Se está probando un nuevo material, que se usará en la fabricación de un conducto eléctrico, para determinar si reduce la varianza en la fuerza de compresión sobre el viejo material. La fuerza de compresión se mide para una mues-

tra de 16 conductos del material viejo y una muestra de 20 del nuevo. Si se demuestra que la fuerza compresiva con el material nuevo tiene una varianza menor, se usará el nuevo material.

- Suponga que ha comprado una máquina de llenado para bolsas de dulces que contendrá 16 oz de éstos. Suponga que los pesos de las bolsas llenas están aproximadamente distribuidos en forma normal. Una muestra aleatoria de diez bolsas produce los siguientes datos (en oz):

15.87	16.02	15.78	15.83	15.69	15.81
16.04	15.81	15.92	16.10		

Con base en estos datos, ¿puede concluir que la media del peso llenado es, en realidad, menor que 16 oz?

- Establezca las hipótesis nulas y alternativas adecuadas.
 - Calcule el valor del estadístico de prueba.
 - Determine el P -valor y establezca su conclusión.
- ¿Las respuestas clave para pruebas de elección múltiple se generan aleatoriamente, o se construyen para que sea menos probable que ocurran dos veces consecutivas para la misma respuesta? En el artículo "Seek Whence: Answer Sequences and Their Consequences in Key-Balanced Multiple-Choice Tests" (M. Bar-Hillel y Y. Attali, *The American Statistician*, 2002:299-303) se formuló tal pregunta. Se estudiaron 1 280 preguntas de diez verdaderas pruebas escolares (PE). Suponga que todas las preguntas tenían cinco opciones (en efecto, 150 de ellas tenían sólo cuatro opciones). Encontraron que para 192 preguntas, la opción correcta (A, B, C, D o E) era la misma que la elección correcta para la pregunta anterior. Si las opciones fueran generadas aleatoriamente, entonces la probabilidad de que una pregunta tuviera la misma opción correcta que la anterior sería de 0.20. ¿Puede concluir que las elecciones para las PE no se pueden generar de manera aleatoria?

- Establezca las hipótesis nula y alternativa adecuadas.
- Calcule el valor del estadístico de prueba.
- Determine el P -valor y establezca su conclusión.

7. Un fabricante de automóviles desea comparar las duraciones de dos marcas de neumáticos. Obtiene muestras de seis llantas de cada marca. En cada uno de seis automóviles instala un neumático de cada marca sobre cada rueda del frente. Los carros se conducen hasta que quede solamente 20% de las huellas originales. Las distancias, en millas, para cada neumático se presentan en la tabla siguiente. ¿Puede concluir que hay una diferencia entre la media de las duraciones de las dos marcas de neumáticos?

Automóvil	Marca 1	Marca 2
1	36 925	34 318
2	45 300	42 280
3	36 240	35 500
4	32 100	31 950
5	37 210	38 015
6	48 360	47 800
7	38 200	33 215

- Establezca las hipótesis nula y alternativa adecuadas.
- Calcule el valor del estadístico de prueba.
- Determine el P -valor y establezca su conclusión.

8. Se tomaron 21 mediciones independientes de la dureza (en la escala C de Rockwell) de metal de base de acero HSLA-100, y se hicieron otras 21 mediciones independientes de la dureza de una soldadura producida sobre este metal básico. La desviación estándar de las mediciones hechas sobre el metal básico era de 3.06, y la estándar de las mediciones hechas sobre la soldadura era de 1.41. Suponga que las mediciones son muestras aleatorias independientes de poblaciones normales. ¿Puede concluir que las mediciones hechas sobre el metal básico son más variables que las mediciones realizadas sobre la soldadura?
9. Existe la preocupación de que al aumentar la industrialización se incremente el contenido de minerales en el agua de los ríos. Hace diez años, el contenido de silicio en el agua de cierto río era de 5 mg/l. Recientemente se tomaron 85 muestras de agua del río que tienen una media de contenido de silicio de 5.6 mg/l y desviación estándar de 1.2 mg/l. ¿Puede concluir que el contenido de silicio del agua es mayor hoy que lo que era hace diez años?

10. El artículo "Modeling of Urban Area Stop-and-Go Traffic Noise" (P. Pamanikabud y C. Tharasawatipat, en *Journal*

of Transportation Engineering 1999:152-159) presenta las mediciones del ruido de tránsito, en dBA, de diez lugares en Bangkok, Tailandia. En la siguiente tabla se presentan mediciones, realizadas en cada lugar, tanto en carriles acelerados como desacelerados.

Lugar	Acelerados	Desacelerados
1	78.1	78.6
2	78.1	80.0
3	79.6	79.3
4	81.0	79.1
5	78.7	78.2
6	78.1	78.0
7	78.6	78.6
8	78.5	78.8
9	78.4	78.0
10	79.6	78.4

¿Puede concluir que hay una diferencia en la media de los niveles de ruido entre los carriles acelerados y desacelerados?

11. Una máquina se configura para producir válvulas cuyas longitudes tengan una media de 100 mm y desviación estándar de 0.1 mm. La máquina se traslada a otra ubicación. Se pensaba que el cambio podría haber afectado la calibración para la media de la longitud, pero que es improbable que cambie la desviación estándar. Sea μ la media de la longitud de válvulas producidas después del cambio. Para probar la calibración se rectifica una muestra de 100 válvulas, se medirán sus longitudes, y se hará una prueba de las hipótesis $H_0: \mu = 100$ contra $H_1: \mu \neq 100$.

- Determine la región de rechazo si la prueba es hecha en un nivel de 5%.
- Determine la región de rechazo si la prueba es hecha en un nivel de 10%.
- Si la media muestral de longitudes es 99.97 mm, ¿ H_0 será rechazada a un nivel de 5%?
- Si la media muestral de longitudes es 100.01 mm, ¿ H_0 será rechazado a un nivel de 10%?
- Un punto crítico es 100.015 mm. ¿Cuál es el nivel de la prueba?

12. Se supone que se calibra un proceso de fabricación de hojas de vidrio con el propósito de que la media μ de los espesores de las hojas sea de más de 4 mm. La desviación estándar de los espesores de las hojas está bien aproximada mediante $\sigma = 0.20$ mm. Se medirán los espesores de cada

hoja en una muestra, y se realizará una prueba de hipótesis $H_0: \mu \leq 4$ contra $H_1: \mu > 4$. Suponga que, efectivamente, la media verdadera del espesor es 4.04 mm.

- Si se muestran 100 láminas, ¿cuál es la potencia de una prueba hecha a un nivel de 5%?
- ¿Cuántas láminas se deben probar con el propósito de que una prueba al nivel de 5% tenga una potencia de 0.95?
- Se prueban 100 láminas, ¿en qué nivel se debe hacer la prueba con el propósito de que la potencia sea de 0.90?
- Si se probaron 100 láminas, y la región de rechazo es $\bar{X} \geq 4.02$, ¿cuál es la potencia de la prueba?

13. Una máquina fabrica pernos que, se supone, son de tres pulgadas de longitud. A diario un ingeniero responsable del control de calidad selecciona una muestra aleatoria de 50 pernos de la producción del día, mide sus longitudes y realiza una prueba de hipótesis de $H_0: \mu = 3$ contra $H_0: \mu \neq 3$, donde μ significa la media de la longitud de todos los pernos fabricados ese día. Suponga que la desviación estándar poblacional para las longitudes de los pernos es de 0.1 pulg. Si H_0 es rechazada a un nivel de 5%, la máquina se detendrá y se recalibrará.

- Suponga que en un día específico, la media verdadera de la longitud de pernos es 3 pulg. ¿Cuál es la probabilidad de que la máquina se detenga? (Esto último se llama tasa de falsa alarma.)
- Si la media verdadera de la longitud de los pernos en un día determinado es 3.01 pulgadas, determine la probabilidad de que el equipo será recalibrado.

14. Se montan motores eléctricos sobre cuatro diferentes líneas de producción. Se toman muestras aleatorias de los motores de cada línea y se inspeccionan. Los grupos que pasan y que reprueban la inspección se cuentan en cada línea, con los siguientes resultados:

	Línea			
	1	2	3	4
Pasa	482	467	458	404
Reprueba	57	59	37	47

¿Puede concluir que la tasa de reprobación es diferente entre las cuatro líneas?

15. Con referencia al ejercicio 14, la ingeniera de proceso observa que la muestra de la línea 3 tiene la proporción más

baja de fallas. Utilice el ajuste de Bonferroni para determinar si puede concluir que la proporción de fallas de la población en la línea 3 es menor que 0.10.

Los ejercicios 16 y 17 muestran que los métodos de distribución libre pueden producir resultados engañosos cuando sus suposiciones se incumplen seriamente.

16. Considere las siguientes dos muestras:

X:	0	2	3	4	10	20	40	100	1 000
Y:	-738	162	222	242	252	258	259	260	262

- Muestre que ambas muestras tienen la misma media y varianza.
- Utilice la prueba de la suma del rango de Wilcoxon para probar la hipótesis de que las medias poblacionales son iguales, ¿qué concluye?
- ¿Parece que las suposiciones de la prueba de la suma del rango se satisfacen? Explique por qué sí o no.

17. A veces se piensa en la prueba de la suma del rango como una prueba para las medianas de población. Bajo las suposiciones de igual extensión y forma, las medias de las dos poblaciones son diferentes si y sólo si las medianas son diferentes; por tanto, probar la igualdad de medias poblacionales son también pruebas para la igualdad de las medianas poblacionales. Este ejercicio muestra que cuando estas suposiciones se incumplen seriamente la prueba de la suma del rango puede dar resultados falsos con respecto a la igualdad de las medianas poblacionales. Considere las siguientes dos muestras:

X:	1	2	3	4	5	6	7	
	20	40	50	60	70	80	90	100
Y:	-10	-9	-8	-7	-6	-5	-4	
	20	21	22	23	24	25	26	27

- Compruebe que ambas muestras tienen la misma mediana.
- Calcule el P -valor para una prueba de la suma del rango de dos colas. Si P -valores pequeños proporcionan evidencia contra la hipótesis nula de que las medianas poblacionales son iguales, ¿concluiría que las medianas poblacionales son diferentes?
- ¿Las suposiciones de la prueba de la suma del rango parecen satisfacerse? Explique por qué sí o no.

18. Se está considerando un nuevo proceso de producción para la fabricación de cojinetes de acero inoxidable. Mediciones de los diámetros de muestras aleatorias de cojinetes de viejos y nuevos procesos produjeron los siguientes datos:

Viejo: 16.3 15.9 15.8 16.2 16.1 16.0
 15.7 15.8 15.9 16.1 16.3 16.1
 15.8 15.7 15.8 15.7

Nuevo: 15.9 16.2 16.0 15.8 16.1 16.1
 15.8 16.0 16.2 15.9 15.7 16.2
 15.8 15.8 16.2 16.3

- a) ¿Puede concluir que en un nivel de 5% que un proceso produce una media de los tamaños de los cojinetes diferente de la otra?
- b) ¿Puede concluir que en un nivel de 5% la varianza del nuevo procedimiento es menor que con el procedimiento antiguo?

19. Se consideran dos formulaciones químicas diferentes de combustible de cohete para el empuje pico que se proporciona a un diseño particular de un motor de cohete. Los cocientes empuje/peso (en kilogramos fuerza por gramo) para cada uno de los dos combustibles se miden varias veces. Los resultados son los siguientes:

Combustible A: 54.3 52.9 57.9 58.2 53.4 51.4
 56.8 55.9 57.9 56.8 58.4 52.9
 55.5 51.3 51.8 53.3

Combustible B: 55.1 55.5 53.1 50.5 49.7 50.1
 52.4 54.4 54.1 55.6 56.1 54.8
 48.4 48.3 55.5 54.7

- a) Suponga que la planta de procesamiento de combustible está configurada para producir combustible B y los costos de cambio son altos. Debido a que el aumento del cociente empuje/peso para el combustible cohete es benéfico, ¿cómo se deben establecer las hipótesis nula y alternativa para una prueba útil para la decisión de cambiar al combustible A?
- b) ¿Puede concluir a un nivel de 5% que se debe cambiar al combustible A?

20. Suponga que la Agencia de Protección al Ambiente está en el proceso de dar seguimiento a la calidad del agua, en un estuario grande en el este de Estados Unidos, con la finalidad de medir la concentración de PCB (en partes por miles de millones).

- a) Suponga que una muestra aleatoria de tamaño 80 tiene una media muestral de 1.59 ppm y una desviación estándar muestral de 0.25 ppm. Pruebe la hipótesis, a un nivel de 5%, de que la media de la concentración de PCB en el estuario es menor o igual a 1.50 ppm contra la alternativa que es mayor. ¿ H_0 se rechaza?

- b) Si la media poblacional es 1.6 ppm y la desviación estándar poblacional es 0.33 ppm, ¿cuál es la probabilidad de que la hipótesis nula $H_0: \mu \leq 1.50$ sea rechazada a un nivel de 5%, si el tamaño de muestra es 80?
- c) Si la media de población es 1.6 ppm, y la desviación estándar poblacional, 0.33 ppm, ¿qué tamaño muestral es necesario con el propósito de que la probabilidad de 0.99 de que $H_0: \mu \leq 1.50$ sea rechazada a un nivel de 5%?

21. Dos máquinas se utilizan para empacar detergente de lavandería. Se sabe que los pesos de las cajas están normalmente distribuidos. Cuatro cajas de cada máquina tienen sus contenidos cuidadosamente pesados, con los siguientes resultados (en gramos):

Máquina 1: 1 752 1 757 1 751 1 754
 Máquina 2: 1 756 1 750 1 752 1 746

Un ingeniero desea para probar la hipótesis nula de que las medias de los pesos de las cajas de las dos máquinas son iguales. Decide suponer que las varianzas poblacionales son iguales, el razonamiento es el siguiente:

Las varianzas muestrales son $s_1^2 = 7.00$ para la máquina número 1 y $s_2^2 = 17.33$ para la máquina número 2. El estadístico F para probar la igualdad de las varianzas poblacionales es $F_{3,3} = s_2^2/s_1^2 = 2.48$. El punto 10% superior de la distribución $F_{3,3}$ es 5.39. Debido a que la hipótesis nula especifica que las varianzas son iguales, I determina que el P -valor es mayor que $2(0.10) = 0.20$. Por tanto, I no rechaza la hipótesis nula, y I concluye que las varianzas son iguales.

- a) ¿La prueba F se ha hecho correctamente?
- b) ¿La conclusión está justificada? Explique.

22. El artículo "Valuing Watershed Quality Improvements Using Conjoint Analysis" (S. Farber y B. Griner, en *Ecological Economics*, 2000:63-76) presenta los resultados de una encuesta de correo diseñada para evaluar opiniones acerca del valor de los esfuerzos para mejorar en una línea divisoria degradada de mina de ácido en el occidente de Pensilvania. De 510 personas encuestadas, 347 era hombres. Los datos del censo muestran que 48% de la población objetivo son hombres. ¿Puede concluir que el método de la encuesta empleado en este estudio tiende a sobremuestrear a los hombres? Explique.

23. Los antropólogos pueden calcular la tasa de natalidad de una antigua civilización estudiando la distribución de edad de los esqueletos encontrados en cementerios. Los números

de esqueletos descubiertos en dos de esos sitios, como informa el artículo “Paleoanthropological Traces of a Neolithic Demographic Transition” (J. Bocquet-Appel, en *Current Anthropology*, 2002:637-650), se muestran en la tabla siguiente:

Sitio	Edad de los esqueletos		
	0–4 años	5–19 años	20 años o más
Casa de Moura	27	61	126
Wandersleben	38	60	118

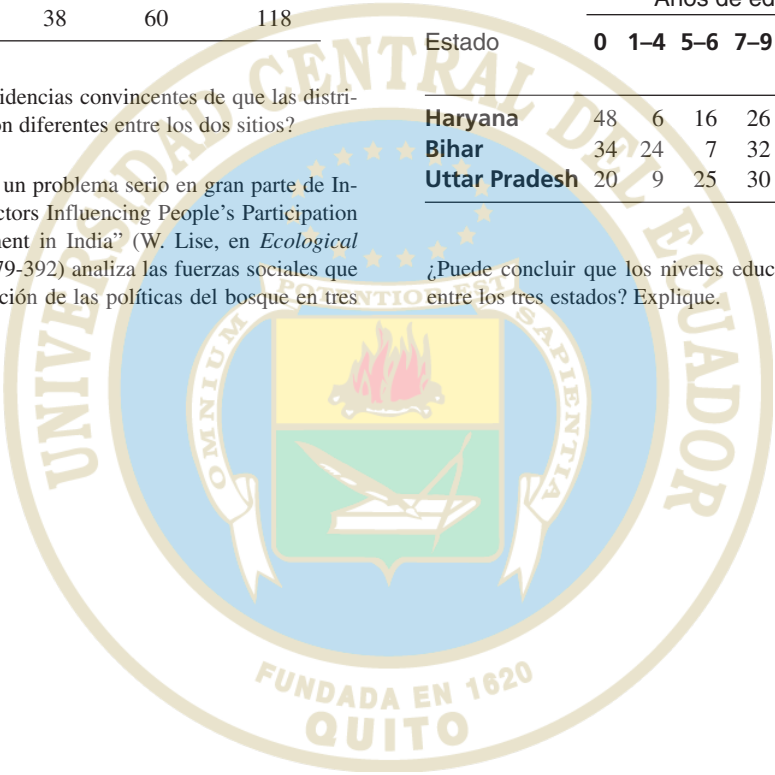
¿Estos datos dan evidencias convincentes de que las distribuciones de edad son diferentes entre los dos sitios?

24. La deforestación es un problema serio en gran parte de India. El artículo “Factors Influencing People’s Participation in Forest Management in India” (W. Lise, en *Ecological Economics*, 2000:379-392) analiza las fuerzas sociales que influyen en la dirección de las políticas del bosque en tres

estados federales indios: Haryana, Bihar y Uttar Pradesh. La calidad de bosque en Haryana está un poco degradada, en Bihar está muy degradada y en Uttar Pradesh está bien conservada. Con el fin de estudiar la relación entre los niveles educacionales y las actitudes hacia la administración del bosque, se examinaron muestras aleatorias de adultos en cada uno de estos estados federales y se investigaron sus niveles de educación. Se registraron los números de adultos en cada uno de algunos niveles educativos. Los datos se presentan en la siguiente tabla.

Estado	Años de educación					
	0	1–4	5–6	7–9	10–11	12 o más
Haryana	48	6	16	26	24	7
Bihar	34	24	7	32	16	10
Uttar Pradesh	20	9	25	30	17	34

¿Puede concluir que los niveles educativos son diferentes entre los tres estados? Explique.



Correlación y regresión lineal simple

Introducción

Con frecuencia, científicos e ingenieros reúnen datos con el propósito de determinar la naturaleza de la relación entre dos cantidades. Por ejemplo, un ingeniero químico puede realizar varias veces un proceso químico para estudiar la relación entre la concentración de cierto catalizador y la producción del proceso. Cada vez que realiza éste, registra la concentración x y la producción y . Por consiguiente, el experimento genera datos **bivariados**; un conjunto de pares ordenados $(x_1, y_1), \dots, (x_n, y_n)$. En muchos casos, al graficar los pares ordenados que se generan en un experimento científico se encontrarán, aproximadamente, a lo largo de una línea recta. En estos casos, los datos son útiles para calcular la ecuación de una recta. Aquélla se puede utilizar con muchos propósitos; por ejemplo, el catalizador contra la producción del experimento que se acaba de describir, esto podría ser útil para pronosticar la producción y , que se obtendrá la próxima vez que se opere el proceso con una concentración específica de un catalizador x .

Los métodos de correlación y la regresión lineal simple constituyen el tema de este capítulo; ambos se utilizan para: analizar datos bivariados con la finalidad de determinar si un ajuste lineal es adecuado, calcular la ecuación de la recta si es adecuado, y usar esta ecuación con el fin de hacer inferencias respecto de la relación entre ambas cantidades.

7.1 Correlación

Una de las primeras aplicaciones de la estadística consistió en estudiar la variación de las características físicas en poblaciones humanas. Con ese fin, los estadígrafos crearon una cantidad llamada **coeficiente de correlación** como forma de describir la relación tan cercana entre

dos características físicas. El primer coeficiente de correlación publicado fue producto del trabajo de sir Francis Galton, quien, en 1888, midió la estatura y longitud del antebrazo, respectivamente, de 348 hombres adultos. (En realidad, midió la distancia del codo a la punta del dedo cordial, lo que se conocía como un cúbito.) Si se denota la estatura del i -ésimo hombre mediante x_i , y la longitud de su antebrazo como y_i , entonces los datos de Galton consistían de 348 pares ordenados (x_i, y_i) . La figura 7.1 presenta una recreación simulada de estos datos, con base en una tabla que construyó Galton.

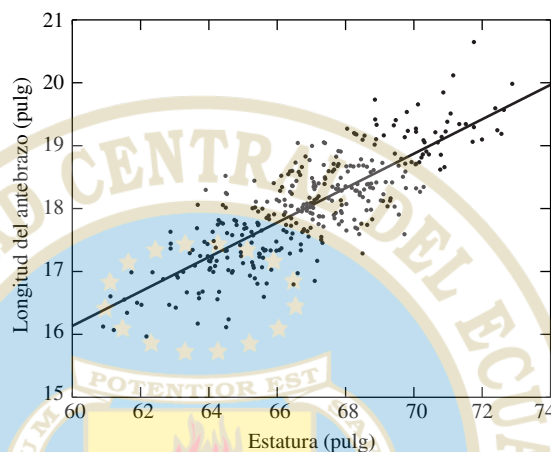


FIGURA 7.1 Estatura y longitud de los antebrazos de 348 hombres.

Los puntos tienden a inclinarse hacia arriba y a la derecha, lo que indica que los hombres más altos tienden a tener antebrazos más largos. Se dice que hay una **asociación positiva** entre la estatura y la longitud del antebrazo. La pendiente es casi constante en toda la gráfica, esto indica que los puntos están agrupados alrededor de una línea recta. La línea sobrepuesta sobre la gráfica representa una recta especial conocida como **recta de mínimos cuadrados**. Ésta es la que ajusta mejor tales datos, en el sentido que se describirá en la sección 7.2. También ahí se aprenderá cómo calcular la recta de mínimos cuadrados.

La figura 7.2 presenta los resultados de un estudio de la relación entre la media de la temperatura diaria y la media de la humedad diaria en un lugar cerca de Riverside, California, durante el reciente invierno. Nuevamente los puntos están agrupados alrededor de la recta de mínimos cuadrados. Esta última tiene una pendiente negativa, ello indica que los días con mayor humedad tienden a tener temperaturas más bajas.

El grado en que los puntos en un diagrama similar tienden a agruparse alrededor de una recta refleja la fuerza de la relación lineal entre x y y . La impresión visual de una gráfica de puntos puede ser engañosa respecto de lo anterior, debido al cambio de escala de los ejes, lo que puede hacer que el agrupamiento parezca más junto o más disperso. En consecuencia, se define el **coeficiente de correlación**, que es una medida numérica de la fuerza de la relación lineal entre dos variables. Este coeficiente se denota con la literal r .

Sean $(x_1, y_1), \dots, (x_n, y_n)$ los n puntos del diagrama de dispersión. Para calcular la correlación, primero se deducen las medias y las desviaciones estándar de las x y de las y , que se representan mediante $\bar{x}$, $\bar{y}$, s_x y s_y . Después se convierte cada x y cada y a las unidades estándar; en otras palabras, se calculan los puntajes z : $(x_i - \bar{x})/s_x$, $(y_i - \bar{y})/s_y$. El coeficiente de co-

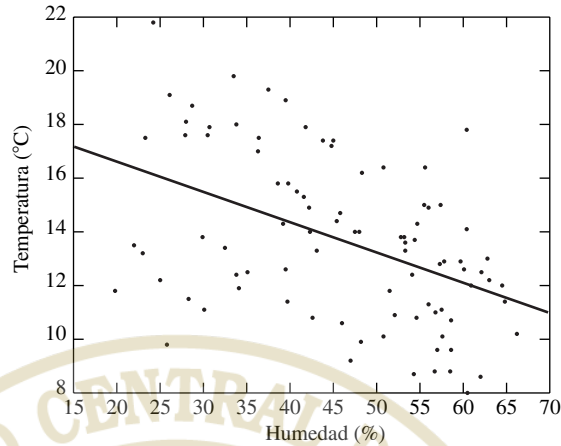


FIGURA 7.2 Humedad (en %) y temperatura (en °C) para los días de un invierno reciente en Riverside, California.

La relación representa el promedio de los productos de los puntajes z , excepto que se divide entre $n - 1$ en lugar de n :

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) \quad (7.1)$$

Se puede reescribir la ecuación 7.1 en una forma que a veces es útil. Al sustituir s_x con $\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$ y a s_y con $\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)}$, se obtiene

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (7.2)$$

En principio, el coeficiente de correlación se puede calcular para cualquier conjunto de puntos. Éstos, en muchos casos, constituyen una muestra aleatoria de una población de puntos. En dichos casos el coeficiente de correlación con frecuencia se llama **correlación muestral**, y es una estimación de la correlación poblacional. (En la sección 2.6 se analizaron formalmente correlaciones poblacionales; en forma intuitiva, se supone que la población consta de un gran número finito de puntos; en esta tesitura, la correlación poblacional significa la cantidad que se calcula mediante la ecuación 7.2 en toda una población, con las medias muestrales sustituidas por las medias poblacionales.) La correlación muestral es útil en la construcción de intervalos de confianza y la realización de pruebas de hipótesis acerca de la correlación poblacional; éstas se tratarán después en esta sección. Se puntualiza que el coeficiente de correlación también se puede utilizar para medir la fuerza de una relación lineal en muchos casos donde los puntos no son una muestra aleatoria de una población; en la sección 7.2 se observa el análisis del coeficiente de determinación.

Es un hecho matemático que el coeficiente de correlación se encuentra siempre entre -1 y 1 . Valores positivos del coeficiente de correlación indican que la recta de mínimos cuadrados tiene pendiente positiva, ello significa que valores mayores de una variable están asociados con valores mayores que los demás. Por el contrario, valores negativos del coeficiente

de correlación indican que la recta de mínimos cuadrados tiene pendiente negativa, lo anterior muestra que valores mayores de una variable están relacionados con valores menores que los demás. En este contexto, valores del coeficiente de correlación cercanos a 1 o a -1 indican fuerte relación lineal; asimismo, valores cercanos a 0 indican débil relación lineal. El coeficiente de correlación es igual a 1 (o a -1) sólo cuando los puntos en el diagrama de dispersión están exactamente sobre una recta de pendiente positiva (o negativa); en otras palabras, cuando hay una relación lineal perfecta. Como nota técnica, tenemos que si los puntos se hallan exactamente sobre una recta horizontal o vertical, el coeficiente de correlación está indeterminado, ya que una de las desviaciones estándar es igual a cero. Por último, un poco de terminología: siempre que $r \neq 0$, se dice que x y y están **correlacionados**. Si $r = 0$, se dice que x y y **no están correlacionados**.

La correlación entre estatura y longitud del antebrazo, en la figura 7.1, es 0.80. La que corresponde a temperatura y humedad en la figura 7.2 es -0.46 . Las figuras 7.3 y 7.4 de las páginas 479 y 480 muestran algunos ejemplos de gráficas de dispersión con diferentes correlaciones. En cada gráfica tanto x como y tienen media 0, y desviación estándar 1. Todas las gráficas están dibujadas en la misma escala.

Cómo funciona el coeficiente de correlación

¿Por qué la fórmula (ecuación 7.1) para el coeficiente de correlación r mide la fuerza de asociación lineal entre dos variables? La figura 7.5 de la página 481 muestra cómo funciona el coeficiente de correlación. En este diagrama de puntos, el origen está colocado en el punto de los promedios $(\bar{x}, \bar{y})$. Por tanto, en el primer cuadrante, los puntajes z $(x_i - \bar{x})/s_x$ y $(y_i - \bar{y})/s_y$ son positivos, por lo que su producto también lo es. Por consiguiente, cada punto en el primer cuadrante aporta una cantidad positiva a la suma de la ecuación (7.1). En el segundo cuadrante los puntajes z para las coordenadas x de los puntos son negativos, mientras que los z para las coordenadas y , positivos. En consecuencia, los productos de los puntajes z son negativos, por lo que cada punto en el segundo cuadrante aporta una cantidad negativa a la suma de la ecuación (7.1). De forma semejante, los puntos en el tercer cuadrante aportan cantidades positivas, y los del cuarto cuadrante, cantidades negativas. Evidentemente, en la figura 7.5 hay más puntos en el primero y tercer cuadrantes que en el segundo y cuarto, así la correlación será positiva. Si la gráfica tuviera una pendiente negativa, habría más puntos en el segundo y cuarto cuadrantes, y el coeficiente de correlación sería negativo.

El coeficiente de correlación es un número puro

En cualquier muestra $x_1, \dots, x_n$, la media $\bar{x}$ y la desviación estándar s_x tienen las mismas unidades que $x_1, \dots, x_n$. Por esta razón los puntajes z $(x_i - \bar{x})/s_x$ son números puros. Puesto que el coeficiente de correlación r constituye el promedio de productos de puntajes z , también es un número puro, sin unidades. Este hecho es crucial para la utilidad de r . Por ejemplo, las unidades para las coordenadas x y y en la figura 7.1 son para ambas pulgadas, mientras que las unidades correspondientes en la figura 7.2 son porcentajes y grados Celsius. Si los coeficientes de correlación para las dos gráficas tuvieran unidades diferentes, sería imposible comparar sus valores para determinar la gráfica que presenta la relación lineal más fuerte. Pero debido a que los coeficientes de correlación son números puros, son directamente comparables; se concluye que la relación entre la estatura de los hombres y la longitud de su antebrazo en la figura 7.1 es linealmente más fuerte que la relación entre la temperatura y la humedad de la figura 7.2.

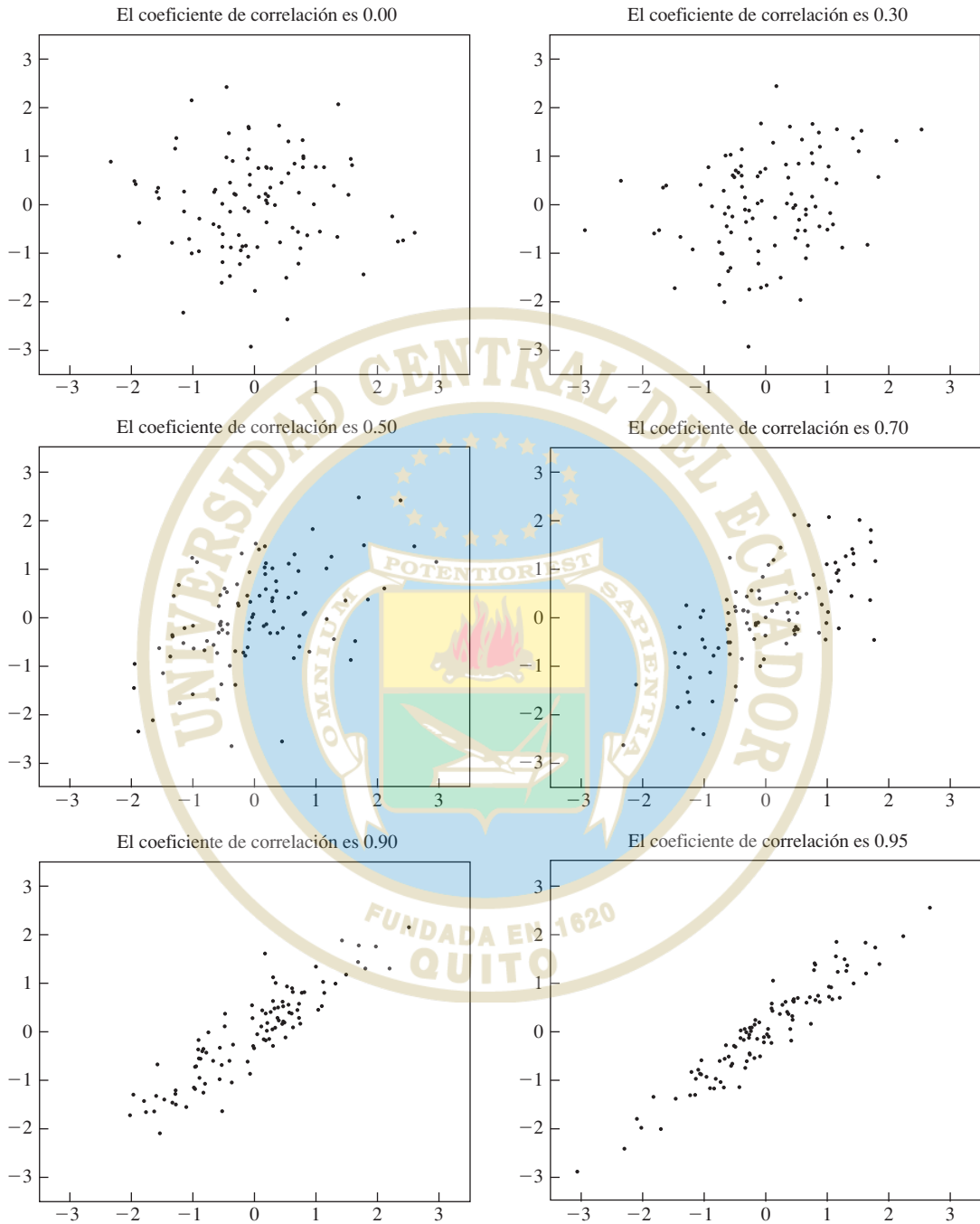


FIGURA 7.3 Ejemplos de diferentes niveles de correlación positiva.

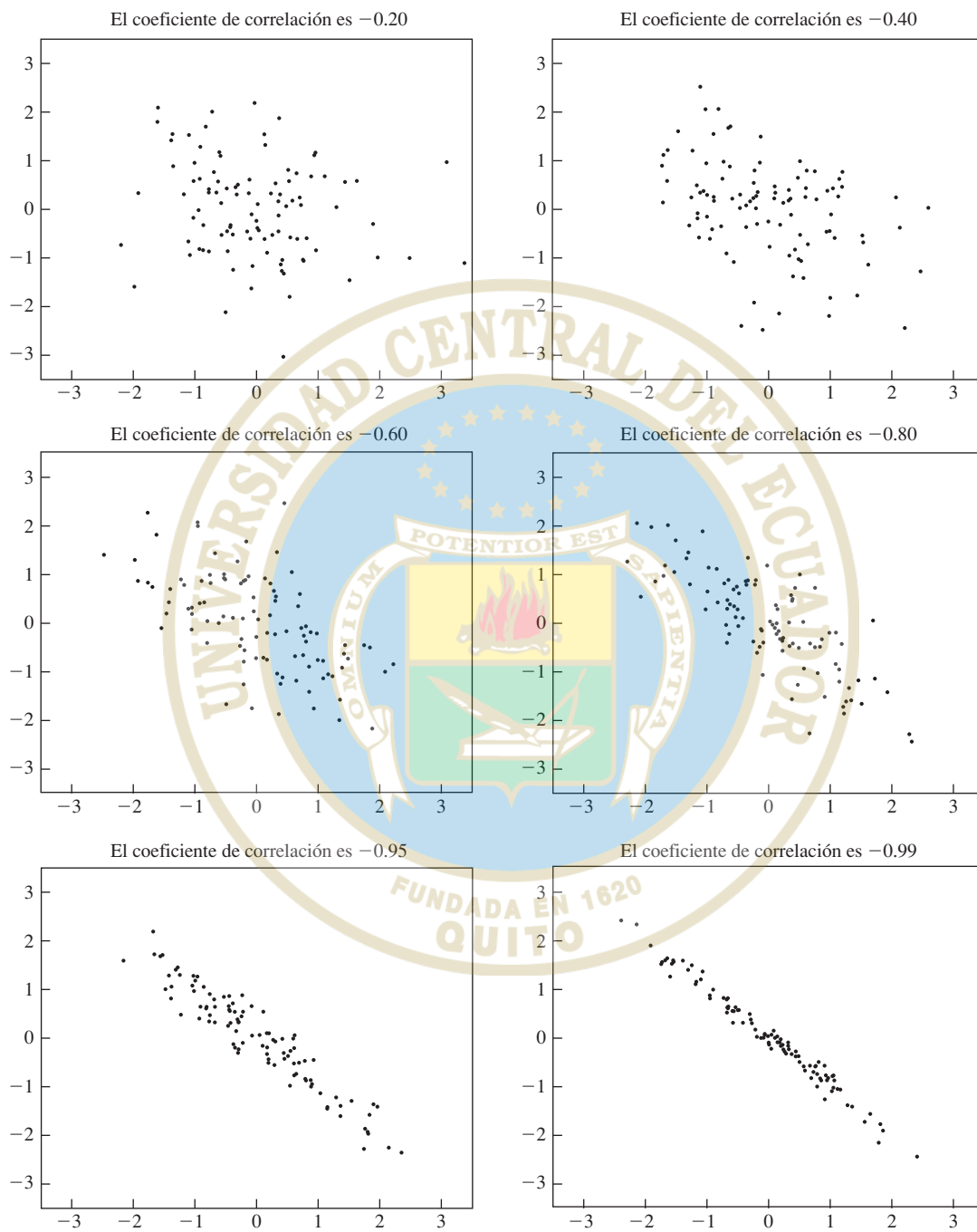


FIGURA 7.4 Ejemplos de diferentes niveles de correlación negativa.

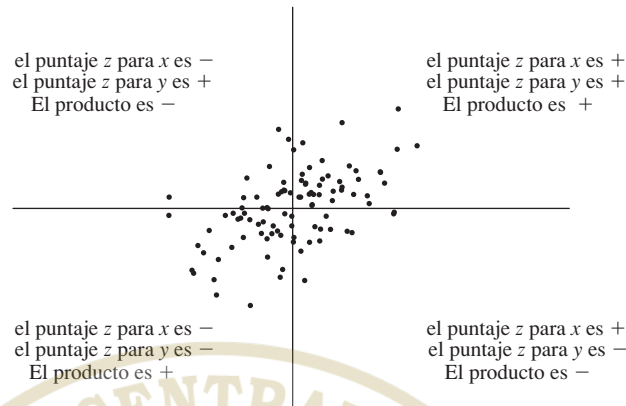


FIGURA 7.5 Cómo funciona el coeficiente de correlación.

Otra importante propiedad del coeficiente de correlación es que no lo afectan las unidades con las que se hacen las mediciones. Por ejemplo, imagine que en la figura 7.1 las estaturas de los hombres se midieron en centímetros en lugar de pulgadas. Entonces cada x_i estaría multiplicada por 2.54. Pero esto ocasionaría que $\bar{x}$ y s_x también estén multiplicadas por 2.54, por lo que los puntajes $z(x_i - \bar{x})/s_x$ permanecerían constantes, por lo que r permanece igual. En un ejemplo más extravagante, imagine que cada hombre estaba sobre una plataforma de dos pulgadas mientras se medía su estatura. Esto aumentaría cada x_i en 2, pero el de valor de $\bar{x}$ también habría aumentado en 2. Por tanto, los puntajes z no cambiarían, y el coeficiente de correlación también permanecería constante. Por último, imagine que se intercambian los valores de x y y , usando a x para denotar las longitudes de antebrazo, y que y denote las estaturas. Dado que el coeficiente de correlación se determina mediante el producto de los puntajes z , no importa cuál variable se representaba por medio de x y cuál por y .

Resumen

El coeficiente de correlación permanece constante en cada una de las siguientes operaciones:

- Multiplicar cada valor de una variable por una constante positiva.
- Sumar una constante a cada valor de una variable.
- Intercambiar los valores de x y y .

La figura 7.6 de la página 482 presenta gráficas de las medias de temperatura durante abril y octubre para varias ciudades de Estados Unidos. Ya sea que las temperaturas se midan en $^{\circ}\text{C}$ o en $^{\circ}\text{F}$, la correlación es la misma. Esto último se debe a que convertir de $^{\circ}\text{C}$ a $^{\circ}\text{F}$ implica multiplicar por 1.8 y sumar 32.

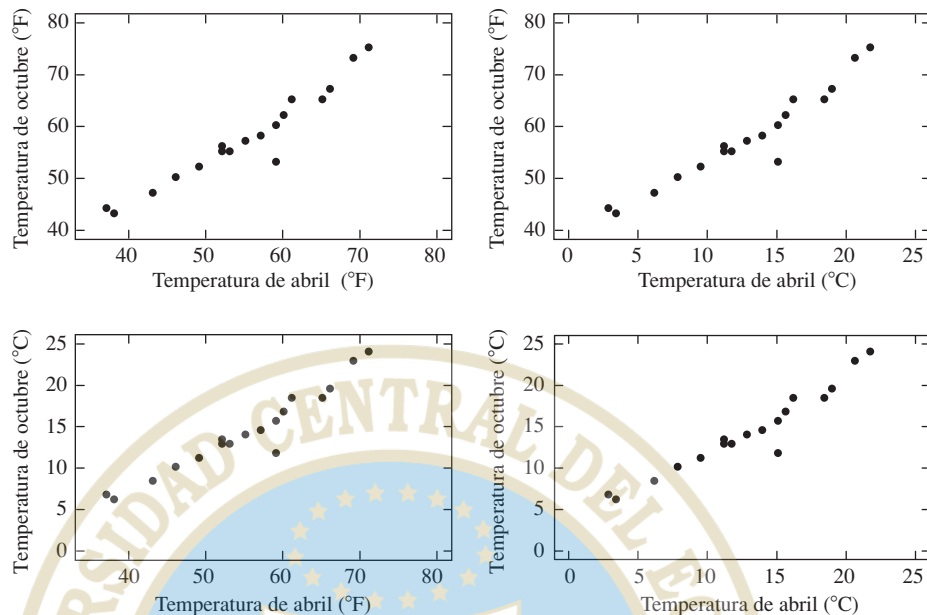


FIGURA 7.6 Medias de las temperaturas de abril y octubre para algunas ciudades de Estados Unidos. El coeficiente de correlación es 0.96 para cada figura; no importan las unidades elegidas.

El coeficiente de correlación mide sólo la asociación lineal

Un objeto se dispara hacia arriba desde el suelo con una velocidad inicial de 64 pies/s. Para cada uno de los tiempos $x_1, \dots, x_n$ se miden las alturas $y_1, \dots, y_n$ desde la superficie de la tierra. En ausencia de fricción, y suponiendo que no hay error de medición, el diagrama de dispersión de los puntos $(x_1, y_1), \dots, (x_n, y_n)$ se parecerá a la figura 7.7. Obviamente existe una fuerte relación entre x y y ; en efecto, el valor de y se determina mediante x a través de la función $y = 64x - 16x^2$. Todavía la correlación entre x y y es igual a 0. ¿Algo está equivocado? No. El valor de 0 para la correlación indica que no hay ninguna relación *lineal* entre x y y , lo que es verdadero. La relación es puramente *cuadrática*. La lección de este ejemplo es que el coeficiente de correlación se debe utilizar sólo cuando la relación entre x y y es lineal. De otra manera los resultados pueden ser engañosos.

Datos atípicos

En la figura 7.8 el punto (0, 3) es un dato atípico porque está apartado del cuerpo principal de datos. La correlación para este diagrama de dispersión es $r = 0.26$, lo que indica una relación lineal débil. Aun así, diez de los 11 puntos tienen una relación lineal perfecta. Los datos atípicos pueden distorsionar el coeficiente de correlación enormemente, en especial con conjuntos pequeños de datos, y presentan un serio problema para los analistas de datos. Algunos datos atípicos los ocasionan errores en el registro de datos o fallas en la secuencia del protocolo experimental. Estos datos atípicos se pueden corregir o eliminar adecuadamente. A veces las personas eliminan los datos atípicos de una gráfica sin justificación, para darle una

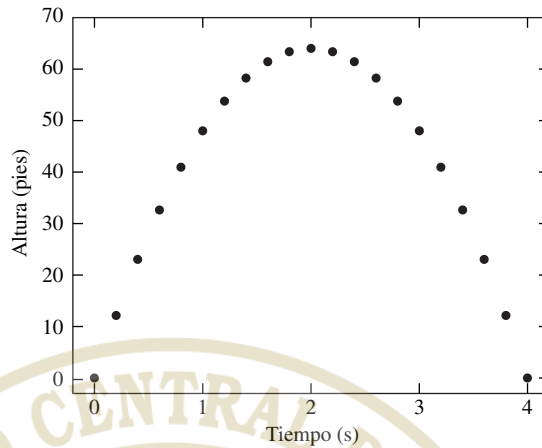


FIGURA 7.7 La relación entre la altura de un objeto en caída libre con una velocidad positiva inicial y el tiempo de caída libre es cuadrático. La correlación es igual a 0.

apariencia más agradable. Esto no es adecuado, ya que es resultado de una subestimación de la variabilidad del proceso que generó los datos. Interpretar los datos que contienen datos atípicos puede ser difícil, porque hay pocas reglas fáciles de seguir.

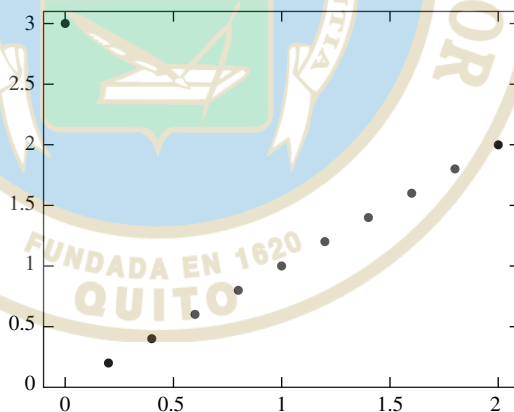


FIGURA 7.8 La correlación es 0.26. Debido a los datos atípicos el coeficiente de correlación es engañoso.

La correlación no es causalidad

En los niños, lo extenso de su vocabulario está fuertemente correlacionado con el tamaño de su zapato. Sin embargo, aprender nuevas palabras no ocasiona que sus pies crezcan, como tampoco pies en crecimiento significa que su vocabulario aumente. Hay un tercer factor, en concreto la edad, que está correlacionada tanto con el tamaño del zapato como con el voca-

bulario. Niños de mayor edad tienden a tener tanto mayor tamaño de zapato como vocabularios más extensos, y esto último propicia una correlación positiva entre el vocabulario y el tamaño del zapato. Este fenómeno se conoce como **confusión**. La confusión ocurre cuando hay una tercera variable correlacionada con las dos variables de interés, lo que da como resultado una correlación entre ambas.

Repitiendo este ejemplo con más detalle: las personas con mayor edad tienden a tener un tamaño de zapato más grande. Las personas de mayor edad también tienden a tener vocabularios más extensos. Se tiene que personas con tamaño de zapato más grande tenderán a tener vocabularios más variados. En otras palabras, dado que tanto el tamaño de zapato como el vocabulario están correlacionados con la edad, su correlación es positiva.

En este ejemplo la confusión se descubrió con facilidad. En muchos casos no es tan fácil. El ejemplo muestra que el solo hecho de que dos variables estén correlacionadas entre sí no permite suponer que el cambio en una tenderá a ocasionar un cambio en la otra. Antes de que se concluya que las dos variables tienen una relación causal se debe descartar la posibilidad de confusión.

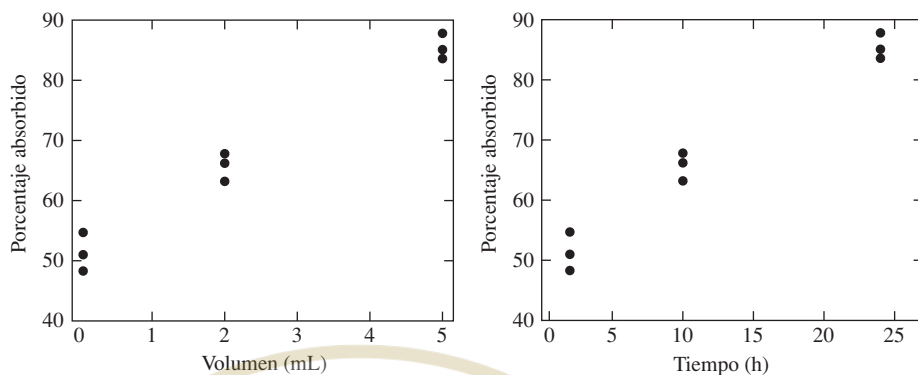
A veces se puede utilizar la **regresión múltiple** (véase el capítulo 8) para detectar confusión. Algunas veces los experimentos se pueden diseñar para reducir la posibilidad de confusión. El tema del **diseño experimental** (véase el capítulo 9) está muy relacionado con este tema. He aquí un ejemplo simple.

Ejemplo
7.1

Una científica ecologista estudia la **tasa de absorción** de cierto compuesto químico sobre la piel. Coloca volúmenes diferentes del **compuesto químico** sobre diferente segmentos de piel y permite que **ésta permanezca en contacto con el compuesto químico** durante diferentes lapsos. Ella entonces mide **el volumen del compuesto químico absorbido** en cada segmento de piel. Obtiene los resultados que se muestran en la tabla siguiente.

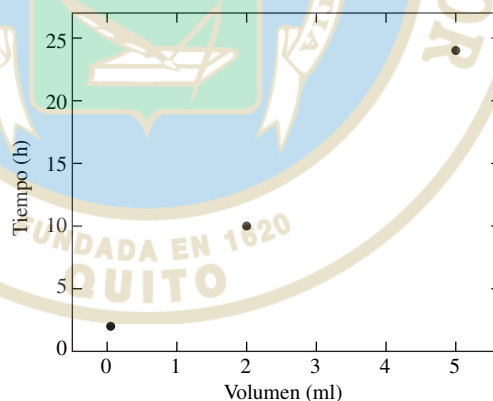
Volumen (mL)	Tiempo (h)	Porcentaje absorbido
0.05	2	48.3
0.05	2	51.0
0.05	2	54.7
2.00	10	63.2
2.00	10	67.8
2.00	10	66.2
5.00	24	83.6
5.00	24	85.1
5.00	24	87.8

La científica grafica el porcentaje absorbido contra el volumen y el tiempo, como se muestra en la siguiente figura. Calcula la correlación entre el volumen y la absorción y obtiene $r = 0.988$. Concluye que aumentar el volumen del compuesto químico ocasiona el incremento de los porcentajes de absorción. Después calcula la correlación entre el tiempo y la absorción, por lo que obtiene $r = 0.987$. Concluye que al aumentar el tiempo que la piel está en contacto con el compuesto químico ocasiona también que los porcentajes de absorción aumenten. ¿Estas conclusiones están justificadas?



Solución

No. La científica debe observar la gráfica tiempo contra volumen que se presenta en la siguiente figura. La correlación entre el tiempo y el volumen es $r = 0.999$, por lo que estas dos variables están casi completamente confundidas. Si el tiempo o el volumen afecta los porcentajes de absorción, pareciese que *ambos* lo hacen porque están muy correlacionados entre sí. En consecuencia, es imposible determinar si es el tiempo o el volumen el que está teniendo un efecto. Esta relación entre tiempo y volumen resulta del diseño experimental y se debió haber evitado.



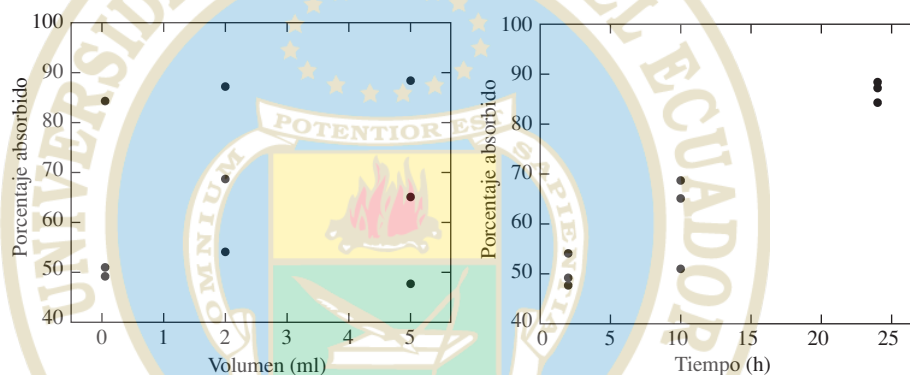
Ejemplo

7.2

La científica del ejemplo 7.1 ha repetido el experimento, esta vez con un nuevo diseño. Los resultados se presentan en la tabla siguiente.

Volumen (mL)	Tiempo (h)	Porcentaje absorbido
0.05	2	49.2
0.05	10	51.0
0.05	24	84.3
2.00	2	54.1
2.00	10	68.7
2.00	24	87.2
5.00	2	47.7
5.00	10	65.1
5.00	24	88.4

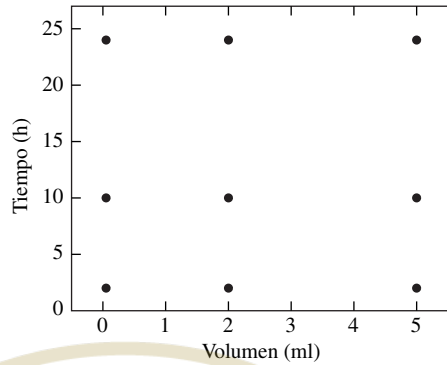
La científica grafica el porcentaje absorbido tanto contra el volumen como contra el tiempo, como se muestra en la siguiente figura.



Después calcula la correlación entre el volumen y la absorción y obtiene $r = 0.121$. Concluye que aumentar el volumen del compuesto químico tiene poco o ningún efecto en los porcentajes de absorción. Después calcula la correlación entre el tiempo y la absorción y obtiene $r = 0.952$. Determina que al aumentar el tiempo que el compuesto químico está en contacto con la piel causará que los porcentajes de absorción aumenten. ¿Estas conclusiones están justificadas?

Solución

Estas conclusiones están mucho mejor justificadas que las del ejemplo 7.1. Se verá por qué. Observe la gráfica de tiempo contra volumen de la siguiente figura. Este experimento ha sido diseñado con el propósito de que el tiempo y el volumen no estén correlacionados. Ahora parece que el tiempo, pero no el volumen, tiene efecto en los porcentajes de absorción. Antes de establecer la conclusión final de que el aumento del tiempo realmente ocasiona que el porcentaje absorbido aumente, la científica debe asegurarse de que no haya algunas otras confusiones potenciales por ahí. Por ejemplo, si la temperatura ambiental varió con cada reproducción del experimento, y si estaba muy correlacionada con el tiempo, entonces podría darse el caso de que la temperatura, más que el tiempo, estaba ocasionando la variación de los porcentajes de absorción.



Inferencia sobre la correlación poblacional

En lo que resta de esta sección se utilizarán algunas ideas de la sección 2.6. Cuando los puntos (x_i, y_i) son una muestra aleatoria de una población de pares ordenados, entonces cada punto se puede considerar como observación de un par ordenado de variables aleatorias (X, Y) . El coeficiente de correlación, o correlación muestral, r es entonces una estimación de la correlación poblacional, $\rho_{X,Y}$.

Si las variables aleatorias X y Y tienen cierta distribución conjunta llamada **distribución normal bivariada**, entonces la correlación muestral r se puede utilizar para construir intervalos de confianza y realizar pruebas de hipótesis sobre la correlación poblacional. En la práctica, si X y Y están distribuidas normalmente, entonces hay una certeza virtual de que tanto X como Y tengan una distribución normal bivariada, por lo que los intervalos de confianza y las pruebas descritas después serán válidas. (Mientras que es matemáticamente posible construir dos variables aleatorias normales que no una normal bivariada de manera conjunta, las condiciones en las que esto último ocurre no son visibles en la práctica.)

Los intervalos de confianza, y la mayoría de las pruebas, sobre $\rho_{X,Y}$ están basados en el resultado siguiente:

Sean X y Y variables aleatorias con distribución normal bivariada.

Sea ρ la correlación de población entre X y Y .

Sea $(x_1, y_1), \dots, (x_n, y_n)$ una muestra aleatoria de la distribución conjunta de X y Y .

Sea r la correlación muestral de los n puntos.

Entonces la cantidad

$$W = \frac{1}{2} \ln \frac{1+r}{1-r} \quad (7.3)$$

está casi normalmente distribuida, con la media dada por

$$\mu_W = \frac{1}{2} \ln \frac{1+\rho}{1-\rho} \quad (7.4)$$

y varianza dada por

$$\sigma_W^2 = \frac{1}{n-3} \quad (7.5)$$

Observe que μ_w es una función de la correlación poblacional ρ . La construcción de intervalos de confianza requerirá despejar ρ de la ecuación (7.4). Se obtiene

$$\rho = \frac{e^{2\mu_w} - 1}{e^{2\mu_w} + 1} \quad (7.6)$$

Ejemplo

7.3

En un estudio de los tiempos de reacción, el tiempo de respuesta a un estímulo visual (x) y el tiempo de respuesta a un estímulo auditivo (y) se registraron para cada una de diez personas. Los tiempos se midieron en m. En la tabla siguiente se presentan los resultados.

x	161	203	235	176	201	188	228	211	191	178
y	159	206	241	163	197	193	209	189	169	201

Determine un intervalo de confianza de 95% para la correlación entre los dos tiempos de reacción.

Solución

Usando la ecuación (7.1) se calcula la correlación muestral, se obtiene $r = 0.8159$. Después se utiliza la ecuación (7.3) para calcular la cantidad W :

$$\begin{aligned} W &= \frac{1}{2} \ln \frac{1+r}{1-r} \\ &= \frac{1}{2} \ln \frac{1+0.8159}{1-0.8159} \\ &= 1.1444 \end{aligned}$$

Dado que W está distribuida normalmente con desviación estándar $\sigma = \sqrt{1/(10-3)} = 0.3780$ (ecuación 7.5), un intervalo de confianza de 95% para μ_w está dado por

$$\begin{aligned} 1.1444 - 1.96(0.3780) &< \mu_w < 1.1444 + 1.96(0.3780) \\ 0.4036 &< \mu_w < 1.8852 \end{aligned}$$

Para obtener un intervalo de confianza de 95% para ρ se transforma la desigualdad mediante la ecuación (7.6), por lo que se obtiene

$$\begin{aligned} \frac{e^{2(0.4036)} - 1}{e^{2(0.4036)} + 1} &< \frac{e^{2\mu_w} - 1}{e^{2\mu_w} + 1} < \frac{e^{2(1.8852)} - 1}{e^{2(1.8852)} + 1} \\ 0.383 &< \rho < 0.955 \end{aligned}$$

Para probar hipótesis nulas de la forma $\rho = \rho_0$, $\rho \leq \rho_0$ y $\rho \geq \rho_0$, donde ρ_0 es una constante que *no es igual a 0*, la cantidad W es la base de una prueba. A continuación se presenta un ejemplo.

Ejemplo**7.4**

Con referencia al ejemplo 7.3, determine el P -valor para probar $H_0 : \rho \leq 0.3$ contra $H_1 : \rho > 0.3$.

Solución

Bajo H_0 se tiene que $\rho = 0.3$, por lo que, al usar la ecuación (7.4),

$$\begin{aligned}\mu_W &= \frac{1}{2} \ln \frac{1 + 0.3}{1 - 0.3} \\ &= 0.3095\end{aligned}$$

La desviación estándar de W es $\sigma = \sqrt{1/(10-3)} = 0.3780$. Por lo que bajo H_0 , $W \sim N(0.3095, 0.3780^2)$. El valor observado de W es $W = 1.1444$. El puntaje z es, por tanto,

$$z = \frac{1.1444 - 0.3095}{0.3780} = 2.21$$

El P -valor es 0.0136. Se concluye que $\rho > 0.3$.

Para la prueba de hipótesis nula de la forma $\rho = 0$, $\rho \leq 0$, $\rho \geq 0$ está disponible un procedimiento parecido. Cuando $\rho = 0$, la cantidad

$$U = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

tiene la distribución t de Student con $n - 2$ grados de libertad. El ejemplo 7.5 muestra cómo utilizar U como un estadístico de prueba.

Ejemplo**7.5**

Con referencia al ejemplo 7.3, pruebe la hipótesis $H_0 : \rho \leq 0$ contra $H_1 : \rho > 0$.

Solución

Bajo H_0 se tiene que $\rho = 0$, por lo que el estadístico de prueba U tiene la distribución t de Student con $n - 2 = 8$ grados de libertad. La correlación muestral es $r = 0.8159$, por lo que el valor de U es

$$\begin{aligned}U &= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \\ &= \frac{0.8159\sqrt{10-2}}{\sqrt{1-0.8159^2}} \\ &= 3.991\end{aligned}$$

Al consultar la tabla t de Student con ocho grados de libertad, se encuentra que el P -valor está entre 0.001 y 0.005. Es seguro concluir que $\rho > 0$.

Ejercicios de la sección 7.1

1. Calcule el coeficiente de correlación para el siguiente conjunto de datos.

x	1	2	3	4	5	6	7
y	2	1	4	3	7	5	6

2. Para cada uno de los siguientes conjuntos de datos, explique por qué el coeficiente de correlación es el mismo que para el conjunto de datos del ejercicio 1.

a)

x	1	2	3	4	5	6	7
y	5	4	7	6	10	8	9

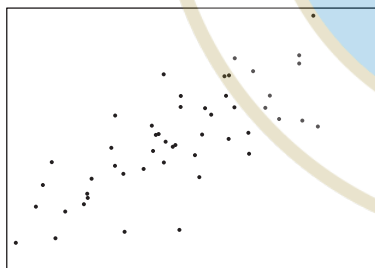
b)

x	11	21	31	41	51	61	71
y	5	4	7	6	10	8	9

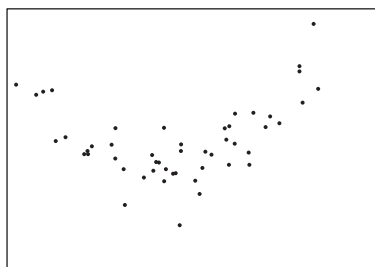
c)

x	53	43	73	63	103	83	93
y	4	6	8	10	12	14	16

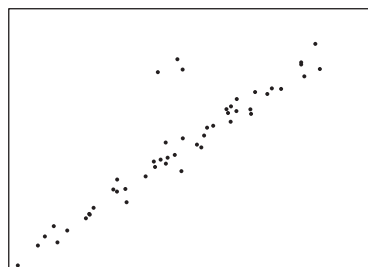
3. Establezca para cada una de las siguientes gráficas de dispersión si el coeficiente de correlación es un resumen adecuado y explíquelo brevemente.



a)



b)



c)

4. Verdadero o falso, y explique brevemente:

a) Si el coeficiente de correlación es positivo, entonces los valores arriba del promedio de una variable están asociados con los valores superiores al promedio de la otra.

b) Si el coeficiente de correlación es negativo, entonces los valores debajo del promedio de una variable están asociados con los valores debajo del promedio de la otra.

c) Si y es usualmente menor que x , entonces la correlación entre y y x será negativa.

5. Un investigador reunió datos acerca de las estaturas y los pesos de estudiantes universitarios. La correlación entre la estatura y el peso para hombres era de 0.6, y para mujeres era igual. Si los hombres y las mujeres se consideran en conjunto, la correlación entre la estatura y el peso ¿es mayor que 0.6, menor que 0.6, o casi igual a 0.6? Puede ser útil hacer un diagrama aproximado de dispersión.

6. En un estudio de movimientos de tierra ocasionados por sismos se registraron para cinco de éstos la velocidad máxima (en m/s) y la aceleración máxima (en m/s^2). Los resultados se presentan en la tabla siguiente.

Velocidad	1.54	1.60	0.95	1.30	2.92
Aceleración	7.64	8.04	8.04	6.37	5.00

a) Calcule el coeficiente de correlación entre la velocidad máxima y la aceleración máxima.

b) Construya un diagrama de dispersión para estos datos.

c) ¿El coeficiente de correlación es un resumen adecuado de estos datos? Explique por qué.

d) Alguien sugiere convertir las unidades de metros a centímetros y de segundos a minutos. ¿Qué efecto tendría sobre la correlación hacer esto?

7. Un ingeniero químico está estudiando el efecto de la temperatura y la tasa de agitación en la producción de cierto producto. El proceso se realiza 16 veces; en la tabla siguiente se muestran los resultados. Las unidades para la producción son porcentajes de un máximo teórico.

Temperatura (°C)	Tasa de agitación (rpm)	Producción (%)
110	30	70.27
110	32	72.29
111	34	72.57
111	36	74.69
112	38	76.09
112	40	73.14
114	42	75.61
114	44	69.56
117	46	74.41
117	48	73.49
122	50	79.18
122	52	75.44
130	54	81.71
130	56	83.03
143	58	76.98
143	60	80.99

- a) Calcule la correlación entre la temperatura y la producción, entre la tasa de agitación y la producción, y entre la temperatura y la tasa de agitación.
- b) ¿Estos datos proporcionan, dentro del rango de los datos, buena evidencia de que el aumento de la temperatura ocasiona que la producción aumente? ¿O el resultado se podría deber a la confusión? Explique.
- c) ¿Estos datos proporcionan, dentro del rango de los datos, buena evidencia de que el aumento de la tasa de agitación ocasiona que la producción aumente? ¿O el resultado se podría deber a la confusión? Explique.

8. Otro ingeniero químico está estudiando el mismo procedimiento que en el ejercicio 7 y utiliza la siguiente matriz experimental.

Temperatura (°C)	Tasa de agitación (rpm)	Producción (%)
110	30	70.27
110	40	74.95
110	50	77.91
110	60	82.69
121	30	73.43
121	40	73.14
121	50	78.27
121	60	74.89
132	30	69.07
132	40	70.83
132	50	79.18
132	60	78.10
143	30	73.71
143	40	77.70
143	50	74.31
143	60	80.99

- a) Calcule la correlación entre la temperatura y la producción, entre la tasa de agitación y la producción, y entre la temperatura y la tasa de agitación.
- b) ¿Estos datos proporcionan, dentro del rango de los datos, buena evidencia de que la producción no está afectada por la temperatura? ¿O el resultado se podría deber a la confusión? Explique.
- c) ¿Estos datos proporcionan, dentro del rango de los datos, buena evidencia de que el aumento de la tasa de agitación ocasiona que la producción aumente? ¿O el resultado se podría deber a la confusión? Explique.
- d) ¿Cuál de los diseños experimentales es mejor, éste o el del ejercicio 7? Explique.

9. Las siguientes son mediciones de la fuerza de tensión en ksi (x) y la dureza de Brinell (y) para diez elementos del cobre extraído en frío. Suponga que la fuerza de tensión y la dureza de Brinell siguen una distribución normal bivariada.

x	y
106.2	35.0
106.3	37.2
105.3	39.8
106.1	35.8
105.4	41.3
106.3	40.7
104.7	38.7
105.4	40.2
105.5	38.1
105.1	41.6

- a) Determine un intervalo de confianza de 95% para ρ , la correlación poblacional entre la fuerza de tensión y la dureza de Brinell.
- b) ¿Puede concluir que $\rho < 0.3$?
- c) ¿Puede concluir que $\rho \neq 0$?

10. En una muestra de 400 cojinetes de bolas, el coeficiente de correlación entre la excentricidad y la suavidad fue de $r = 0.10$.

- a) Determine el P -valor para probar que $H_0 : \rho \leq 0$ contra $H_1 : \rho > 0$. ¿Puede concluir que $\rho > 0$?
- b) ¿El resultado del inciso a) le permite concluir que hay una fuerte correlación entre excentricidad y suavidad? Explique.

11. Un diagrama de dispersión contiene cuatro puntos: $(-2, -2)$, $(-1, -1)$, $(0, 0)$ y $(1, 1)$. Se agrega un quinto punto, $(2, y)$, a la gráfica. Sea r la correlación entre x y y .

- a) Determine el valor de y para que $r = 1$.
 b) Determine el valor de y para que $r = 0$.
 c) Determine el valor de y para que $r = 0.5$.
 d) Determine el valor de y para que $r = -0.5$.
 e) Dé un argumento geométrico para mostrar que no existe un valor y para el cual $r = -1$.

7.2 La recta de mínimos cuadrados

Cuando dos variables tienen una relación lineal, el diagrama de dispersión tiende a estar agrupado alrededor de la recta de mínimos cuadrados (véanse las figuras 7.1 y 7.2 en la sección 7.1). En esta sección se aprenderá cómo calcular la recta de mínimos cuadrados y cómo se puede utilizar para obtener conclusiones de los datos.

Se empieza describiendo un experimento hipotético. Los resortes se usan en aplicaciones por su capacidad para alargarse (contraerse) bajo carga. La rigidez de un resorte se mide con la “constante del resorte”, que es la longitud del resorte que se alargará por unidad de la fuerza o de carga.¹ Para asegurarse de que un resorte dado funciona adecuadamente es necesario calcular la constante de resorte con exactitud y precisión.

En este experimento hipotético un resorte se cuelga verticalmente con un extremo fijo, y los pesos se cuelgan uno tras otro del otro extremo. Después de colgar cada peso se mide la longitud del resorte. Sean $x_1, \dots, x_n$ los pesos, y sea l_i la longitud del resorte bajo la carga x_i . La ley de Hooke establece que

$$l_i = \beta_0 + \beta_1 x_i \quad (7.7)$$

donde β_0 representa la longitud del resorte cuando no tiene carga y β_1 es la constante del resorte.

Sea y_i la longitud *medida* del resorte bajo carga x_i . Debido al error de medición, y_i será diferente de la longitud verdadera l_i . Se escribe como

$$y_i = l_i + \varepsilon_i \quad (7.8)$$

donde ε_i constituye el error en la i -ésima medición. Al combinar las ecuaciones (7.7) y (7.8) se obtiene

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (7.9)$$

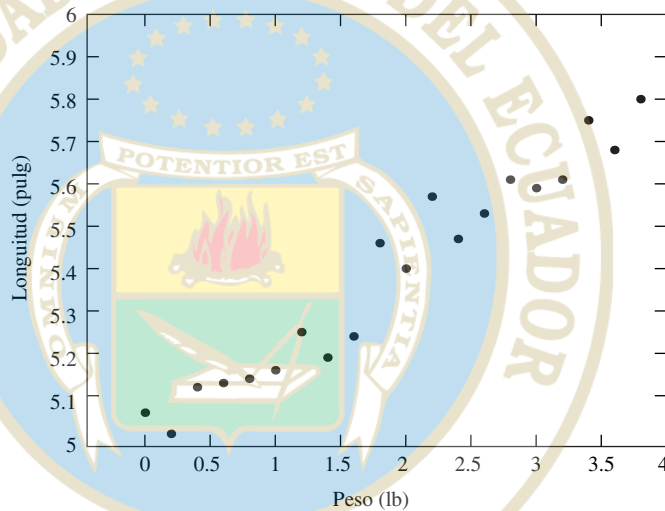
En la ecuación (7.9) y_i se llama **variable dependiente**, x_i se conoce como **variable independiente**, β_0 y β_1 son los **coeficientes de regresión**, y ε_i se denomina **error**. A la ecuación (7.9) se le llama **modelo lineal**.

La tabla 7.1 presenta los resultados del experimento hipotético y la figura 7.9 el diagrama de dispersión de y contra x . Se quieren utilizar estos datos para calcular la constante β_1 del resorte y la longitud β_0 sin carga. Si no hubiera error en la medición, los puntos se encontrarían en una línea recta con pendiente β_1 e intercepto β_0 , y estas cantidades serían fáciles de determinar. Debido al error de medición, β_0 y β_1 no se pueden determinar exactamente, pero se pueden estimar cuando se calcula la recta de mínimos cuadrados.

¹La definición más tradicional de la constante del resorte es el recíproco de esta cantidad, a saber: la fuerza que se requiere para extender el resorte una unidad de longitud.

TABLA 7.1 Longitudes medidas de un resorte bajo diferentes cargas

Peso (lb)	Longitud medida (pulg)	Peso (lb)	Longitud medida (pulg)
x	y	x	y
0.0	5.06	2.0	5.40
0.2	5.01	2.2	5.57
0.4	5.12	2.4	5.47
0.6	5.13	2.6	5.53
0.8	5.14	2.8	5.61
1.0	5.16	3.0	5.59
1.2	5.25	3.2	5.61
1.4	5.19	3.4	5.75
1.6	5.24	3.6	5.68
1.8	5.46	3.8	5.80

**FIGURA 7.9** Gráfica de longitudes medidas de un resorte contra carga.

La figura 7.10 de la página 494 presenta el diagrama de dispersión de y contra x con la recta de mínimos cuadrados sobrepuesta. Se escribe la ecuación de la recta como

$$y = \hat{\beta}_0 + \hat{\beta}_1 x \quad (7.10)$$

Las cantidades $\hat{\beta}_0$ y $\hat{\beta}_1$ se llaman **coeficientes de mínimos cuadrados**. El coeficiente $\hat{\beta}_1$ representa la pendiente de la recta de mínimos cuadrados, es una estimación de la constante verdadera del resorte β_1 , y el coeficiente $\hat{\beta}_0$, el intercepto de la recta de mínimos cuadrados, significa una estimación de la verdadera longitud sin carga β_0 .

La recta de mínimos cuadrados es la que ajusta “mejor” los datos. Ahora se define lo que significa “mejor”. Para cada punto de datos (x_i, y_i) la distancia vertical al punto $(x_i, \hat{y}_i)$ en la recta de mínimos cuadrados es $e_i = y_i - \hat{y}_i$ (véase la figura 7.10). La cantidad $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ se llama **valor ajustado**, y la cantidad e_i se llama **residuo** asociado con el punto (x_i, y_i) . El residuo e_i representa la diferencia entre el valor observado y_i en los datos y el valor ajusta-

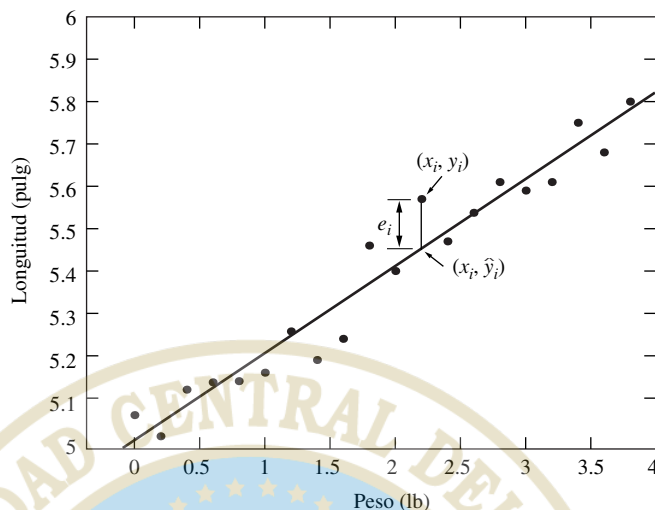


FIGURA 7.10 Gráfica de longitudes medidas de un resorte contra carga. La recta de mínimos cuadrados $y = \hat{\beta}_0 + \hat{\beta}_1 x$ está sobrepuesta. La distancia vertical de un punto de datos (x_i, y_i) para el punto $(x_i, \hat{y}_i)$ en la recta es el i -ésimo residuo e_i . La recta de mínimos cuadrados es la recta que minimiza la suma de los cuadrados de los residuos.

do $\hat{y}_i$ pronosticado por la recta de mínimos cuadrados. Éste es la distancia vertical del punto a la recta. Los puntos arriba de la recta de mínimos cuadrados tienen residuos positivos, y los puntos debajo de la recta de mínimos cuadrados tienen residuos negativos. Entre más cercanos estén los valores ajustados a las observaciones, mejor ajustará la recta a los datos. La recta de mínimos cuadrados se define como aquella para la que la suma de cuadrados de los residuos $\sum_{i=1}^n e_i^2$ se minimiza. En este sentido, la recta de mínimos cuadrados ajusta mejor los datos que cualquier otra recta.

En el ejemplo de ley de Hooke hay solamente una variable independiente (peso), ya que es razonable suponer que la única variable que afecta la longitud del resorte es la cantidad del peso colgada de él. En otros casos se necesita utilizar algunas variables independientes. Por ejemplo, para pronosticar la cosecha de cierto cultivo se necesitará saber la cantidad de fertilizante usado, la cantidad de agua aplicada y diferentes mediciones de las propiedades químicas del suelo. Modelos lineales como la ley de Hooke, con sólo una variable independiente, se conocen como modelos de **regresión lineal simple**. Estos últimos, con más de una variable independiente, se llaman modelos de **regresión múltiple**. Este capítulo trata acerca de la regresión lineal simple. La regresión múltiple se trata en el capítulo 8.

Cálculo de la ecuación de la recta de mínimos cuadrados

Para calcular la ecuación de la recta de mínimos cuadrados se deben determinar los valores para la pendiente $\hat{\beta}_1$ y el intercepto $\hat{\beta}_0$ que minimizan la suma del residuo $\sum_{i=1}^n e_i^2$.

Para hacer esto primero se expresa e_i en función de $\hat{\beta}_0$ y $\hat{\beta}_1$:

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \quad (7.11)$$

Por tanto, $\hat{\beta}_0$ y $\hat{\beta}_1$ son las cantidades que minimizan la suma

$$S = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \quad (7.12)$$

Estas cantidades son

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (7.13)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (7.14)$$

Las deducciones de estos resultados se proporcionan al final de esta sección.

Fórmulas para el cálculo

Es necesario calcular las cantidades $\sum_{i=1}^n (x_i - \bar{x})^2$ y $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ con el fin de determinar la ecuación de la recta de mínimos cuadrados y, como pronto se verá, la cantidad $\sum_{i=1}^n (y_i - \bar{y})^2$ que se necesita calcular con el fin de determinar qué tan bien ajusta la recta a los datos. Cuando se calculan estas cantidades a mano existen fórmulas alternas que con frecuencia son más fáciles de utilizar. Éstas se proporcionan en el cuadro siguiente.

Fórmulas para el cálculo

Las expresiones de la derecha son equivalentes a las de la izquierda, y son, con frecuencia, más fáciles de calcular:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 \quad (7.15)$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2 \quad (7.16)$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \quad (7.17)$$

Ejemplo

7.6

Con los datos de la ley de Hooke de la tabla 7.1 calcule los estimadores de mínimos cuadrados de la constante del resorte y la longitud sin carga del resorte. Escriba la ecuación de la recta de mínimos cuadrados.

Solución

La estimación de la constante del resorte es $\hat{\beta}_1$ y la estimación de la longitud sin carga es $\hat{\beta}_0$. De la tabla 7.1 se calcula:

$$\bar{x} = 1.9000 \quad \bar{y} = 5.3885$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = 26.6000$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} = 5.4430$$

Mediante las ecuaciones (7.13) y (7.14) se calcula

$$\hat{\beta}_1 = \frac{5.4430}{26.6000} = 0.2046$$

$$\hat{\beta}_0 = 5.3885 - (0.2046)(1.9000) = 4.9997$$

La ecuación de la recta de mínimos cuadrados es $y = \hat{\beta}_0 + \hat{\beta}_1 x$. Al sustituir los valores calculados para $\hat{\beta}_0$ y $\hat{\beta}_1$, se obtiene

$$y = 4.9997 + 0.2046 x$$

Con la ecuación de la recta de mínimos cuadrados se pueden calcular los valores ajustados $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ y los residuos $e_i = y_i - \hat{y}_i$ para cada punto (x_i, y_i) en el conjunto de datos de la ley de Hooke. Los resultados se presentan en la tabla 7.2. El punto cuyo residuo se muestra en la figura 7.10 es $x = 2.2$.

En el ejemplo de la ley de Hooke la cantidad $\beta_0 + \beta_1 x$ representa la longitud verdadera del resorte bajo una carga x . Puesto que $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores de los valores verdaderos β_0 y β_1 la cantidad $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ es una estimación de $\beta_0 + \beta_1 x$. Los ejemplos 7.7 y 7.8 muestran esto último.

Ejemplo

7.7

Con los datos de la ley de Hooke calcule la longitud del resorte bajo una carga de 1.3 lb.

Solución

En el ejemplo 7.6 la ecuación de la recta de mínimos cuadrados se calculó de $y = 4.9997 + 0.2046x$. Con el valor $x = 1.3$ se estima que la longitud del resorte bajo una carga de 1.3 lb es de

$$\hat{y} = 4.9997 + (0.2046)(1.3) = 5.27 \text{ pulg.}$$

TABLA 7.2 Longitudes medidas de un resorte bajo diferentes cargas, con valores ajustados y residuos

Peso x	Longitud medida y	Valor ajustado $\hat{y}$	Residuo e	Peso x	Longitud medida y	Valor ajustado $\hat{y}$	Residuo e
0.0	5.06	5.00	0.06	2.0	5.40	5.41	-0.01
0.2	5.01	5.04	-0.03	2.2	5.57	5.45	0.12
0.4	5.12	5.08	0.04	2.4	5.47	5.49	-0.02
0.6	5.13	5.12	0.01	2.6	5.53	5.53	-0.00
0.8	5.14	5.16	-0.02	2.8	5.61	5.57	0.04
1.0	5.16	5.20	-0.04	3.0	5.59	5.61	-0.02
1.2	5.25	5.25	0.00	3.2	5.61	5.65	-0.04
1.4	5.19	5.29	-0.10	3.4	5.75	5.70	0.05
1.6	5.24	5.33	-0.09	3.6	5.68	5.74	-0.06
1.8	5.46	5.37	0.09	3.8	5.80	5.78	0.02

Ejemplo

7.8

Utilizando la ley de Hooke con los datos, estime la longitud de un resorte bajo una carga de 1.4 lb.

Solución

La estimación es $\hat{y} = 4.9997 + (0.2046)(1.4) = 5.29$ pulg.

En el ejemplo 7.8, observe que la longitud medida en una carga de 1.4 era de 5.19 pulg (véase la tabla 7.2). Pero la estimación de mínimos cuadrados de 5.29 pulg está basada en todos los datos y es más precisa (tiene menor incertidumbre). En la sección 7.3 se aprenderá cómo calcular incertidumbres en los estimadores $\hat{y}$.

Los estimadores no son lo mismo que los valores verdaderos

Es importante entender la diferencia entre los *estimadores* de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$, y los *valores verdaderos* β_0 y β_1 . Los valores verdaderos son constantes cuyos valores no se conocen. Los estimadores son cantidades que se calculan a partir de los datos. Se pueden utilizar los estimadores como aproximaciones de los valores verdaderos.

En principio, un experimento como el de la ley de Hooke se podría repetir muchas veces. Los valores verdaderos β_0 y β_1 permanecerían constantes en todas las reproducciones del experimento, pero cada reproducción produciría datos diferentes; por consiguiente, valores diferentes de los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$. Los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ son, por tanto, *variables aleatorias*, ya que sus valores varían de experimento a experimento. Para usar completamente estos estimadores será preciso calcular sus desviaciones estándar. Este tema se analizará en la sección 7.3.

Los residuos no son lo mismo que los errores

Un conjunto de puntos $(x_1, y_1), \dots, (x_n, y_n)$ sigue un modelo lineal si las coordenadas x y y están relacionadas por medio de la ecuación $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. Es importante entender la diferencia entre los residuos e_i y los errores ε_i . Cada residuo e_i es la diferencia $y_i - \hat{y}_i$ entre el

valor observado, o valor medido, el valor y_i y el valor ajustado $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ calculado por la recta de mínimos cuadrados. Debido a que los valores y_i se conocen y los valores $\hat{y}_i$ se pueden calcular a partir de los datos, es posible calcular los residuos. En contraste, los errores ε_i representan las diferencias entre y_i y los valores $\beta_0 + \beta_1 x_i$. Puesto que los valores verdaderos β_0 y β_1 no se conocen, los errores son también desconocidos. Otra manera de considerar la diferencia consiste en que los residuos constituyen las distancias verticales de los valores observados y_i y la recta de mínimos cuadrados $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$, y los errores son las distancias del y_i a la recta verdadera $y = \beta_0 + \beta_1 x$.

Resumen

Dados los puntos $(x_1, y_1), \dots, (x_n, y_n)$:

- La recta de mínimos cuadrados es $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$.
- $$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$
- $$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$
- Las cantidades que $\hat{\beta}_0$ y $\hat{\beta}_1$ se consideran como los estimadores de una pendiente verdadera β_1 y de un intercepto verdadero β_0 .
- Para cualquier x , $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ significa una estimación de la cantidad $\beta_0 + \beta_1 x$.

No haga una extrapolación fuera del rango de los datos

¿Qué sucede si se quería estimar la longitud del resorte bajo una carga de 100 lb? La estimación de mínimos cuadrados es $4.9997 + (0.2046)(100) = 25.46$ pulg. ¿Se debe creer esto? No. Ninguno de los pesos en el conjunto de datos era tan grande. Es probable que el resorte se deformaría, por lo que la ley de Hooke no valdría. Para muchas variables las relaciones lineales valen dentro de cierto rango, pero no fuera de él. Si se extrapola una recta de mínimos cuadrados fuera del rango de los datos, no habrá garantía de que describirá la relación adecuadamente. Si se quiere saber cómo responderá el resorte a una carga de 100 lb se deben incluir pesos de 100 lb o mayores en el conjunto de datos.

Resumen

No extrapole una recta ajustada (como la recta de mínimos cuadrados) fuera del rango de los datos. La relación lineal puede no ser válida ahí.

No use la recta de mínimos cuadrados cuando los datos no sean lineales

En la sección 7.1 se aprendió que el coeficiente de correlación se debía usar solamente cuando la relación entre x y y sea lineal. Lo mismo es verdad para la recta de mínimos cuadrados. Cuando el diagrama de dispersión sigue un patrón curvado no tiene caso presentarlo con una línea recta. Para mostrar esto, la figura 7.11 presenta una gráfica de la relación entre la altu-

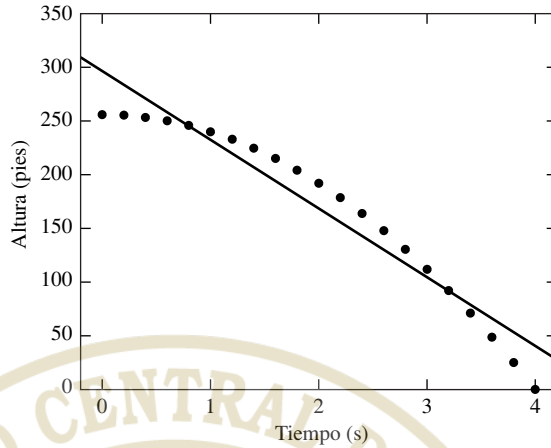


FIGURA 7.11 Relación entre la altura de un objeto en caída libre con una velocidad inicial positiva y el tiempo de caída libre no es lineal. La recta de mínimos cuadrados no se ajusta bien a los datos y no se debe utilizar para pronosticar la altura del objeto para un tiempo dado.

ra y de un objeto desde su lanzamiento a una altura de 256 pies y el tiempo x desde que se lanzó. La relación entre x y y no es lineal. La recta de mínimos cuadrados no se ajusta bien a los datos.

En algunos casos la recta de mínimos cuadrados se puede utilizar para datos no lineales, después de que se ha aplicado un proceso conocido como transformación variable. En la sección 7.4 se analiza este tema.

Otra mirada a la recta de mínimos cuadrados

La expresión (7.13) para $\hat{\beta}_1$ puede describirse en una forma que proporciona una interpretación útil. Iniciando con la definición del coeficiente de correlación (ecuación 7.2 en la sección 7.1) y al multiplicar ambos lados por $\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = s_y / s_x$ se obtiene el resultado

$$\hat{\beta}_1 = r \frac{s_y}{s_x} \quad (7.18)$$

La ecuación 7.18 permite interpretar la pendiente de la recta de mínimos cuadrados en función del coeficiente de correlación. Las unidades de $\hat{\beta}_1$, la pendiente de la recta de mínimos cuadrados, deben ser unidades de y/x . El coeficiente de correlación r es un número puro que mide la fuerza de la relación lineal entre x y y . La ecuación (7.18) muestra que la pendiente $\hat{\beta}_1$ es proporcional al coeficiente de correlación, donde la constante de proporcionalidad es la cantidad s_y/s_x para que las unidades concuerden con las unidades en las que se miden x y y .

Con la ecuación (7.13) se puede escribir la recta de mínimos cuadrados en su forma punto pendiente: Sustituyendo $\bar{y} - \hat{\beta}_1 \bar{x}$ por $\hat{\beta}_0$ en la ecuación para la recta de mínimos cuadrados $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ y reordenando los términos se obtiene

$$\hat{y} - \bar{y} = \hat{\beta}_1 (x - \bar{x}) \quad (7.19)$$

Combinando las ecuaciones (7.18) y (7.19) se tiene que

$$\hat{y} - \bar{y} = r \frac{s_y}{s_x} (x - \bar{x}) \quad (7.20)$$

Por consiguiente, la recta de mínimos cuadrados es la que pasa a través del centro de masa del diagrama de dispersión $(\bar{x}, \bar{y})$, con pendiente $\hat{\beta} = r(s_y/s_x)$.

Medición de la bondad del ajuste

Un estadístico de la bondad del ajuste representa una cantidad que mide qué tan bien un modelo explica un conjunto específico de datos. Un modelo lineal ajusta bien si hay una fuerte relación lineal entre x y y . En la sección 7.1 se mencionó que el coeficiente de correlación r mide la fuerza de la relación lineal entre x y y . Por tanto, r es un estadístico de la bondad del ajuste para el modelo lineal. Ahora se describirá cómo mide r la bondad del ajuste. La figura 7.12 presenta los datos de Galton acerca de las longitudes de antebrazos contra la estatura. Los puntos en el diagrama de dispersión son (x_i, y_i) , donde x_i significa la estatura del hombre del i -ésimo y y_i es la longitud de su antebrazo. Tanto la recta de mínimos cuadrados como la horizontal $y = \bar{y}$ están sobrepuestas. Ahora imagine que se debe predecir la longitud de uno de los antebrazos. Si no se conoce la estatura del hombre, se debe predecir que la longitud del antebrazo es el promedio de $\bar{y}$. Los errores de predicción son $y_i - \bar{y}$. Si se pronostica de esta manera la longitud de cada antebrazo, la suma de los errores pronosticados al cuadrado será $\sum_{i=1}^n (y_i - \bar{y})^2$. Si, por otro lado, se conoce la estatura de cada hombre antes de pronosticar

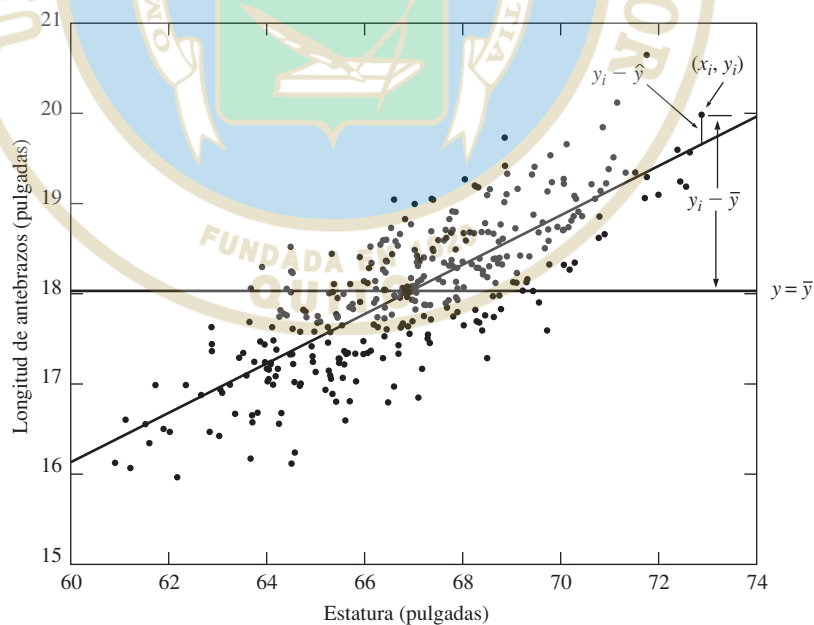


FIGURA 7.12 Estatura y longitud de antebrazos de hombres. La recta de mínimos cuadrados y la horizontal $y = \bar{y}$ están sobrepuestas.

la longitud de su antebrazo se puede utilizar la recta de mínimos cuadrados, y predecir que la longitud del i -ésimo antebrazo sea $\hat{y}_i$. El error pronosticado será el residuo $y_i - \hat{y}_i$, y la suma de los errores pronosticados al cuadrado es: $\sum_{i=1}^n (y_i - \hat{y}_i)^2$. La fuerza de la relación lineal se mide al calcular la reducción obtenida en la suma de los errores pronosticados al cuadrado con $\hat{y}_i$ en lugar de $\bar{y}$. Esto es la diferencia $\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$. Mientras más grande sea la diferencia, más fuertes serán la agrupación de puntos alrededor de la recta de mínimos cuadrados y la relación lineal entre x y y . Por tanto, $\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$ es un estadístico de la bondad del ajuste.

Sin embargo, existe un problema al utilizar $\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$ como un estadístico de la bondad del ajuste. Esta cantidad tiene unidades, a saber, unidades cuadradas de y . No se podía usar este estadístico para comparar la bondad del ajuste de dos modelos que ajusten diferentes conjuntos de datos, puesto que las unidades serían diferentes. Se necesita utilizar un estadístico de la bondad del ajuste que sea un número puro para que se pueda medir la bondad del ajuste en una escala absoluta.

Aquí es donde entra el **coeficiente de correlación** r . Se demuestra al final de esta sección que

$$r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (7.21)$$

La cantidad r^2 , el cuadrado del coeficiente de correlación, se llama **coeficiente de la determinación**. Ésta es la reducción obtenida en la suma de los errores pronosticados al cuadrado al utilizar $\hat{y}_i$ en lugar de $\bar{y}$, expresado como una fracción de la suma de los errores pronosticados al cuadrado $\sum_{i=1}^n (y_i - \bar{y})^2$, obtenidos al usar $\bar{y}$. Es importante conocer esta interpretación de r^2 . En el capítulo 8 se verá cómo se puede generalizar para proporcionar una medida de la bondad del ajuste para relaciones lineales que impliquen varias variables.

Para una interpretación visual de r^2 , observe la figura 7.12. Para cada punto (x_i, y_i) en el diagrama de dispersión, la cantidad $y_i - \bar{y}$ constituye la distancia vertical del punto a la recta horizontal $y = \bar{y}$. La cantidad $y_i - \hat{y}_i$ es la distancia vertical del punto a la recta de mínimos cuadrados. Por tanto, la cantidad $\sum_{i=1}^n (y_i - \bar{y})^2$ mide la dispersión total de los puntos alrededor de la recta $y = \bar{y}$, y la cantidad $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ mide la dispersión total de los puntos alrededor de la recta de mínimos cuadrados. Por consiguiente, la cantidad $\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$ mide la reducción en la dispersión de los puntos obtenidos usando mínimos cuadrados en lugar de $y = \bar{y}$. El coeficiente de determinación r^2 expresa esta reducción como una proporción de la dispersión alrededor de $y = \bar{y}$.

Las sumas de los cuadrados que aparecen en este análisis se utilizan tan a menudo que los estadígrafos les han dado nombres. Llamamos a $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ la **suma de los cuadrados de los errores** y a $\sum_{i=1}^n (y_i - \bar{y})^2$ la **suma total de los cuadrados**. Su diferencia $\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$ se denomina **suma de los cuadrados de la regresión**. Evidentemente, la siguiente relación es válida:

$$\begin{array}{l} \text{Suma total} \\ \text{de los cuadrados} \end{array} = \begin{array}{l} \text{Suma de los cuadrados} \\ \text{de la regresión} \end{array} + \begin{array}{l} \text{Suma de los cuadrados} \\ \text{de los errores} \end{array}$$

Mediante la terminología anterior se puede escribir la ecuación (7.21) como

$$r^2 = \frac{\text{Suma de los cuadrados de la regresión}}{\text{Suma total de los cuadrados}}$$

Puesto que la suma total de cuadrados es exactamente la varianza muestral de los y_i sin dividir entre $n - 1$, los estadígrafos (y otros) con frecuencia se refieren a r^2 como la **proporción de la varianza en y explicada por la regresión**.

Deducción de los coeficientes de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$

Se deducirán las ecuaciones (7.13) y (7.14). Los coeficientes de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$ son las cantidades que minimizan la suma

$$S = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Se calculan estos valores mediante derivadas parciales de S con respecto a $\hat{\beta}_0$ y $\hat{\beta}_1$ e igualándolas a 0. Por tanto, $\hat{\beta}_0$ y $\hat{\beta}_1$ son las cantidades que resuelven las ecuaciones simultáneas

$$\frac{\partial S}{\partial \hat{\beta}_0} = - \sum_{i=1}^n 2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (7.22)$$

$$\frac{\partial S}{\partial \hat{\beta}_1} = - \sum_{i=1}^n 2x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (7.23)$$

Estas cantidades se pueden escribir como un sistema de ecuaciones lineales con dos incógnitas:

$$n\hat{\beta}_0 + \left(\sum_{i=1}^n x_i \right) \hat{\beta}_1 = \sum_{i=1}^n y_i \quad (7.24)$$

$$\left(\sum_{i=1}^n x_i \right) \hat{\beta}_0 + \left(\sum_{i=1}^n x_i^2 \right) \hat{\beta}_1 = \sum_{i=1}^n x_i y_i \quad (7.25)$$

Se despeja a $\hat{\beta}_0$ de la ecuación (7.24), obteniendo

$$\begin{aligned} \hat{\beta}_0 &= \frac{\sum_{i=1}^n y_i}{n} - \hat{\beta}_1 \frac{\sum_{i=1}^n x_i}{n} \\ &= \bar{y} - \hat{\beta}_1 \bar{x} \end{aligned}$$

Esto demuestra la ecuación (7.14). Ahora se sustituye $\bar{y} - \hat{\beta}_1 \bar{x}$ en lugar de $\hat{\beta}_0$ en la ecuación (7.25), con lo que se tiene que

$$\left(\sum_{i=1}^n x_i \right) (\bar{y} - \hat{\beta}_1 \bar{x}) + \left(\sum_{i=1}^n x_i^2 \right) \hat{\beta}_1 = \sum_{i=1}^n x_i y_i \quad (7.26)$$

Despejando a $\hat{\beta}_1$ de la ecuación (7.26), se tiene que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$$

Para demostrar la ecuación (7.13) se debe mostrar que $\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$ y que $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$. (Éstas son las ecuaciones 7.15 y 7.17.)

Ahora

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \sum_{i=1}^n (x_i y_i - \bar{x} y_i - \bar{y} x_i + \bar{x} \bar{y}) \\ &= \sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i - \bar{y} \sum_{i=1}^n x_i + \sum_{i=1}^n \bar{x} \bar{y} \\ &= \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} - n\bar{y}\bar{x} + n\bar{x}\bar{y} \\ &= \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \end{aligned}$$

También

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n (x_i^2 - 2\bar{x}x_i + \bar{x}^2) \\ &= \sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + \sum_{i=1}^n \bar{x}^2 \\ &= \sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum_{i=1}^n x_i^2 - n\bar{x}^2 \end{aligned}$$

Deducción de la ecuación (7.21)

Primero se demuestra que

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (7.27)$$

Este resultado se conoce como la **identidad del análisis de varianza**. Para deducirla se comienza por sumar y restar $\hat{y}_i$ en el lado izquierdo:

$$\begin{aligned} \sum_{i=1}^n (y_i - \bar{y})^2 &= \sum_{i=1}^n [(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})]^2 \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) \end{aligned}$$

Ahora sólo se necesita demostrar que $\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 0$. Puesto que $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ y $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$,

$$\hat{y}_i = \bar{y} + \hat{\beta}_1 (x_i - \bar{x}) \quad (7.28)$$

Por tanto,

$$\begin{aligned}\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) &= \sum_{i=1}^n [(y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})][\hat{\beta}_1(x_i - \bar{x})] \\ &= \hat{\beta}_1 \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) - \hat{\beta}_1^2 \sum_{i=1}^n (x_i - \bar{x})^2\end{aligned}\quad (7.29)$$

Ahora $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$, por lo que

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \hat{\beta}_1 \sum_{i=1}^n (x_i - \bar{x})^2$$

Sustituyendo en la ecuación (7.29) se obtiene

$$\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = \hat{\beta}_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 - \hat{\beta}_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 = 0$$

Esto último demuestra la identidad del análisis de varianza.

Para deducir la ecuación (7.21), la ecuación (7.28) implica que

$$\hat{y}_i - \bar{y} = \hat{\beta}_1(x_i - \bar{x})\quad (7.30)$$

Elevando al cuadrado ambos lados de la ecuación (7.30) y sumando, se obtiene

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \hat{\beta}_1^2 \sum_{i=1}^n (x_i - \bar{x})^2$$

Ahora $\hat{\beta}_1 = r \frac{s_y}{s_x}$ (ecuación 7.18), por lo que

$$\hat{\beta}_1^2 = r^2 \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Sustituyendo y eliminando, se obtiene

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = r^2 \sum_{i=1}^n (y_i - \bar{y})^2$$

por lo que

$$r^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Por la identidad del análisis de varianza, $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$. Por tanto,

$$r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Ejercicios de la sección 7.2

1. Todos los meses, durante cierto periodo, se midieron la temperatura promedio en °C (x) y el número de libras de vapor (y) consumidas por cierta planta química. La recta de mínimos cuadrados calculada de los datos resultantes es $y = 245.82 + 1.13x$.
 - a) Pronostique el número de libras del vapor consumido en un mes donde la temperatura promedio es 65°C.
 - b) Si dos meses difieren en sus temperaturas promedio por 5°C, ¿cuánto predice que será diferente el número de libras del vapor consumido?
2. En un estudio de la relación entre la dureza de Brinell (x) y la tensión de compresión en ksi (y) de elementos de cobre extraídos en frío, la recta de mínimos cuadrados fue $y = -196.32 + 2.42x$.
 - a) Pronostique la fuerza de tensión de un elemento cuya dureza de Brinell es 102.7.
 - b) Si las dos muestras difieren en su dureza de Brinell en 3, ¿cuánto predice que serán diferentes sus tensiones de compresión?
3. Una recta de mínimos cuadrados está ajustando a un conjunto de puntos. Si la suma total de los cuadrados es $\sum (y_i - \bar{y})^2 = 9\,615$, y la suma de los cuadrados de los errores es $\sum (y_i - \hat{y}_i)^2 = 1\,450$, calcule el coeficiente de determinación r^2 .
4. Una recta de mínimos cuadrados está ajustando un conjunto de puntos. Si la suma total de cuadrados es $\sum (y_i - \bar{y})^2 = 181.2$, y la suma de los cuadrados de los errores es $\sum (y_i - \hat{y}_i)^2 = 33.9$, calcule el coeficiente de determinación r^2 .
5. Con los datos de las estaturas de Galton (figura 7.1, en la sección 7.1), la recta de mínimos cuadrados para pronosticar la longitud del antebrazo (y) de la estatura (x) es $y = -0.2967 + 0.2738x$.
 - a) Pronostique la longitud del antebrazo de un hombre cuya estatura es 70 pulg.
 - b) ¿Qué estatura debe tener un hombre con el propósito de que se pronostique que su longitud de antebrazo sea de 19 pulg?
 - c) Todos los hombres en cierto grupo tienen mayor estatura que las que se calcularon en el inciso b). ¿Puede concluir que todos sus antebrazos serán de al menos 19 pulg de longitud? Explique.
6. En un estudio que relaciona el grado de deformación, en mm, de una placa de cobre (y) a la temperatura en °C (x) se calculó el siguiente resumen estadístico: $n = 40$, $\sum_{i=1}^n (x_i - \bar{x})^2 = 98\,775$, $\sum_{i=1}^n (y_i - \bar{y})^2 = 19.10$, $\bar{x} = 26.36$, $\bar{y} = 0.5188$, $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 826.94$.
 - a) Calcule la correlación r entre el grado de deformación y la temperatura.
 - b) Calcule la suma de cuadrados del error, la suma de cuadrados de regresión y la suma total de cuadrados.
 - c) Calcule la recta de mínimos cuadrados para pronosticar la deformación de la temperatura.
 - d) Pronostique la deformación a una temperatura de 40°C.
 - e) ¿A qué temperatura se pronostica que la deformación será de 0.5 mm?
 - f) Suponga que es importante que la deformación no sea mayor de 0.5 mm. Un ingeniero sugiere que si la temperatura se conserva debajo del nivel calculado en el inciso e) se puede tener la seguridad de que la deformación no será mayor de 0.5 mm. ¿Esta es una conclusión correcta? Explique.
7. Se midió el peso inercial (en toneladas) y el ahorro de combustible (en milla/galón) para una muestra de siete camiones de diesel. En la tabla siguiente se presentan los resultados. (De "In-Use Emissions from Heavy-Duty Diesel Vehicles," J. Yanowitz, tesis de doctorado, Escuela de Minas de Colorado, 2001.)

Peso	Millaje
8.00	7.69
24.50	4.97
27.00	4.56
14.50	6.49
28.50	4.34
12.75	6.24
21.25	4.45

 - a) Construya un diagrama de puntos del millaje (y) contra el peso (x). Compruebe que un modelo lineal es adecuado.
 - b) Calcule la recta de mínimos cuadrados para pronosticar el millaje a partir del peso.
 - c) Si los dos camiones son diferentes en peso por cinco toneladas, ¿cuánto predeciría que son diferentes sus millajes?
 - d) Pronostique el millaje para camiones con un peso de 15 toneladas.

- e) ¿Cuáles son las unidades de la pendiente estimada $\hat{\beta}_1$?
- f) ¿Cuáles son las unidades del intercepto estimado $\hat{\beta}_0$?

8. El procesamiento de carbón natural implica el “lavado” durante el cual se elimina ceniza de carbón (no orgánico, material no combustible). El artículo “Quantifying Sampling Precision for Coal Ash Using Gy’s Discrete Model of the Fundamental Error” (*Journal of Coal Quality*, 1989, 33-39) proporciona los datos relacionados con los porcentajes de ceniza con la densidad de una partícula de carbón. Se midió el promedio de porcentaje de ceniza para cinco densidades de partículas de carbón. En la tabla siguiente se presentan los datos:

Densidad (g/cm ³)	Porcentaje de ceniza
1.25	1.93
1.325	4.63
1.375	8.95
1.45	15.05
1.55	23.31

- a) Construya un diagrama de dispersión del porcentaje de ceniza (y) contra la densidad (x). Verifique que es adecuado un modelo lineal.
- b) Calcule la recta de mínimos cuadrados para pronosticar porcentaje de ceniza a partir de la densidad.
- c) Si las dos partículas de carbón difieren en densidad por 0.1 g/cm³, ¿cuánto predeciría que será diferente el porcentaje de ceniza?
- d) Pronostique el porcentaje de ceniza para partículas con 1.40 g/cm³.
- e) Calcule los valores ajustados.
- f) Calcule los residuos. ¿Qué punto tiene el residuo con magnitud mayor?
- g) Calcule la correlación entre la densidad y el porcentaje de ceniza.
- h) Calcule la suma de cuadrados de regresión, la suma de cuadrados del error y la suma total de cuadrados.
- i) Divida la suma de cuadrados de regresión entre la suma total de cuadrados. ¿Cuál es la relación entre esta cantidad y el coeficiente de correlación?

9. En pruebas diseñadas para medir el efecto de cierto aditivo en el tiempo de secado de pintura se obtuvieron los siguientes datos.

Concentración de aditivo (%)	Tiempo de secado (h)
4.0	8.7
4.2	8.8
4.4	8.3
4.6	8.7
4.8	8.1
5.0	8.0
5.2	8.1
5.4	7.7
5.6	7.5
5.8	7.2

- a) Construya un diagrama de dispersión del tiempo de secado (y) contra la concentración del aditivo (x). Verifique que es adecuado un modelo lineal.
 - b) Calcule la recta de mínimos cuadrados para pronosticar el tiempo de secado a partir de la concentración del aditivo.
 - c) Calcule el valor ajustado y el residuo para cada punto.
 - d) Si la concentración del aditivo está aumentando en 0.1%, ¿en cuánto predeciría que aumente o disminuya el tiempo de secado?
 - e) Pronostique el tiempo de secado para una concentración de 4.4%.
 - f) ¿Puede utilizarse la recta de mínimos cuadrados para pronosticar el tiempo de secado respecto a una concentración de 7%? Si es así, pronostique el tiempo de secado. Si no, explique por qué.
 - g) ¿Para qué concentración pronosticaría un tiempo de secado de 8.2 horas?
 - h) El objetivo de este proyecto es reducir el tiempo de secado a seis horas. Con base en los datos dados, ¿puede especificar una concentración que probablemente ocasiona este resultado? Si es así, especifique la concentración. Si no, explique por qué.
10. Los tiempos de curado en días (x) y las fuerzas compresivas en MPa (y) se registraron para varios elementos de concreto. Las medias y las desviaciones estándar de los valores x y y fueron $\bar{x} = 5$, $s_x = 2$, $\bar{y} = 1\,350$, $s_y = 100$. La correlación entre tiempos de curado y las fuerzas compresivas se calculó de $r = 0.7$. Encuentre la ecuación de la recta de mínimos cuadrados para pronosticar la fuerza compresiva a partir de los tiempos de curado.
11. Diferentes cantidades de pectina se agregaron a gelatina enlatada para estudiar la relación entre la concentración de

pectina en % (x) y el índice de firmeza (y). Las medias y las desviaciones estándar de los valores de x y de y fueron $\bar{x} = 3$, $s_x = 0.5$, $\bar{y} = 50$, $s_y = 10$. La correlación entre los tiempos de cura y la firmeza se calculó de $r = 0.5$. Determine la ecuación de la recta de mínimos cuadrados para pronosticar la firmeza a partir de la concentración de pectina.

12. Un ingeniero quiere pronosticar el valor de y cuando $x = 4.5$ utilizando el siguiente conjunto de datos.

x	y	$z = \ln y$	x	y	$z = \ln y$
1	0.2	-1.61	6	2.3	0.83
2	0.3	-1.20	7	2.9	1.06
3	0.5	-0.69	8	4.5	1.50
4	0.5	-0.69	9	8.7	2.16
5	1.3	0.26	10	12.0	2.48

- a) Construya un diagrama de dispersión de los puntos (x, y).
- b) ¿La recta de mínimos cuadrados se debe utilizar para pronosticar el valor de y cuando $x = 4.5$? Si es así, calcule la recta de mínimos cuadrados y el valor pronosticado. Si no, explique.
- c) Construya un diagrama de dispersión de los puntos (x, z), donde $z = \ln y$.
- d) Utilice la recta de mínimos cuadrados para pronosticar el valor de z cuando $x = 4.5$. ¿Este es un método adecuado de pronóstico? Explique por qué sí o no.
- e) Sea $\hat{z}$ el valor pronosticado de z calculado en el inciso d). Sea $\hat{y} = e^{\hat{z}}$. Explique por qué $\hat{y}$ es un pronosticador razonable del valor de y cuando $x = 4.5$.

13. Una muestra aleatoria simple de 100 hombres entre 25 y 34 años de edad promediaron 70 pulgadas de estatura, y tenían desviación estándar de tres pulgadas. Sus ingresos promediaron \$34 900 y tenían desviación estándar de \$17 200. Complete el espacio en blanco: a partir de la recta de mínimos cuadrados se predeciría que los ingresos de un hombre de 70 pulgadas de estatura serían de_____

- i. Menor que \$34 900.
- ii. Mayor que \$34 900.
- iii. Igual a \$34 900.
- iv. No se puede saber a menos que se conozca la correlación.

14. Una mezcla de sacarosa y agua se calentó en una plancha y la temperatura (en °C) se registró cada minuto durante 20 minutos con tres termómetros de termopar. Los resultados se muestran en la tabla siguiente.

Tiempo	T_1	T_2	T_3
0	20	18	21
1	18	22	11
2	29	22	26
3	32	25	35
4	37	37	33
5	36	46	35
6	46	45	44
7	46	44	43
8	56	54	63
9	58	64	68
10	64	69	62
11	72	65	65
12	79	80	80
13	84	74	75
14	82	87	78
15	87	93	88
16	98	90	91
17	103	100	103
18	101	98	109
19	103	103	107
20	102	103	104

- a) Calcule la recta de mínimos cuadrados para estimar la temperatura como función del tiempo usando T_1 como el valor para la temperatura.
- b) Calcule la recta de mínimos cuadrados para estimar la temperatura como una función del tiempo usando T_2 como el valor para la temperatura.
- c) Calcule la recta de mínimos cuadrados para estimar la temperatura como una función del tiempo usando T_3 como el valor para la temperatura.
- d) Se desea calcular una sola recta para calcular la temperatura como una función del tiempo. Una persona sugiere promediar los tres estimadores de la pendiente para obtener una sola estimación de la pendiente, y promediar los tres estimadores del intercepto para obtener una sola estimación del intercepto. Determine la ecuación de la recta que es resultado de este método.
- e) Otra persona sugiere promediar las tres mediciones de la temperatura cada vez para obtener $\bar{T} = (T_1 + T_2 + T_3)/3$. Calcule la recta de mínimos cuadrados usando $\bar{T}$ como el valor para la temperatura.
- f) ¿Los resultados de los incisos d) y e) son diferentes?

7.3 Incertidumbres en los coeficientes de mínimos cuadrados

En la sección 7.2 se presentó el modelo lineal (ecuación 7.9):

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Aquí ε_i es el error en la i -ésima observación y_i . En la práctica, ε_i representa la acumulación del error de muchas fuentes. Por ejemplo, en los datos de la ley de Hooke puede estar afectado por los errores al medir la longitud del resorte, al medir los pesos de las cargas puestas en el resorte, las variaciones en la elasticidad del resorte debidas a los cambios en la temperatura ambiental o al cansancio mental, etcétera. Si no hubiera error los puntos se encontrarían exactamente en la recta de mínimos cuadrados, y la pendiente $\hat{\beta}_1$ y el intercepto $\hat{\beta}_0$ de la recta de mínimos cuadrados serían iguales a los valores verdaderos β_0 y β_1 . Debido al error los puntos están dispersos alrededor de la recta, y las cantidades $\hat{\beta}_0$ y $\hat{\beta}_1$ no igualan los valores verdaderos. Cada vez que se repite el proceso, los valores de ε_i y, por tanto, los valores de $\hat{\beta}_0$ y $\hat{\beta}_1$ serán diferentes. En otras palabras, ε_i , $\hat{\beta}_0$, y $\hat{\beta}_1$ son variables aleatorias. Más específicamente, el error ε_i crea incertidumbre en los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$. Es obvio que si el ε_i tiende a tener una magnitud pequeña, los puntos estarán muy agrupados alrededor de la recta, y la incertidumbre en los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$ será pequeña. Por otro lado, si el ε_i tiende a tener mayor magnitud, los puntos estarán muy dispersos alrededor de la recta, y la incertidumbre (desviaciones estándar) en los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$ será mayor.

Suponga que se tienen n puntos de datos $(x_1, y_1), \dots, (x_n, y_n)$ y se planea ajustar la recta de mínimos cuadrados. Con la finalidad de que los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ sean útiles, se necesita estimar qué tan grandes son sus incertidumbres. Para realizar este cálculo se necesita conocer algo acerca de la naturaleza de los errores ε_i . Se iniciará por estudiar la situación más simple, en la cual se satisfacen cuatro supuestos importantes. Éstos se presentan en el cuadro siguiente.

Supuestos para los errores en los modelos lineales

En la situación más simple los siguientes supuestos se satisfacen:

1. Los errores $\varepsilon_1, \dots, \varepsilon_n$ son aleatorios e independientes. En particular, la magnitud de cualquier error ε_i no influye en el valor del siguiente error ε_{i+1} .
2. Todos los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen media 0.
3. Todos los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen la misma varianza, que se denota por σ^2 .
4. Los errores $\varepsilon_1, \dots, \varepsilon_n$ están distribuidos normalmente.

Estos supuestos son restrictivos, por lo que es útil analizar brevemente el grado con el cual es aceptable incumplirlos en la práctica. Cuando el tamaño de muestra es grande, el supuesto de normalidad (4) será menos importante. Los incumplimientos moderados del supuesto de la varianza constante (3) no importa demasiado, pero se deben corregir los graves. En la sección 7.4 se analizarán los métodos para corregir algunos incumplimientos de estos supuestos.

Bajo estos supuestos, el efecto del ε_i lo gobierna en gran parte la magnitud de la varianza σ^2 , ya que ésta determina qué tan grandes son los errores. Por tanto, con la finalidad de calcular las incertidumbres en $\hat{\beta}_0$ y $\hat{\beta}_1$, primero se debe estimar la varianza del error σ^2 . Debido a que la magnitud de la varianza se refleja en el grado de dispersión de los puntos alrededor de la recta de mínimos cuadrados, se tiene que midiendo esta dispersión se puede calcular la varianza. Específicamente, la distancia vertical desde cada punto de datos (x_i, y_i) a la recta de mínimos cuadrados está dada por el residuo e_i (véase la figura 7.10 en la sección 7.2). La dispersión de los puntos alrededor de la recta se puede medir con la suma de los residuos al cuadrado $\sum_{i=1}^n e_i^2$. El estimador de la varianza del error σ^2 es la cantidad s^2 dada por

$$s^2 = \frac{\sum_{i=1}^n e_i^2}{n-2} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2} \quad (7.31)$$

El estimador de la varianza del error es, por tanto, el promedio de los residuos al cuadrado, excepto que se divide entre $n-2$ en lugar de n . La razón para esto se debe a que la recta de mínimos cuadrados minimiza la suma $\sum_{i=1}^n e_i^2$, y los residuos tienden a ser un poco más pequeños que los errores ε_i . Al dividir entre $n-2$ en lugar de n se compensa aproximadamente este efecto.

Hay una fórmula equivalente para s^2 , que implica el coeficiente de correlación r , que es con frecuencia más fácil calcular.

$$s^2 = \frac{(1-r^2) \sum_{i=1}^n (y_i - \bar{y})^2}{n-2} \quad (7.32)$$

Se presenta una deducción breve de este resultado. La ecuación (7.21) (de la sección 7.2) muestra que $1-r^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / \sum_{i=1}^n (y_i - \bar{y})^2$. Entonces $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = (1-r^2) \sum_{i=1}^n (y_i - \bar{y})^2$, y se tiene que

$$s^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2} = \frac{(1-r^2) \sum_{i=1}^n (y_i - \bar{y})^2}{n-2}$$

Bajo los supuestos 1 a 4, las observaciones y_i son también variables aleatorias. Efectivamente, puesto que $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, en consecuencia y_i tiene una distribución normal con media $\beta_0 + \beta_1 x_i$ y la varianza σ^2 . En particular, β_1 representa el cambio en la media de y asociada con un aumento de una unidad en el valor de x .

Resumen

En el modelo lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, bajo los supuestos del 1 al 4, las observaciones $y_1, \dots, y_n$ son variables aleatorias independientes que siguen la distribución normal. La media y la varianza de y_i están dadas por

$$\mu_{y_i} = \beta_0 + \beta_1 x_i$$

$$\sigma_{y_i}^2 = \sigma^2$$

La pendiente β_1 representa el cambio en la media de y asociada con un aumento de una unidad en el valor de x .

Ahora se pueden calcular las medias y las desviaciones estándar de $\hat{\beta}_0$ y $\hat{\beta}_1$. Las desviaciones estándar son, por supuesto, las incertidumbres. Tanto $\hat{\beta}_0$ como $\hat{\beta}_1$ se pueden expresar como combinaciones lineales de las y_i , por lo que sus medias se pueden determinar mediante la ecuación (2.49) y sus desviaciones estándar se determinan con la ecuación (2.53) (ambas ecuaciones de la sección 2.5). Específicamente, al manipular en forma algebraica las ecuaciones (7.13) y (7.14) (en la sección 7.2) se obtiene

$$\hat{\beta}_1 = \sum_{i=1}^n \left[\frac{(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] y_i \quad (7.33)$$

$$\hat{\beta}_0 = \sum_{i=1}^n \left[\frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] y_i \quad (7.34)$$

Utilizando el hecho de que cada una de las y_i tiene una media $\beta_0 + \beta_1 x_i$ y una varianza σ^2 , de las ecuaciones (2.49) y (2.53) se obtienen los siguientes resultados, después de mucha manipulación:

$$\begin{aligned} \mu_{\hat{\beta}_0} &= \beta_0 & \mu_{\hat{\beta}_1} &= \beta_1 \\ \sigma_{\hat{\beta}_0} &= \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} & \sigma_{\hat{\beta}_1} &= \frac{\sigma}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \end{aligned}$$

Los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ no tienen sesgo, ya que sus medias son iguales a los valores verdaderos. También están distribuidos normalmente, porque son combinaciones lineales de las variables aleatorias normales independientes y_i . En la práctica, cuando se calculan las desviaciones estándar, por lo general no se conoce el valor de σ , por eso no se aproxima con s .

Resumen

Bajo los supuestos 1 al 4 (p. 508),

- Las cantidades $\hat{\beta}_0$ y $\hat{\beta}_1$ son variables aleatorias normalmente distribuidas.
- Las medias de $\hat{\beta}_0$ y $\hat{\beta}_1$ son los valores verdaderos β_0 y β_1 , respectivamente.
- Las desviaciones estándar de $\hat{\beta}_0$ y $\hat{\beta}_1$ se calculan con

$$s_{\hat{\beta}_0} = s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7.35)$$

y

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7.36)$$

donde $s = \sqrt{\frac{(1 - r^2) \sum_{i=1}^n (y_i - \bar{y})^2}{n - 2}}$ representa un estimador de la desviación estándar del error σ .

Ejemplo

7.9

Para los datos de la ley de Hooke, calcule s , $s_{\hat{\beta}_1}$, y $s_{\hat{\beta}_0}$. Estime la constante del resorte y la longitud sin carga, y determine sus incertidumbres.

Solución

En el ejemplo 7.6 (en la sección 7.2) se calculó $\bar{x} = 1.9000$, $\bar{y} = 5.3885$, $\sum_{i=1}^n (x_i - \bar{x})^2 = 26.6000$, y $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 5.4430$. Ahora calcule $\sum_{i=1}^n (y_i - \bar{y})^2 = 1.1733$. La correlación es $r = 5.4430 / \sqrt{(26.6000)(1.1733)} = 0.9743$.

$$\text{Utilizando la ecuación (7.32), } s = \sqrt{\frac{(1 - 0.9743^2)(1.1733)}{18}} = 0.0575.$$

$$\text{Utilizando la ecuación (7.35), } s_{\hat{\beta}_0} = 0.0575 \sqrt{\frac{1}{20} + \frac{1.9000^2}{26.6000}} = 0.0248.$$

$$\text{Utilizando la ecuación (7.36), } s_{\hat{\beta}_1} = \frac{0.0575}{\sqrt{26.6000}} = 0.0111.$$

Los valores x más dispersos son mejores (dentro del razonamiento)

En las expresiones para las dos incertidumbres $s_{\hat{\beta}_0}$ y $s_{\hat{\beta}_1}$ en las ecuaciones (7.35) y (7.36), la cantidad $\sum_{i=1}^n (x_i - \bar{x})^2$ aparece en un denominador. Esta cantidad mide la dispersión de los valores x ; cuando se divide por la constante $n - 1$, es exactamente la varianza muestral de los valores x . Como resultado, todas las otras cosas siguen igual, un experimento realizado con mayor dispersión de los valores x dará como resultado incertidumbres más pequeñas para $\hat{\beta}_0$ y $\hat{\beta}_1$; y por tanto, una estimación más precisa de los valores verdaderos β_0 y β_1 . Por supuesto, es importante no utilizar los valores x tan grandes o tan pequeños que estén fuera del rango para el cual es legítimo el modelo lineal.

Resumen

Cuando se pueden elegir los valores x es mejor que estén muy dispersos. Entre más dispersos estén, más pequeñas serán las incertidumbres en $\hat{\beta}_0$ y $\hat{\beta}_1$.

Específicamente, la incertidumbre $\sigma_{\hat{\beta}_1}$ en $\hat{\beta}_1$ es inversamente proporcional a $\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}$, o de manera equivalente, a la desviación estándar muestral de $x_1, x_2, \dots, x_n$.

Precaución: Si el rango de valores de x se extiende más allá del rango donde el modelo lineal es válido, los resultados no lo serán.

Hay otras dos maneras de mejorar la exactitud de la recta estimada de regresión. Primero, se puede aumentar el tamaño de la suma $\sum_{i=1}^n (x_i - \bar{x})^2$ tomando más observaciones; por consiguiente, sumando más términos a la suma. Y segundo, se puede reducir el tamaño de la varianza σ^2 del error, por ejemplo, midiendo con mayor precisión. Estos dos métodos comúnmente aumentan el costo del proyecto; sin embargo, si sólo se eligen valores de x más ampliamente dispersos con frecuencia no lo hacen.

Ejemplo

7.10

Dos ingenieros están realizando experimentos independientes para estimar una constante del resorte para un resorte dado. El primer ingeniero sugiere medir la longitud del resorte sin carga, y luego poner cargas de 1, 2, 3 y 4 lb. El segundo ingeniero sugiere utilizar cargas de 0, 2, 4, 6 y 8 lb. ¿Cuál resultado será más preciso? ¿Por qué factor?

Solución

La desviación estándar muestral de los números 0, 2, 4, 6, 8 es dos veces mayor que la de los números, 0, 1, 2, 3, 4. Por tanto, la incertidumbre σ_{β_1} para el primer ingeniero es dos veces mayor que para el segundo, por lo que éste estima con una precisión del doble.

Se han hecho dos supuestos en la solución para este ejemplo. Primero, se supuso que la varianza del error σ^2 es la misma para ambos ingenieros. Si ambos están usando los mismos aparatos y el mismo procedimiento de medición, éste podría ser un supuesto seguro. Pero si un ingeniero puede medir con mayor precisión, se necesita considerar lo anterior. Segundo, se ha supuesto que una carga de 8 lb está dentro de la zona elástica del resorte, por lo que el modelo lineal es aplicable en todo el rango de los datos.

Inferencias acerca de la pendiente y del intercepto

Dado un diagrama de dispersión con puntos $(x_1, y_1), \dots, (x_n, y_n)$, se puede calcular la pendiente, $\hat{\beta}_1$ y el intercepto $\hat{\beta}_0$ de la recta de mínimos cuadrados. Se considera que éstos son los estimadores de una pendiente verdadera β_1 e intercepto β_0 . Ahora se explicará cómo usar estos estimadores para determinar intervalos de confianza y realizar pruebas de hipótesis respecto de los valores verdaderos β_1 y β_0 . Los métodos presentados para una media poblacional, basados en la distribución t de Student, se pueden adaptar fácilmente para este propósito.

Se ha visto que bajo los supuestos 1 al 4, $\hat{\beta}_0$ y $\hat{\beta}_1$ con medias β_0 y β_1 , y desviaciones estándar que se han estimado por $s_{\hat{\beta}_0}$ y $s_{\hat{\beta}_1}$. Las cantidades $(\hat{\beta}_0 - \beta_0)/s_{\hat{\beta}_0}$ y $(\hat{\beta}_1 - \beta_1)/s_{\hat{\beta}_1}$ tienen distribuciones t de Student con $n - 2$ grados de libertad. El número de grados de libertad es $n - 2$ porque en el cálculo de $s_{\hat{\beta}_0}$ y $s_{\hat{\beta}_1}$, se divide la suma de los residuos al cuadrado entre $n - 2$. Cuando el tamaño muestral n es lo suficientemente grande, la distribución normal es casi indistinguible de la t de Student y se puede utilizar en su lugar. Sin embargo, la mayoría de los paquetes de programas computacionales usan la distribución t de Student sin considerar el tamaño muestral.

Resumen

Bajo los supuestos 1 al 4, las cantidades $\frac{\hat{\beta}_0 - \beta_0}{s_{\hat{\beta}_0}}$ y $\frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$ tienen la distribución t de Student con $n - 2$ grados de libertad.

Los intervalos de confianza para β_0 y β_1 se pueden deducir de la misma manera como los intervalos de confianza basados en la t de Student para una media poblacional. Sea $t_{n-2, \alpha/2}$

el punto en la curva t de Student con $n - 2$ grados de libertad que le corresponde un área de $\alpha/2$ en la cola derecha.

Los intervalos con un nivel de confianza de $100(1 - \alpha)\%$ para β_0 y β_1 están dados por

$$\hat{\beta}_0 \pm t_{n-2, \alpha/2} \cdot s_{\beta_0} \quad \hat{\beta}_1 \pm t_{n-2, \alpha/2} \cdot s_{\beta_1} \quad (7.37)$$

donde

$$s_{\beta_0} = s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad s_{\beta_1} = \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Se ilustra el método anterior con algunos ejemplos.

Ejemplo

7.11

Determine un intervalo de confianza de 95% para la constante del resorte de los datos de la ley de Hooke.

Solución

La constante del resorte es β_1 . Se ha calculado antes que $\hat{\beta}_1 = 0.2046$ (ejemplo 7.6 de la sección 7.2) y $s_{\beta_1} = 0.0111$ (ejemplo 7.9).

El número de grados de libertad es $n - 2 = 20 - 2 = 18$, por lo que el valor de t para un intervalo de confianza de 95% es $t_{18, .025} = 2.101$. Por tanto, el intervalo de confianza para β_1 es

$$0.2046 \pm (2.101)(0.0111) = 0.2046 \pm 0.0233 = (0.181, 0.228)$$

Se tiene una confianza de 95% de que el aumento de la longitud del resorte dará como resultado un aumento de 1 lb en la carga que se encuentra entre 0.181 y 0.228 pulg. Por supuesto, este intervalo de confianza es válido solamente dentro del rango de los datos (0 a 3.8 lb).

Ejemplo

7.12

En los datos de la ley de Hooke determine un intervalo de confianza de 99% para la longitud del resorte no cargado.

Solución

La longitud no cargada del resorte es β_0 . Se ha calculado $\hat{\beta}_0 = 4.9997$ (ejemplo 7.6) y $s_{\beta_0} = 0.0248$ (ejemplo 7.9).

El número de grados de libertad es $n - 2 = 20 - 2 = 18$, por lo que el valor de t para un intervalo de confianza de 99% es $t_{18, .005} = 2.878$. Por tanto, el intervalo de confianza para β_0 es

$$4.9997 \pm (2.878)(0.0248) = 4.9997 \pm 0.0714 = (4.928, 5.071)$$

Se tiene confianza de 99% de que la longitud del resorte no cargado se encuentra entre 4.928 y 5.071 pulg.

También se pueden realizar pruebas de hipótesis sobre β_0 y β_1 . Se presentan algunos ejemplos.

Ejemplo

7.13

El fabricante del resorte de los datos de la ley de Hooke afirma que la constante del resorte β_1 es de al menos 0.23 pulg/lb. Se ha calculado que la constante del resorte es $\hat{\beta}_1 = 0.2046$ pulg/lb. ¿Se puede concluir que la afirmación del fabricante es falsa?

Solución

Esto último requiere de una prueba de hipótesis. Las hipótesis nula y alternativa son

$$H_0 : \beta_1 \geq 0.23 \quad \text{contra} \quad H_1 : \beta_1 < 0.23$$

La cantidad

$$\frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$$

tiene la distribución t de Student con $n - 2 = 20 - 2 = 18$ grados de libertad. Bajo H_0 se tiene que $\beta_1 = 0.23$. Por tanto, el estadístico de prueba es,

$$\frac{\hat{\beta}_1 - 0.23}{s_{\hat{\beta}_1}}$$

que se ha calculado anteriormente $\hat{\beta}_1 = 0.2046$ y $s_{\hat{\beta}_1} = 0.0248$. Por tanto, el valor del estadístico de prueba es

$$\frac{0.2046 - 0.23}{0.0248} = -1.024$$

Al consultar la tabla t de Student se encuentra que el P -valor se encuentra entre 0.10 y 0.25. No se puede rechazar la afirmación del fabricante con base en estos datos.

La hipótesis nula más comúnmente probada es $H_0 : \beta_1 = 0$. Si esta hipótesis es verdadera, entonces no hay ninguna tendencia de que y aumente o disminuya cuando x aumenta. Esto implica que x y y no tienen ninguna relación lineal. En general, si la hipótesis de que $\beta_1 = 0$ no es rechazada, no se debe utilizar el modelo lineal para pronosticar y a partir de x .

Ejemplo

7.14

La capacidad de una unión soldada de elongarse bajo tensión está afectada por el compuesto químico del metal de soldadura. En un experimento para determinar el efecto del contenido de carbono (x) sobre la elongación (y) se elongaron 39 soldaduras hasta la fractura, y se midió tanto el contenido de carbono (en partes por mil) como la elongación (en %). Se calcularon los siguientes resúmenes estadísticos:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 0.6561 \quad \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = -3.9097 \quad s = 4.3319$$

Suponiendo que x y y siguen un modelo lineal, calcule el cambio estimado en la elongación debido a un aumento de una parte por mil en el contenido de carbono. ¿Se debe utilizar el modelo lineal para pronosticar la elongación del contenido de carbono?

Solución

El modelo lineal es $y = \beta_0 + \beta_1 x + \varepsilon$, y el cambio en la elongación (y) debido a una parte por mil aumentada en el contenido de carbono (x) es β_1 . Las hipótesis nula y alternativa son

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

La hipótesis nula establece que incrementar el contenido de carbono no afecta la elongación, mientras que la hipótesis alternativa establece que lo hace. La cantidad

$$\frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$$

tiene la distribución t de Student con $n - 2 = 39 - 2 = 37$ grados de libertad. Bajo H_0 , $\beta_1 = 0$. Por tanto, el estadístico de prueba es

$$\frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}}$$

Se calculan $\hat{\beta}_1$ y $s_{\hat{\beta}_1}$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{-3.9097}{0.6561} = -5.959$$

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} = 5.348$$

El valor del estadístico de prueba es

$$\frac{-5.959 - 0}{5.348} = -1.114$$

La tabla t muestra que el P -valor es mayor a 0.20. No se puede concluir que el modelo lineal sea útil para pronosticar la elongación a partir del contenido de carbono.

Inferencias sobre la respuesta media

En el ejemplo 7.8 (sección 7.2) se calculó que la longitud de un resorte bajo una carga de 1.4 lb era de 5.29 pulg. Debido a que esta estimación estaba basada en mediciones sujetas a la incertidumbre, la estimación en sí misma está sujeta a la incertidumbre. Para que la estimación sea más útil, se debe construir un intervalo de confianza alrededor de ésta con el propósito de reflejar su incertidumbre. Ahora se describe cómo hacer esto para el caso general donde la carga sobre el resorte es x lb.

Si se hubiera tomado una medición de la longitud y del resorte bajo una carga de x libras, la media de y sería la longitud verdadera (o “respuesta media”) $\beta_0 + \beta_1 x$, donde β_1 es la constante del resorte verdadera y β_0 es la longitud sin carga verdadera del resorte. Calculamos esta longitud con $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. Puesto que $\hat{\beta}_0$ y $\hat{\beta}_1$ están distribuidos por lo general con medias β_0 y β_1 , respectivamente, por lo cual $\hat{y}$ de manera habitual está distribuido con media $\beta_0 + \beta_1 x$.

Al utilizar $\hat{y}$ para determinar un intervalo de confianza se debe conocer su desviación estándar. Esta última se obtiene mediante $\hat{y}$ como una combinación lineal de los y_i y utilizando la ecuación (2.53) (en la sección 2.5). Las ecuaciones (7.33) y (7.34) expresan $\hat{\beta}_1$ y $\hat{\beta}_0$ co-

mo combinaciones lineales de las y_i . Puesto que $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$, estas ecuaciones, después de algunas manipulaciones algebraicas, se tiene que

$$\hat{y} = \sum_{i=1}^n \left[\frac{1}{n} + (x - \bar{x}) \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] y_i \quad (7.38)$$

Ahora se puede usar la ecuación (2.53) para obtener una expresión por la desviación estándar de $\hat{y}$. Aquella depende del error de varianza σ^2 . Puesto que en la práctica no se conoce por lo general el valor de σ , se aproxima con s . La desviación estándar de $\hat{y}$ es aproximada por

$$s_{\hat{y}} = s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7.39)$$

La cantidad $[\hat{y} - (\beta_0 + \beta_1 x)]/s_{\hat{y}}$ tiene una distribución t de Student con $n - 2$ grados de libertad. Ahora se puede proporcionar la expresión para un intervalo de confianza destinado a la respuesta media.

Un intervalo con un nivel de confianza de $100(1 - \alpha)\%$ para la cantidad $\beta_0 + \beta_1 x$ está dado por

$$\hat{\beta}_0 + \hat{\beta}_1 x \pm t_{n-2, \alpha/2} \cdot s_{\hat{y}} \quad (7.40)$$

donde $s_{\hat{y}} = s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

Ejemplo

7.15

Mediante los datos de la ley de Hooke calcule un intervalo de confianza de 95% para la longitud de un resorte bajo una carga de 1.4 lb.

Solución

Se calculará $\hat{y}$, $s_{\hat{y}}$, $\hat{\beta}_0$, y $\hat{\beta}_1$, y se utilizará la expresión (7.40). El número de puntos es $n = 20$. En el ejemplo 7.9 se calculó $s = 0.0575$. En el ejemplo 7.6 (de la sección 7.2) se calculó a $\bar{x} = 1.9$, $\sum_{i=1}^n (x_i - \bar{x})^2 = 26.6$, $\hat{\beta}_1 = 0.2046$ y $\hat{\beta}_0 = 4.9997$. Utilizando $x = 1.4$, ahora se calcula

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 = 4.9997 + (0.2046)(1.4) = 5.286$$

Utilizando la ecuación (7.39) con $x = 1.4$ se obtiene

$$s_{\hat{y}} = 0.0575 \sqrt{\frac{1}{20} + \frac{(1.4 - 1.9)^2}{26.6}} = 0.0140$$

El número de grados de libertad es $n - 2 = 20 - 2 = 18$. Se encuentra que el valor de t es $t_{18, .025} = 2.101$. Sustituyendo en la ecuación (7.40) se determina el intervalo de confianza de 95% para la longitud $\beta_0 + \beta_1(1.4)$ de

$$5.286 \pm (2.101)(0.0140) = 5.286 \pm 0.0294 = (5.26, 5.32)$$

Ejemplo

7.16

En un estudio de la relación entre el contenido de oxígeno (x) y fuerza de prueba final (y) de soldaduras se obtuvieron los datos que se presentan en la tabla siguiente para 29 soldaduras. Aquí el contenido de oxígeno se mide en partes por mil, y la fuerza se mide en ksi. Utilizando un modelo lineal, determine un intervalo de confianza de 95% para la media de la fuerza de las soldaduras con contenido de oxígeno de 1.7 partes por mil. (Del artículo “Advances in Oxygen Equivalence Equations for Predicting the Properties of Titanium Welds”, D. Harwig, W. Ittiwattana y H. Castner, en *The Welding Journal*, 2001:126s–136s.)

Contenido de oxígeno	Fuerza	Contenido de oxígeno	Fuerza	Contenido de oxígeno	Fuerza
1.08	63.00	1.16	68.00	1.17	73.00
1.19	76.00	1.32	79.67	1.40	81.00
1.57	66.33	1.61	71.00	1.69	75.00
1.72	79.67	1.70	81.00	1.71	75.33
1.80	72.50	1.69	68.65	1.63	73.70
1.65	78.40	1.78	84.40	1.70	91.20
1.50	72.00	1.50	75.05	1.60	79.55
1.60	83.20	1.70	84.45	1.60	73.95
1.20	71.85	1.30	70.25	1.30	66.05
1.80	87.15	1.40	68.05		

Solución

Se calculan las cantidades siguientes (se pueden utilizar las fórmulas para el cálculo de la p. 495):

$$\bar{x} = 1.51966 \quad \bar{y} = 75.4966 \quad \sum_{i=1}^n (x_i - \bar{x})^2 = 1.33770 \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 1\,304.23$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 22.6377 \quad \hat{\beta}_0 = 49.7796 \quad \hat{\beta}_1 = 16.9229 \quad s = 5.84090$$

La estimación de la media de la fuerza para soldaduras con contenido de oxígeno de 1.7 es

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(1.7) = 49.7796 + (16.9229)(1.7) = 78.5485$$

La desviación estándar estimada de $\hat{y}$ es

$$\begin{aligned} s_{\hat{y}} &= s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \\ &= 5.84090 \sqrt{\frac{1}{29} + \frac{(1.7 - 1.51966)^2}{1.33770}} \\ &= 1.4163 \end{aligned}$$

Hay $n - 2 = 29 - 2 = 27$ grados de libertad. Por tanto, el valor t es $t_{27, .025} = 2.052$. El intervalo de confianza de 95% es

$$78.5485 \pm (2.052)(1.4163) = 78.5485 \pm 2.9062 = (75.64, 81.45)$$

Se pueden realizar pruebas de hipótesis sobre la respuesta media, utilizando una distribución t de Student. A continuación se presenta un ejemplo.

Ejemplo

7.17

Con referencia al ejemplo 7.15, sea μ_0 la longitud verdadera del resorte bajo una carga de 1.6 lb. Pruebe la hipótesis $H_0 : \mu_0 \leq 5.3$ contra $H_1 : \mu_0 > 5.3$.

Solución

Ya que μ_0 es la longitud verdadera del resorte bajo una carga de 1.6 lb, $\mu_0 = \beta_0 + \beta_1(1.6)$. Ahora sea $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(1.6)$. La cantidad

$$\frac{\hat{y} - [\beta_0 + \beta_1(1.6)]}{s_{\hat{y}}} = \frac{\hat{y} - \mu_0}{s_{\hat{y}}}$$

tiene la distribución t de Student con $n - 2 = 18$ grados de libertad. Bajo H_0 se tiene que $\mu_0 = 5.3$. Por tanto, el estadístico de prueba es

$$\frac{\hat{y} - 5.3}{s_{\hat{y}}}$$

Se calculan $\hat{y}$ y $s_{\hat{y}}$:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(1.6) = 4.9997 + (0.2046)(1.6) = 5.3271$$

$$s_{\hat{y}} = 0.0575 \sqrt{\frac{1}{20} + \frac{(1.6 - 1.9)^2}{26.6000}} = 0.0133$$

El valor del estadístico de prueba es

$$\frac{5.3271 - 5.3}{0.0133} = 2.04$$

El P -valor se encuentra entre 0.025 y 0.05. Es razonable concluir que la longitud verdadera es mayor a 5.3 pulg.

Intervalos de predicción para futuras observaciones

En el ejemplo 7.16 se encontró un intervalo de confianza para la media de la fuerza de soldaduras con contenido de oxígeno de 1.7 partes por mil. He aquí un escenario algo diferente: suponga que se desea pronosticar la fuerza de una soldadura dada cuyo contenido de oxígeno es 1.7, en lugar de la media de la fuerza de todas estas soldaduras.

Utilizando los valores calculados en el ejemplo 7.16, se pronostica que la fuerza de esta soldadura es $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(1.7) = 49.7796 + (16.9229)(1.7) = 78.5485$. Este pronóstico es el mismo que la estimación de la media de la fuerza de todas las soldaduras con un contenido de oxígeno de 1.7. Ahora se quiere poner un intervalo alrededor de este pronóstico para indicar su incertidumbre. Con el propósito de calcular este **intervalo de predicción** es preciso determinar la incertidumbre en el pronóstico.

La media de la fuerza de las soldaduras con contenido de oxígeno de 1.7 es $\beta_0 + \beta_1(1.7)$. La fuerza verdadera de una soldadura dada es igual a $\beta_0 + \beta_1(1.7) + \varepsilon$, donde ε representa la diferencia aleatoria entre la fuerza de la soldadura dada y la media de la fuerza de

todas las soldaduras cuyo contenido de oxígeno es 1.7. El error en la predicción de la fuerza de una soldadura dada con $\hat{y}$ representa el error de la predicción

$$\hat{y} - [\beta_0 + \beta_1(1.7)] = \varepsilon \quad (7.41)$$

La incertidumbre en la predicción de la fuerza de la soldadura dada es la desviación estándar de este error de predicción. Se mostrará brevemente cómo calcular esta desviación estándar. La cantidad $\beta_0 + \beta_1(1.7)$ es constante y no afecta la desviación estándar. Las cantidades $\hat{y}$ y ε son independientes, ya que $\hat{y}$ se calcula de los datos del ejemplo 7.16, mientras que ε se aplica a una soldadura que no es parte de ese conjunto de datos. Ello da como resultado que la desviación estándar del error de la predicción (7.41) se aproxime por

$$s_{\text{pred}} = \sqrt{s_{\hat{y}}^2 + s^2}$$

Utilizando la ecuación (7.39) al sustituir $s_{\hat{y}}$ se obtiene

$$s_{\text{pred}} = s \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7.42)$$

Ahora se puede determinar la expresión apropiada para el intervalo de predicción.

Un intervalo de predicción con un nivel de $100(1 - \alpha)\%$ para la cantidad $\beta_0 + \beta_1 x$ está dado por

$$\hat{\beta}_0 + \hat{\beta}_1 x \pm t_{n-2, \alpha/2} \cdot s_{\text{pred}} \quad (7.43)$$

donde $s_{\text{pred}} = s \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$.

Observe que el intervalo de predicción es más amplio que el de confianza, porque el valor 1 se suma a la cantidad bajo la raíz cuadrada para considerar la incertidumbre adicional.

Ejemplo 7.18

Para los datos de la soldadura del ejemplo 7.16, determine un intervalo de predicción de 95% para la fuerza de una soldadura específica cuyo contenido de oxígeno es 1.7 partes por mil.

Solución

La fuerza pronosticada es $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(1.7)$, que se calculó en el ejemplo 7.16 de 78.5485.

Utilizando las cantidades que se presentan en el ejemplo 7.16, se calculó el valor de s_{pred} de

$$s_{\text{pred}} = 5.84090 \sqrt{1 + \frac{1}{29} + \frac{(1.7 - 1.51966)^2}{1.33770}} = 6.0102$$

Hay $n - 2 = 29 - 2 = 27$ grados de libertad. El valor de t es, por tanto, $t_{27, .025} = 2.052$. Por tanto, el intervalo de predicción de 95% es

$$78.5485 \pm (2.052)(6.0102) = 78.5485 \pm 12.3329 = (66.22, 90.88)$$

Tanto los intervalos de confianza como los intervalos de predicción que se han descrito corresponden a un valor específico de x . En los ejemplos 7.16 y 7.18 se tomó $x = 1.7$. Al calcular los intervalos para muchos valores de x y conectar los puntos con una curva suave, se obtienen las **bandas de confianza** o **bandas de predicción**, respectivamente. La figura 7.13 muestra las bandas de confianza de 95% y bandas de predicción para los datos de la soldadura presentados en el ejemplo 7.16. Para cualquier contenido de oxígeno dado, las bandas de confianza de 95% y las de predicción se pueden leer de la figura.

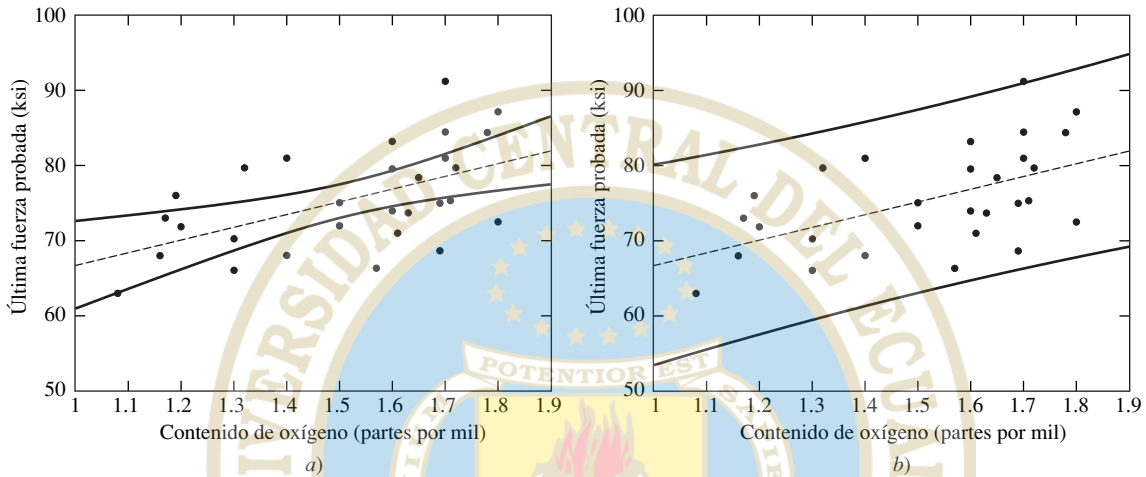


FIGURA 7.13 Contenido de oxígeno contra la última fuerza probada para 29 soldaduras. En ambas figuras, la recta punteada representa la recta de mínimos cuadrados. *a)* Las dos curvas sólidas son las bandas de confianza de 95%. Dado cualquier contenido de oxígeno, se tiene una confianza de 95% de que la media de la fuerza para soldaduras con ese contenido de oxígeno se encuentre entre los límites de confianza superior e inferior. *b)* Las dos curvas sólidas son las bandas de predicción de 95%. Dada cualquier soldadura específica, se tiene una confianza de 95% de que la fuerza para esa soldadura específica se encuentre entre los límites de predicción superior e inferior correspondiendo al contenido de oxígeno de esa soldadura.

Las bandas de confianza y de predicción proporcionan una presentación visual agradable de la forma en que la incertidumbre depende del valor de la variable independiente. Observe que tanto el intervalo de confianza como el de predicción son más angostos cuando $x = \bar{x}$, y aumentan el ancho conforme x se aleja de $\bar{x}$. Esto último se debe al término $(x - \bar{x})^2$ que aparece en el numerador en las expresiones para s_y y s_{pred} . Se concluye que las predicciones basadas en la recta de mínimos cuadrados son más precisas cerca del centro del diagrama de dispersión y lo son menos cerca de los extremos.

Observe que las bandas de confianza indican intervalos de confianza para valores individuales de x . No proporcionan una región de confianza para la recta verdadera $y = \beta_0 + \beta_1 x$. En otras palabras, no se puede tener una confianza de 95% de que la recta verdadera se encuentre en las bandas de confianza de 95%.

Interpretación de los resultados de computadora

En la actualidad, los cálculos de mínimos cuadrados usualmente se hacen en una computadora empleando un software como MINITAB. El siguiente resultado de MINITAB corresponde a los datos de la ley de Hooke.

Regression Analysis: Length versus Weight

The regression equation is

$$\text{Length} = 5.00 + 0.205 \text{ Weight} \quad (1)$$

Predictor	Coef (2)	SE Coef (3)	T (4)	P (5)
Constant	4.99971	0.02477	201.81	0.000
Weight	0.20462	0.01115	18.36	0.000

S = 0.05749 (6) R-Sq = 94.9% (7) R-Sq(adj) = 94.6%

Analysis of Variance (8)

Source	DF	SS	MS	F	P
Regression	1	1.1138	1.1138	337.02	0.000
Residual Error	18	0.0595	0.0033		
Total	19	1.1733			

Unusual Observations (9)

Obs	Weight	Length	Fit	SE Fit	Residual	St Resid
12	2.20	5.5700	5.4499	0.0133	0.1201	2.15R

R denotes an observation with a large standardized residual

Predicted Values for New Observations (10)

New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	5.2453	0.0150	(5.2137, 5.2769)	(5.1204, 5.3701)

Values of Predictors for New Observations (11)

New Obs	Weight
1	1.20

Ahora se explicarán las cantidades etiquetadas en el resultado:

- (1) Es la ecuación de la recta de mínimos cuadrados.
- (2) **Coef:** los coeficientes $\hat{\beta}_0 = 4.99971$ y $\hat{\beta}_1 = 0.20462$.
- (3) **SE Coef:** las desviaciones estándar s_{β_0} y s_{β_1} . (“SE” significa el error estándar, otro término por la desviación estándar.)
- (4) **T:** los valores de los estadísticos t de Student para probar las hipótesis $\beta_0 = 0$ y $\beta_1 = 0$. El estadístico t es igual al coeficiente dividido entre su desviación estándar.
- (5) **P:** Los P -valores para las pruebas de las hipótesis $\beta_0 = 0$ y $\beta_1 = 0$. El P -valor más importante es el de β_1 . Si este P -valor no es lo suficientemente pequeño para rechazar la hipótesis de que $\beta_1 = 0$, el modelo lineal no es útil para predecir y a partir de x . En este ejemplo, los P -valores son extremadamente pequeños, indicando que β_0 ni β_1 son iguales a 0.

- (6) **S:** la estimación s de la desviación estándar del error.
- (7) **R-Sq:** éste es r^2 , el cuadrado del coeficiente de correlación r , también llamado coeficiente de determinación.
- (8) **Análisis de varianza:** esta tabla no es tan importante en la regresión lineal simple cuando sólo hay una variable independiente. Es más importante en la regresión múltiple, donde hay algunas variables independientes. Sin embargo, es notable que los tres números de la columna etiquetada con “SS” son la suma de cuadrados de regresión $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$, la suma de cuadrados del error $\sum_{i=1}^n (y_i - \hat{y}_i)^2$, y su suma, la suma total de cuadrados $\sum_{i=1}^n (y_i - \bar{y})^2$.
- (9) **Observaciones inusuales:** aquí MINITAB trata de alertarlo acerca de los puntos de datos que pueden incumplir algunos de los supuestos del 1 al 4 analizados. MINITAB es conservador y con frecuencia lista varios de esos puntos, incluso cuando los datos están bien descritos con un modelo lineal. En la sección 7.4 se aprenderán algunos métodos gráficos para comprobar los supuestos del modelo lineal.
- (10) **Valores predichos para las nuevas observaciones:** éstos son los intervalos de confianza y los de predicción para valores de x especificados por el usuario. Aquí se especifica $x = 1.2$ para el peso. El “Fit” es el valor ajustado $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ y “SE Fit” es la desviación estándar $s_{\hat{y}}$. Después se presentan los intervalos de confianza y de predicción del 95%, respectivamente.
- (11) **Valores de los indicadores para las nuevas observaciones:** ésta es sólo una lista de los valores de x para los cuales se han calculado los intervalos de confianza y de predicción. Éste muestra que estos intervalos se refieren a un peso de $x = 1.2$.

Ejercicios para la sección 7.3

1. Se está evaluando un nuevo resorte. Se han colgado veinticinco pesas x_i , y se mide para cada una la longitud del resorte y_i . Los pesos se miden en libras, y la longitud, en pulgadas. Se registró el siguiente resumen estadístico.

$$\bar{x} = 2.40 \quad \bar{y} = 12.18 \quad \sum_{i=1}^{25} (x_i - \bar{x})^2 = 52.00$$

$$\sum_{i=1}^{25} (y_i - \bar{y})^2 = 498.96 \quad \sum_{i=1}^{25} (x_i - \bar{x})(y_i - \bar{y}) = 160.27$$

Sea β_0 la longitud del resorte en reposo, y β_1 el aumento en la longitud ocasionado por una carga de 1 lb. Suponga que los supuestos del 1 al 4 de la p. 508 son válidos.

- Calcule los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$.
- Calcule la estimación de la varianza del error s^2 .
- Determine los intervalos de confianza del 95% para β_0 y β_1 .
- El fabricante del resorte afirma que la constante del resorte no es mayor de 3 pulg/lb. ¿Los datos proporcionan evidencias suficientes para rechazar esta afirmación?
- También se afirma que la longitud sin carga del resorte es de al menos 5.5 pulg. ¿Los datos proporcionan pruebas suficientes para rechazar esta afirmación?
- Determine un intervalo de confianza de 99% para la longitud del resorte bajo una carga de 1.5 lb.

- g) Determine un intervalo de predicción de 99% para la longitud medida del resorte bajo una carga de 1.5 lb.
- h)Cuál es más útil en este caso, ¿el intervalo de confianza o el intervalo de predicción? Explique.
2. Para muchos compuestos químicos, la cantidad que se disolverá en un volumen de agua específico depende de la temperatura. El siguiente resultado de MINITAB describe el ajuste de un modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$ que expresa el número de gramos de cierto compuesto químico disueltos por litro de agua (y) con respecto a la temperatura en °C de (x). Hay $n = 6$ observaciones.

Predictor	Coef	SE Coef	T	P
Constant	1.4381	0.62459	2.30	0.083
Temperature	0.30714	0.02063	14.9	0.000

- a) Hay $n = 6$ observaciones. ¿Cuántos grados de libertad hay para los estadísticos t de Student?
- b) Determine un intervalo de confianza de 95% para β_1 .
- c) Determine un intervalo de confianza de 95% para β_0 .
- d) Alguien afirma que si la temperatura del agua se aumentara en 1°C, la media del número de gramos disueltos aumentará en exactamente 0.40. Utilice el resultado para realizar una prueba de hipótesis con el fin de determinar si esta afirmación puede ser rechazada.
- e) Alguien afirma que la media del número de gramos que se puede disolver en un litro de agua a 0°C es menor a 1.0. Utilice el resultado dado para realizar una prueba de hipótesis con el propósito de determinar si esta afirmación puede ser rechazada.
3. Los óxidos de nitrógeno (NO_x) constituyen un componente muy importante de la contaminación atmosférica producida por vehículos automotores. En un estudio para determinar la relación entre la carga sobre un motor y la producción de NO_x , un vehículo fue manipulado en un laboratorio de prueba a diferentes velocidades. Se realizaron mediciones periódicas de los caballos de fuerza (x) y de las emisiones de NO_x (y). Las emisiones de NO_x se midieron en mg/s. El siguiente resultado de MINITAB describe el ajuste de un modelo lineal con estos datos. Suponga que los supuestos del 1 al 4 de la p. 508 son válidos.

The regression equation is

$$\text{Chassis NO}_x = 44.5 + 0.845 \text{ Chassis HP}$$

Predictor	Coef	SE Coef	T	P
Constant	44.534	2.704	16.47	0.000
Chassis HP	0.84451	0.03267	25.85	0.000

S = 24.62

R-Sq = 84.7%

R-Sq(adj) = 84.5%

Predicted Values for New Observations

New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	78.31	2.23	(73.89, 82.73)	(29.37, 127.26)

Values of Predictors for New Observations

New Obs	Chassis HP
1	40.0

- a) ¿Cuáles son la pendiente y el intercepto de la recta de mínimos cuadrados?
- b) ¿El modelo lineal es útil para predecir las tasas de emisión de NO_x a partir de los caballos de fuerza? Explique.
- c) Pronostique la tasa de emisión de NO_x si el motor se detiene a los 10 caballos de fuerza.
- d) ¿Cuál es la correlación entre caballos de fuerza y tasa de emisiones de NO_x ?
- e) El resultado proporciona un intervalo de confianza de 95% para la media de la tasa de emisiones de NO_x a 40 caballos de fuerza. Hay $n = 123$ observaciones en el conjunto de datos. Utilizando el valor “SE Fit”, determine un intervalo de confianza de 90%.

- f) Alguien planea operar el motor a 40 caballos de fuerza, y predice que la tasa de emisiones de NO_x será de 180 mg/s. ¿Esta es una predicción razonable? Si es así, explique por qué. Si no, dé un rango razonable de valores predichos.
4. En un experimento similar al del ejercicio 3, el motor se retira del chasis y se acciona. Se miden los caballos de fuerza y la tasa de emisión de NO_x . A continuación se muestra el resultado de MINITAB. Suponga que los supuestos del 1 al 4 de la p. 508 son válidos.

The regression equation is

$$\text{Engine NO}_x = 33.0 + 0.753 \text{ Engine HP}$$

Predictor	Coef	SE Coef	T	P
Constant	33.042	2.371	13.94	0.000
Engine HP	0.75269	0.02912	25.85	0.000
S = 22.79 R-Sq = 84.7% R-Sq(adj) = 84.5%				

- a) ¿Cuál es la pendiente de la recta de mínimos cuadrados?
- b) Hay 123 puntos en el conjunto de datos. Determine un intervalo de confianza de 95% para la pendiente.
- c) Realice una prueba con la hipótesis nula de que la pendiente es mayor que o igual a 0.8. ¿Cuál es el P -valor?
5. Con referencia a los ejercicios 3 y 4, el ingeniero nota que la pendiente de la recta de mínimos cuadrados en el experimento descrito en el ejercicio 4 es menor que en el experimento descrito en el 3. Quiere probar la hipótesis de que el efecto de los caballos de fuerza en la tasa de emisiones de NO_x depende de si el motor está dentro o fuera del chasis. Sea β_C el aumento en la tasa de emisiones debido a un aumento de 1 hp con el motor dentro del chasis, y que β_E denote el aumento correspondiente con el motor fuera del chasis.
- a) Exprese la hipótesis nula que debe probarse en función de β_C y β_E .
- b) Sean $\hat{\beta}_C$ y $\hat{\beta}_E$ las pendientes de las rectas de mínimos cuadrados. Suponga que estas pendientes son independientes. Hay 123 observaciones en cada conjunto de datos. Pruebe la hipótesis nula en el inciso a). ¿Puede concluir que el efecto de los caballos de fuerza es diferente en los dos casos?
6. El artículo "Withdrawal Strength of Threaded Nails" (D. Rammer, S. Winistorfer y D. Bender, en *Journal of Structural Engineering* 2001:442–449) describe un experimento para investigar la relación entre el diámetro de un clavo (x) y su fuerza retirada final (y). Se colocaron clavos de forma anular enhebrados en madera de abeto de Douglas, y después se midieron sus fuerzas de retirada en N/mm. Se obtuvieron los resultados siguientes para diez diámetros diferentes (en mm).

x	2.52	2.87	3.05	3.43	3.68	3.76	3.76	4.50	4.50	5.26
y	54.74	59.01	72.92	50.85	54.99	60.56	69.08	77.03	69.97	90.70

- a) Calcule la recta de mínimos cuadrados para predecir la fuerza a partir del diámetro.
- b) Calcule la estimación s de la desviación estándar del error.
- c) Calcule un intervalo de confianza de 95% para la pendiente.
- d) Determine un intervalo de confianza de 95% para la media de la fuerza de retirada de clavos de 4 mm de diámetro.
- e) ¿Puede concluir que la media de la fuerza de retirada de clavos de 4 mm de diámetro es mayor que 60 N/mm? Realice una prueba de hipótesis y reporte el P -valor.
- f) Determine un intervalo de predicción de 95% para la fuerza de retirada de un clavo especial cuyo diámetro es de 4 mm.
7. En un estudio para determinar la relación entre la permeabilidad (en cm/m) y la resistencia eléctrica (en $\text{k}\Omega \cdot \text{cm}^2$) de piel humana se reunieron los datos de la tabla siguiente. Suponga que los supuestos del 1 al 4 de la p. 508 son válidos.

Perm.	Res.	Perm.	Res.	Perm.	Res.	Perm.	Res.
1.39	0.90	1.08	1.35	1.53	0.94	1.57	0.72
1.21	1.21	1.79	0.40	1.79	0.41	1.49	0.66
1.58	0.65	1.54	0.80	1.45	0.82	1.68	0.67
1.35	0.97	2.04	0.17	1.46	0.83	1.52	0.79
1.76	0.40	1.51	0.94	1.47	0.79	1.29	0.96
1.57	0.65	1.16	1.13	1.67	0.59	1.90	0.27
1.32	0.95	1.71	0.57	1.81	0.47	1.10	1.28
1.13	1.33	1.64	0.65	1.40	0.90	1.67	0.57
1.70	0.44	1.67	0.62	1.82	0.37	1.66	0.54
1.79	0.37	1.19	1.22	1.49	0.78	1.46	0.82
1.09	1.34	1.56	0.78	1.23	1.01	1.85	0.40
1.29	0.99	1.63	0.64				

- Calcule la recta de mínimos cuadrados para pronosticar la permeabilidad (y) a partir de la resistencia (x).
 - Calcule los intervalos de confianza de 95% para β_0 y β_1 .
 - Pronostique la permeabilidad de la piel cuya resistencia es 1.7 k Ω .
 - Determine un intervalo de confianza de 95% para la media de la permeabilidad para todas las muestras de piel con una resistencia de 1.7 k Ω .
 - Determine un intervalo de confianza de 95% para la permeabilidad de la piel cuya resistencia es 1.7 k Ω .
8. Tres ingenieros están calculando de manera independiente la constante de un resorte usando el modelo lineal especificado por la ley de Hooke. El ingeniero A mide la longitud del resorte bajo cargas de 0, 1, 3, 4 y 6 lb para un total de cinco mediciones. El ingeniero B utiliza las mismas cargas, pero repite el experimento dos veces para un total de diez mediciones independientes. El ingeniero C usa cargas de 0, 2, 6, 8 y 12 lb midiendo una vez para cada carga. Todos los ingenieros usan los mismos aparatos de medición y el mismo procedimiento. Cada ingeniero calcula un intervalo de confianza de 95% para la constante del resorte.
- Si el ancho del intervalo del ingeniero A se divide entre el ancho del intervalo del ingeniero B, el cociente será aproximadamente_____.
 - Si el ancho del intervalo del ingeniero A se divide entre el ancho del intervalo del ingeniero C, el cociente será aproximadamente_____.
 - Cada ingeniero calcula un intervalo de confianza de 95% para la longitud del resorte bajo una carga de 2.5 libras. ¿Cuál intervalo es probable que sea el más corto? ¿Cuál será el más largo?
9. En los datos de las soldaduras (ejemplo 7.16), imagine que los intervalos de confianza de 95% se calculan para la media de la fuerza de las soldaduras con contenido de oxígeno de 1.3, 1.5 y 1.8 partes por mil. ¿Cuál de los intervalos de confianza sería el más corto? ¿Cuál sería el más largo?
10. Con referencia al ejercicio 1, si se construyen intervalos de confianza de 95% para la longitud del resorte con cargas de 2.15, 2.57 y 2.45 libras, ¿qué intervalo de confianza sería el más corto? ¿Cuál sería el más largo?
11. En un estudio de barras de cobre, la relación entre la tensión de corte en ksi (x) y la de corte en % (y) se resumió por la recta de mínimos cuadrados $y = -20.00 + 2.56x$. Había un total de $n = 17$ observaciones, y el coeficiente de la determinación era $r^2 = 0.9111$. Si la suma total de cuadrados fuera $\sum (y_i - \bar{y})^2 = 234.19$, calcule la estimación de la varianza del error s^2 .
12. En la fabricación de fibras sintéticas con frecuencia es “puesta” a temperaturas altas. El propósito es mejorar las propiedades de encogimiento de la fibra. En una prueba de 25 elementos de hilo, la relación entre la temperatura en $^{\circ}\text{C}$ (x) y el encogimiento en % (y) se resumió mediante la recta de mínimos cuadrados $y = -12.789 + 0.133x$. La suma total de cuadrados fue

$\sum (y_i - \bar{y})^2 = 57.313$, y la estimación de la varianza del error aproximadamente fue $s^2 = 0.0670$. Calcule el coeficiente de la determinación r^2 .

13. En el siguiente resultado de MINITAB, algunos de los números han sido accidentalmente borrados. Calcúlelos de nuevo empleando los números aún disponibles. Hay $n = 25$ puntos en el conjunto de datos.

The regression equation is

$$Y = 1.71 + 4.27 X$$

Predictor	Coef	SE Coef	T	P
Constant	1.71348	6.69327	(a)	(b)
X	4.27473	(c)	3.768	(d)

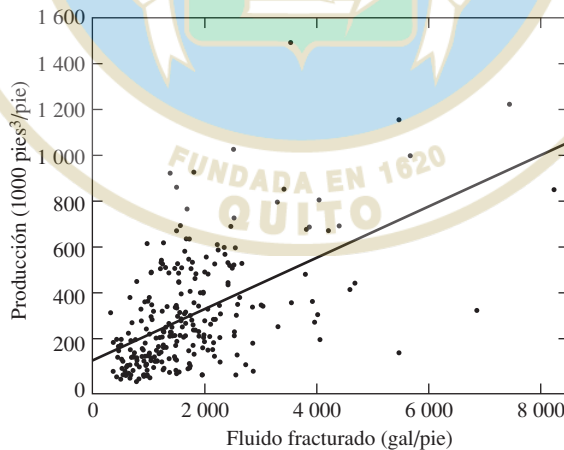
S = 0.05749 R-Sq = 38.2%

14. En el siguiente resultado de MINITAB, algunos de los números han sido accidentalmente borrados. Calcúlelos otra vez aprovechando los números aún disponibles. Hay $n = 20$ puntos en el conjunto de datos.

Predictor	Coef	SE Coef	T	P
Constant	(a)	0.43309	0.688	(b)
X	0.18917	0.065729	(c)	(d)

S = 0.67580 R-Sq = 31.0%

15. Con la finalidad de aumentar la producción de pozos de gas, con frecuencia se usa un procedimiento conocido como “tratamiento de fractura”. Fracturar el fluido consiste en mezclarlo con arena y bombearla dentro del pozo. La siguiente figura presenta un diagrama de dispersión de la producción mensual contra el volumen del fluido fracturado bombeado para 255 pozos de gas. Tanto la producción como el fluido están expresados en unidades de volumen por pie de profundidad del pozo. La recta de mínimos cuadrados está sobrepuesta. La ecuación de la recta de mínimos cuadrados es $y = 106.11 + 0.1119x$.



- a) De la recta de mínimos cuadrados, calcule la producción para un pozo en el cual se han bombeado 4 000 gal/pie.
- b) De la recta de mínimos cuadrados, calcule la producción para un pozo en el cual se han bombeado 500 gal/pie.
- c) Se ha excavado un nuevo pozo y se han bombeado 500 gal/pie. Con base en el diagrama de dispersión, ¿es más probable que la producción de este pozo se encontrará arriba o debajo de lo estimado con mínimos cuadrados?
- d) ¿Qué característica del diagrama de dispersión indica que no se cumple el supuesto 3 de la pág. 508?

7.4 Comprobación de supuestos y transformación de datos

Los métodos que se han analizado hasta ahora son válidos bajo el supuesto de que la relación entre las variables x y y satisfacen el modelo lineal $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, donde los errores ε_i satisfacen los supuestos 1 al 4. Estos supuestos se repiten aquí.

Supuestos para los errores en los modelos lineales

1. Los errores $\varepsilon_1, \dots, \varepsilon_n$ son aleatorios e independientes. En particular, la magnitud de cualquier error ε_i no influye en el valor del siguiente error ε_{i+1} .
2. Todos los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen media 0.
3. Todos los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen la misma varianza, que se denota mediante σ^2 .
4. Los errores $\varepsilon_1, \dots, \varepsilon_n$ están distribuidos normalmente.

Como ya se mencionó, el supuesto de normalidad (4) es menos importante cuando el tamaño de muestra es grande. Mientras que los incumplimientos moderados del supuesto de varianza constante (3) no importan demasiado, los graves se deben corregir.

Se necesita algo para comprobar estos supuestos y asegurarse de que los métodos son adecuados. Han existido innumerables herramientas diagnósticas propuestas para este propósito. Se han escrito muchos libros sobre el tema. Aquí se restringirá a algunos de los procedimientos más básicos.

La gráfica de residuos contra valores ajustados

El mejor diagnóstico para la regresión de mínimos cuadrados es una gráfica de residuos e_i contra valores ajustados $\hat{y}_i$, en ocasiones llamada **gráfica de residuos**. La figura 7.14 de la página 528 presenta esta gráfica para la estatura de Galton contra los datos de antebrazo (véase la figura 7.1 de la sección 7.1 para los datos originales). Por necesidad matemática, los residuos tienen media 0, y también la correlación entre los residuos y los valores ajustados es 0. Por tanto, la recta de mínimos cuadrados es horizontal, pasando a través del 0 sobre el eje vertical. Cuando el modelo lineal es válido, y se satisfacen los supuestos del 1 al 4, la gráfica no indicará un patrón importante. No hay curva en la gráfica, y la dispersión vertical de los puntos no debe variar demasiado de la dispersión horizontal de la gráfica, excepto quizás cerca de los bordes. Estas condiciones están bien satisfechas para los datos de Galton. No existe razón para dudar de los supuestos del modelo lineal.

Un poco de terminología: cuando la dispersión vertical en un diagrama de dispersión no varía demasiado, se dice que el diagrama de dispersión es **homoscedástico**. Lo contrario de homoscedástico es **heteroscedástico**.

Una gráfica de residuos que se ve bien no prueba por sí misma que el modelo lineal es adecuado, porque los supuestos del modelo lineal pueden fallar de otras maneras. Por otra parte, una gráfica de residuos con un defecto serio indica claramente que el modelo lineal es inadecuado.

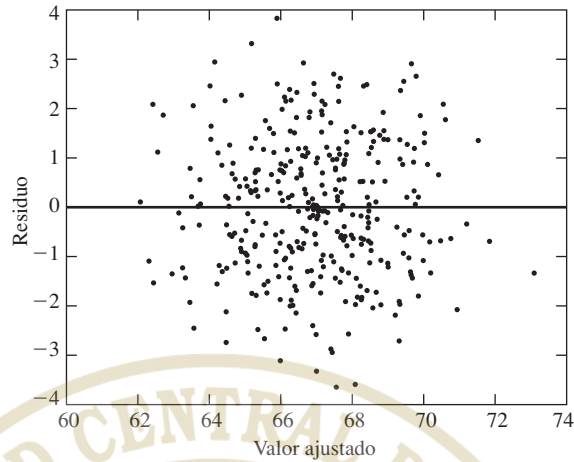


FIGURA 7.14 Gráfica de residuos (e_i) contra valores ajustados ($\hat{y}_i$) para la estatura de Galton contra los datos del antebrazo. No hay ningún patrón importante en la gráfica, y la extensión vertical no varía demasiado, excepto quizás cerca de los bordes. Esto es consistente con los supuestos del modelo lineal.

Resumen

Si la gráfica de residuos contra valores ajustados

- No muestra una tendencia importante de la curva, y
- Es **homoscedástica**; es decir, la dispersión vertical no varía demasiado a lo largo de la longitud horizontal de la gráfica, excepto quizás cerca de los bordes,

entonces es *probable*, pero no se tiene la *seguridad*, de que los supuestos del modelo lineal sean válidos.

Sin embargo, si la gráfica de residuos *sí* muestra tendencia importante o se curva, o es **heteroscedástica**, se tiene la *seguridad* de que los supuestos del modelo lineal *no* son válidos.

En muchos casos, la gráfica de los residuos presentará curvatura o heteroscedasticidad, ello indica el incumplimiento de los supuestos. Se presentarán tres ejemplos. Después se presentará un método llamado **transformación de variables**, que algunas veces puede arreglar los incumplimientos de los supuestos y permite usar el modelo lineal.

Ejemplo

7.19

La figura 7.15 presenta una gráfica de las concentraciones de ozono atmosféricas contra concentraciones de NO_x medidas en 359 días en un año reciente cerca de Riverside, California. (NO_x significa óxido de nitrógeno, y se refiere a la suma de NO y NO_2 .) Ambas concentraciones están medidas en partes por miles de millones (ppmm). Ahora esta gráfica es una gráfica de residuos. La gráfica es evidentemente **heteroscedástica**; es decir, la dispersión vertical va-

ría de manera considerable con el valor ajustado. Específicamente, cuando éste (la estimación de la concentración de ozono) es grande, el residuo tiende a estar más lejos de 0. En virtud de que la magnitud de la dispersión en los residuos depende de la varianza del error σ^2 , se concluye que la varianza del error es grande en días donde el valor ajustado también lo es. Ésta es una violación del supuesto 3, que establece que la varianza σ^2 es igual para todas las observaciones. La gráfica también contiene un dato atípico (donde la concentración de ozono está cerca de 100). La gráfica de residuos indica que no se debe usar este modelo lineal para predecir la concentración de ozono de la concentración de NO_x .

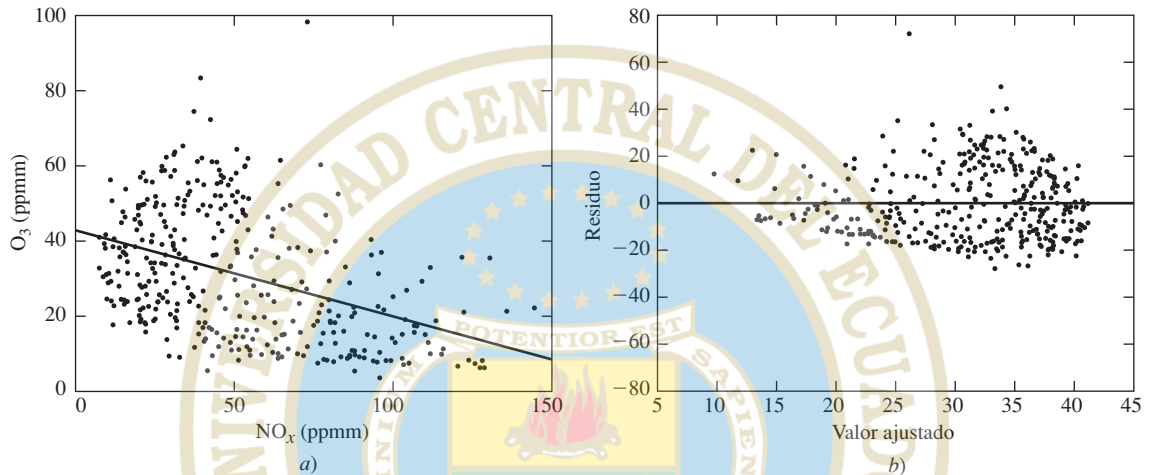


FIGURA 7.15 a) Gráfica de la concentración de ozono contra la concentración de NO_x . La recta de mínimos cuadrados está sobrepuesta. b) Gráfica de residuos (e_i) contra valores ajustados ($\hat{y}_i$) para estos datos. La dispersión vertical aumenta evidentemente el valor ajustado. Esto último indica un incumplimiento del supuesto de la varianza del error constante.

Ejemplo

7.20

(Basado en el artículo “Advances in Oxygen Equivalence Equations for Predicting the Properties of Titanium Welds,” D. Harwig, W. Ittiwattana y H. Castner, en *The Welding Journal*, 2001: 126s-136s.) Las propiedades físicas de una soldadura están influenciadas por la composición química del material de la soldadura. Una medida de la composición química es el número de Ogden-Jaffe, que es una suma ponderada de los porcentajes de carbono, oxígeno y nitrógeno en la soldadura. En un estudio de 63 soldaduras, la dureza de la soldadura (medida en la escala de Rockwell B) se graficó contra el número de Ogden-Jaffe. En la figura 7.16 de la página 530 se presenta la gráfica junto con otra de residuos. Esta última muestra un patrón con residuos positivos de la concentración en la parte central de la gráfica, y de residuos negativos en los extremos. Técnicamente, lo anterior indica que no todos los errores ε_i tienen una media de 0. Esto por lo general ocurre por cualquiera de dos razones: ya sea que la relación entre las variables no sea lineal, o que hay otras variables que se necesitan incluir en el modelo. Se concluye que no se debe usar este modelo para predecir la dureza de la soldadura a partir del número de Ogden-Jaffe.

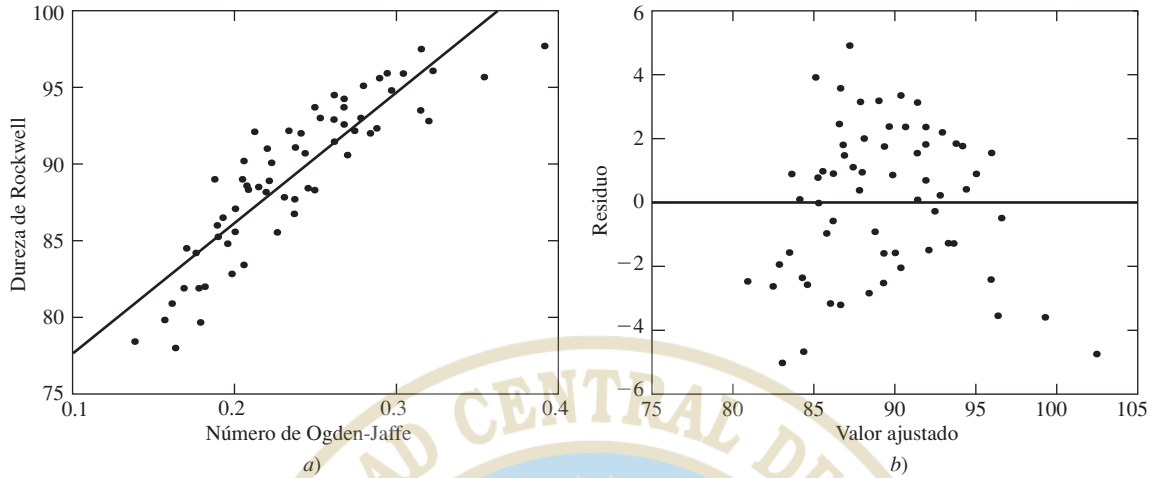


FIGURA 7.16 *a)* Gráfica de la dureza de Rockwell (B) contra el número de Ogden-Jaffe. La recta de mínimos cuadrados está sobrepuesta. *b)* Gráfica de residuos (e_i) contra valores ajustados ($\hat{y}_i$) para estos datos. La gráfica de residuos muestra una tendencia, con residuos positivos en el centro y negativos en los extremos.

Ejemplo 7.21

Estos datos se presentaron en el ejercicio 15 en la sección 7.3. Para un grupo de 255 pozos de gas, la producción mensual por pie de profundidad del pozo está graficada contra el volumen de fluido fracturado bombeado en el pozo. Esta gráfica, junto con la de residuos, se presenta en la figura 7.17. Esta última es fuertemente heteroscedástica, lo que indica que la varianza del error es mayor para pozos de gas cuya estimación de la producción es mayor. Por supues-

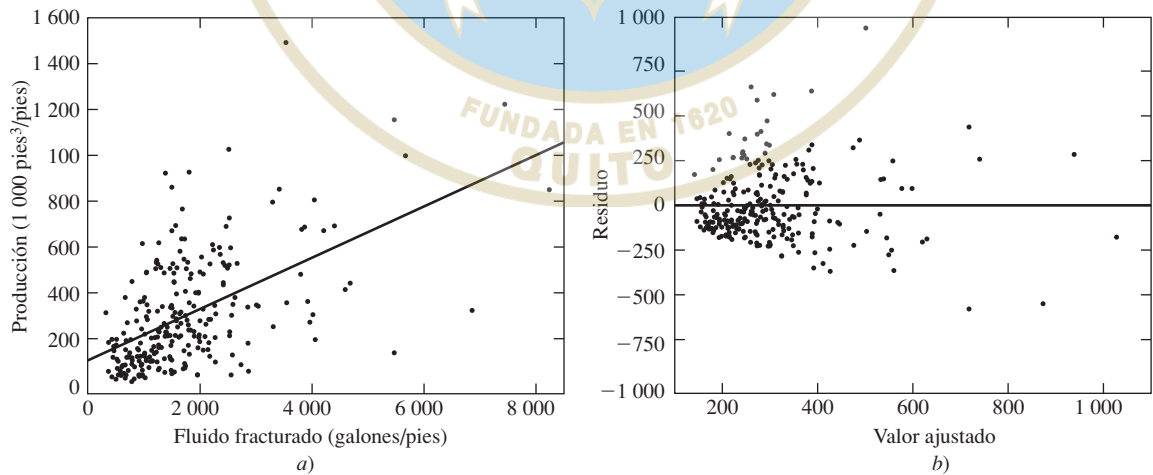


FIGURA 7.17 *a)* Gráfica de la producción mensual contra volumen de fluido fracturado para 255 pozos de gas. *b)* Gráfica de residuos (e_i) contra valores ajustados ($\hat{y}_i$) para los datos de los pozos de gas. La dispersión vertical aumenta con el valor ajustado. Esto último indica incumplimiento del supuesto de la varianza del error constante.

to, éstos son los pozos dentro de los cuales se ha bombeado mayor cantidad de fluido fracturado. Se concluye que no se debe utilizar este modelo para predecir la producción de los pozos a partir de la cantidad de fluido fracturado bombeado.

Transformación de variables

Si se ajusta el modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$ y se determina que la gráfica de residuos es heteroscedástica, o presenta una tendencia o patrón, a veces se puede arreglar el problema elevando x , y , o ambos a una potencia. Podría ser el caso de que un modelo de la forma $y^a = \beta_0 + \beta_1 x^b + \varepsilon$ ajuste los datos bien. En general, sustituir una variable con una función de la misma se llama **transformación** de la variable. Específicamente, elevar una variable a una potencia se llama **transformación potencia**. Tomar el logaritmo de una variable también se considera como transformación potencia, aunque el logaritmo no es una potencia.

He aquí un ejemplo simple que muestra cómo funciona una transformación de potencia. La tabla siguiente presenta valores para las variables hipotéticas x , y y y^2 .

x	y	y^2	x	y	y^2
1.0	2.2	4.84	11.0	31.5	992.25
2.0	9.0	81.00	12.0	32.7	1 069.29
3.0	13.5	182.25	13.0	34.9	1 218.01
4.0	17.0	289.00	14.0	36.3	1 317.69
5.0	20.5	420.25	15.0	37.7	1 421.29
6.0	23.3	542.89	16.0	38.7	1 497.69
7.0	25.2	635.04	17.0	40.0	1 600.00
8.0	26.4	696.96	18.0	41.3	1 705.69
9.0	27.6	761.76	19.0	42.5	1 806.25
10.0	30.2	912.04	20.0	43.7	1 909.69

El diagrama de dispersión de y contra x se presenta en la figura 7.18, junto con la gráfica de residuos. Evidentemente el modelo lineal no es adecuado.

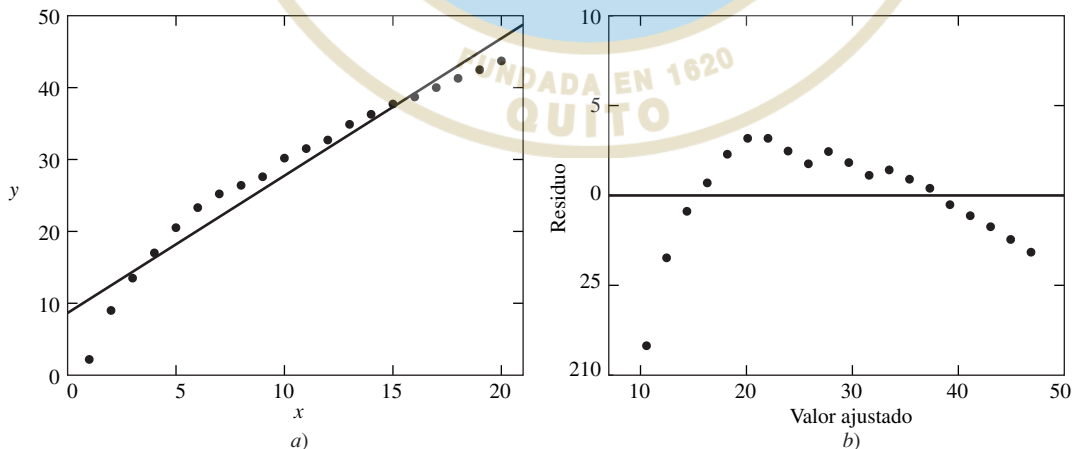


FIGURA 7.18 a) Gráfica de y contra x con la recta de mínimos cuadrados sobrepuesta. b) Gráfica de residuos contra valores ajustados. Hay un fuerte patrón en la gráfica de residuos, lo que indica que el modelo lineal no es adecuado.

El modelo $y = \beta_0 + \beta_1 x + \varepsilon$ no ajusta los datos. Sin embargo, se puede solucionar este problema usando y^2 en lugar de y . La figura 7.19 presenta el diagrama de dispersión de y^2 contra x , junto con la gráfica de residuos. La gráfica de residuos es homoscedástica, sin tendencia o patrón perceptible.

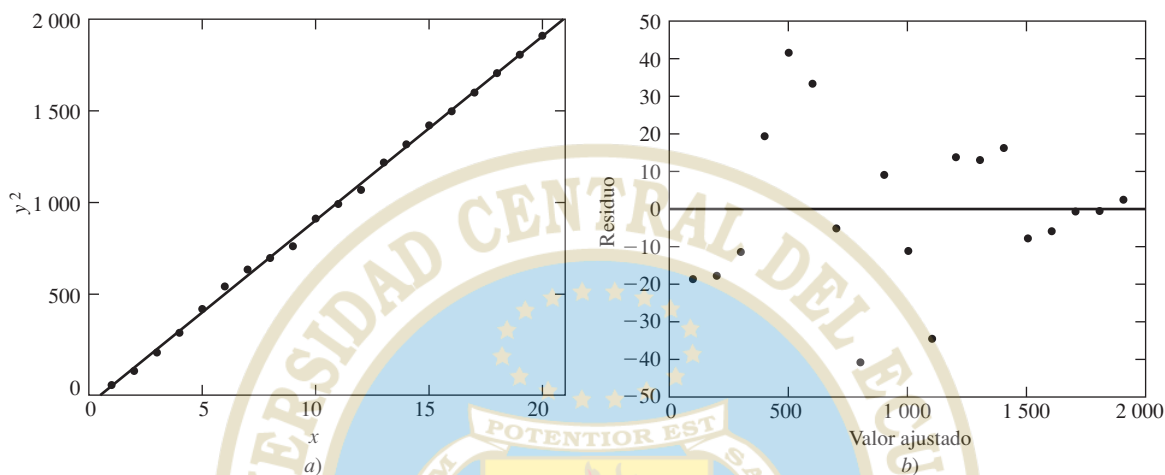


FIGURA 7.19 a) Gráfica de y^2 contra x con la recta de mínimos cuadrados sobrepuesta. b) Gráfica de residuos contra valores ajustados. La gráfica de residuos es casi homoscedástica, sin tendencia o patrón perceptible.

Se concluye que el modelo $y^2 = \beta_0 + \beta_1 x + \varepsilon$ es un modelo creíble para estos datos. En este ejemplo se transforma y , pero no se necesita transformar x . En otros casos, se puede transformar solamente x , o tanto x como y .

Determinación de cuál transformación se debe aplicar

Con la experiencia es posible observar un diagrama de dispersión, o una gráfica de residuos, y hacer un supuesto fundamentado respecto a cómo transformar las variables. También hay métodos matemáticos disponibles para determinar una buena transformación. Sin embargo, es satisfactorio proceder por prueba y error. Intente varias potencias tanto para x como para y (incluyendo $\ln x$ y $\ln y$), observe las gráficas de residuos, y espere encontrar una que sea homoscedástica, sin patrón perceptible. Un análisis más avanzado de la selección de la transformación se encuentra en Draper y Smith (1998).

Las transformaciones no siempre funcionan

Es importante recordar que las transformaciones de las potencias no siempre funcionan. A veces ninguna de las gráficas de residuos parece buena, no importa qué transformaciones se prueben. En estos casos, se deben usar otros métodos. Uno de éstos es la regresión múltiple, que se analizará en el capítulo 8. Algunos otros se mencionan brevemente al final de esta sección.

Las gráficas de residuos con pocos puntos son difíciles de interpretar

Cuando solamente hay algunos puntos en una gráfica de residuos, es difícil determinar si se satisfacen los supuestos del modelo lineal. A veces tal gráfica parecerá heteroscedástica o presentar un patrón, pero con una inspección cercana se encontrará que esta impresión visual la causa la colocación de sólo uno o dos puntos. En ocasiones es difícil determinar si tal gráfica contiene un dato atípico. Cuando uno se enfrenta a una gráfica de residuos dispersa difícil de interpretar es razonable ajustar un modelo lineal, pero considerando los resultados con cautela, en el entendido de que no se ha establecido la propiedad del modelo. Siempre y cuando se tengan más datos disponibles, se puede tomar una decisión fundamentada. Por supuesto, no todas las gráficas de residuos dispersas son difíciles de interpretar. Algunas veces hay un patrón obvio, que no se puede cambiar sólo por el corrimiento de uno o dos puntos. En estos casos no se debe usar el modelo lineal.

Ejemplo 7.22

Con referencia al ejemplo 7.19, la figura 7.15 presenta una gráfica de ozono contra concentraciones de NO_x . Se tiene que la transformación del ozono es su logaritmo natural, $\ln$ ozono, produce una gráfica lineal satisfactoria. La figura 7.20 presenta el diagrama de dispersión de $\ln$ ozono contra NO_x , y la gráfica de residuos correspondiente. La gráfica de residuos es homoscedástica, sin patrón perceptible. El dato atípico que estaba presente en los datos originales es menos prominente. El modelo lineal se ve bien.

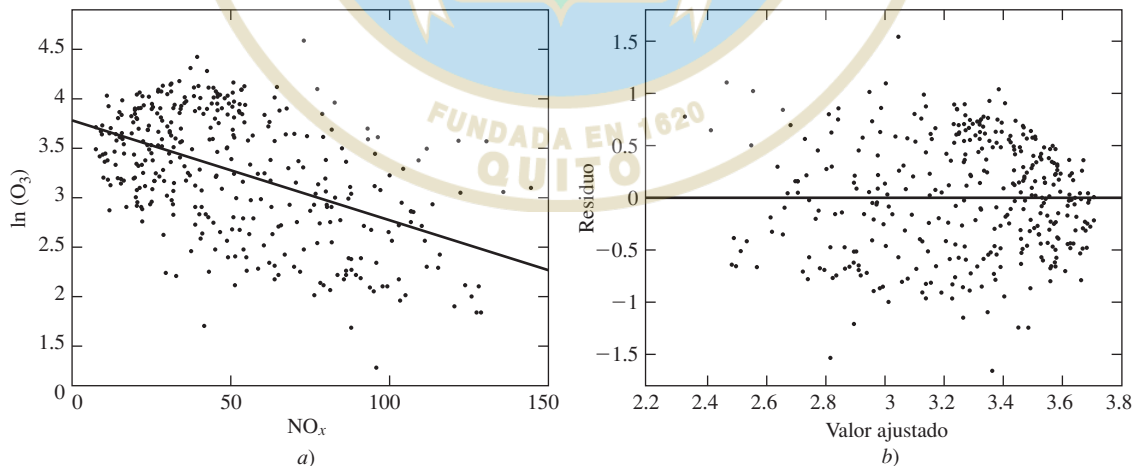


FIGURA 7.20 a) Gráfica del logaritmo natural de la concentración de ozono contra la concentración de NO_x . La recta de mínimos cuadrados está sobrepuesta. b) Gráfica de residuos (e_i) contra valores ajustados ($\hat{y}_i$) para estos datos. El modelo lineal se ve bien.

El siguiente resultado de MINITAB corresponde a los datos transformados.

Regression Analysis: LN OZONE versus NOx					
The regression equation is					
LN OZONE = 3.78 - 0.0101 NOx					
Predictor	Coef	SE Coef	T	P	
Constant	3.78238	0.05682	66.57	0.000	
NOx	-0.0100976	0.0009497	-10.63	0.000	
S = 0.5475		R-Sq = 24.1%		R-Sq(adj) = 23.8%	
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	33.882	33.882	113.05	0.000
Residual Error	357	106.996	0.300		
Total	358	140.878			
Predicted Values for New Observations					
New Obs	Fit	SE Fit	95.0% CI	95.0% PI	
1	3.2775	0.0289	(3.2206, 3.3344)	(2.1994, 4.3556)	
Values of Predictors for New Observations					
New Obs	NOx				
1	50.0				

El análisis de los datos transformados da resultados para el logaritmo natural de la concentración de ozono. Para algunos propósitos se puede regresar a las unidades originales. Por ejemplo, se pueden utilizar los datos transformados con el fin de encontrar intervalos de predicción para valores de ozono dando un valor particular de NO_x . Con este propósito se usan los métodos de la sección 7.3 para encontrar el intervalo para $\ln$ ozono, y después transformar este intervalo de regreso a las unidades originales. El ejemplo 7.23 muestra cómo.

Ejemplo
7.23

Usando el resultado anterior de MINITAB, pronostique el nivel de ozono cuando el nivel de NO_x es 50 ppmm, y determine un intervalo de predicción de 95% para el nivel de ozono en un día cuando el nivel de NO_x es 50 ppmm.

Solución

Sea y el nivel de ozono en un día cuando el nivel de NO_x es 50 ppmm, y $\hat{y}$ el valor pronosticado para ese nivel. Primero se calcula el valor de $\ln \hat{y}$, que es el valor pronosticado para $\ln$ ozono, utilizando los estimadores de los coeficientes del resultado de MINITAB. Para un valor de NO_x de 50, la predicción es

$$\ln \hat{y} = 3.78238 - 0.0100976(50) = 3.2775$$

Por tanto, el valor de ozono pronosticado es

$$\hat{y} = e^{3.2775} = 26.51$$

Se calcula un intervalo de predicción de 95% para el nivel de ozono y , se consulta el intervalo de predicción para $\ln y$ del resultado de MINITAB:

$$2.1994 < \ln y < 4.3556$$

Tomando exponenciales a toda la desigualdad se obtiene el intervalo de predicción de 95% para el nivel de ozono:

$$e^{2.1994} < y < e^{4.3556}$$

$$9.02 < y < 77.91$$

Es importante observar que el método utilizado en el ejemplo 7.23 funciona solamente para el intervalo de predicción. No funciona respecto del intervalo de confianza para la respuesta media. Cuando la variable dependiente se ha transformado, no se puede obtener aquél en las unidades originales.

Ejemplo 7.24

Con referencia del ejemplo 7.20, la figura 7.16 presenta una gráfica de la dureza de Rockwell (B) contra el número de Ogden-Jaffe para un grupo de soldaduras. En este caso, al tomar el recíproco del número de Ogden-Jaffe (elevándolo a la potencia -1) se obtiene una relación casi lineal. La figura 7.21 presenta los resultados. Observe que, en este caso, se transforma la

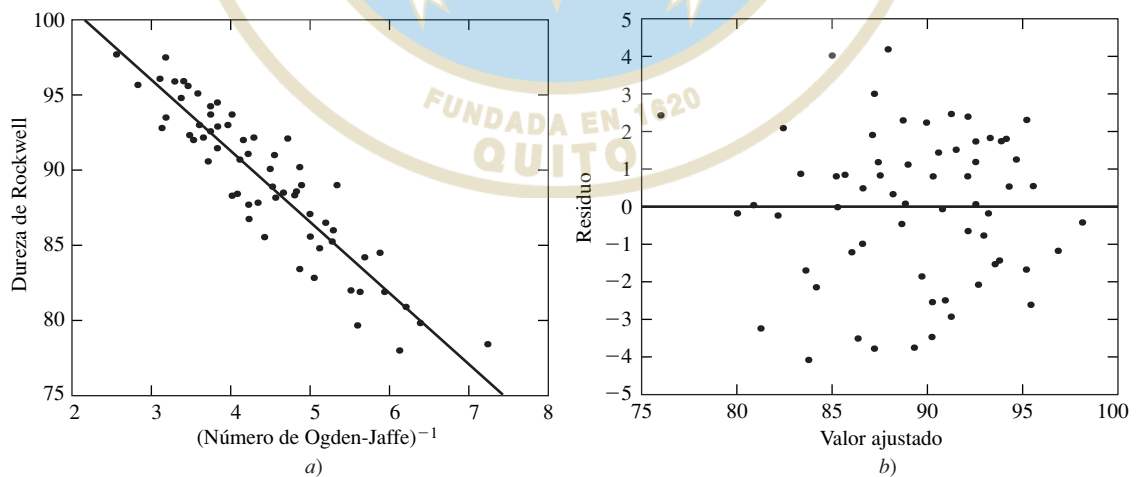


FIGURA 7.21 a) Gráfica de la dureza contra $(\text{número de Ogden-Jaffe})^{-1}$. La recta de mínimos cuadrados está sobrepuesta. b) Gráfica de residuos (e_i) contra valores ajustados $(\hat{y}_i)$ para estos datos. El modelo lineal queda bien.

variable independiente (x), mientras que en el ejemplo 7.22 se transforma la variable dependiente (y).

Ejemplo

7.25

Con referencia al ejemplo 7.21, la figura 7.17 presentó una gráfica de la producción contra el volumen de fluido fracturado para 255 pozos de gas. Ésta presenta una relación casi lineal válida entre el logaritmo de la producción y el del volumen del fluido fracturado. La figura 7.22 presenta los resultados. Observe que en este caso ambas variables fueron transformadas.

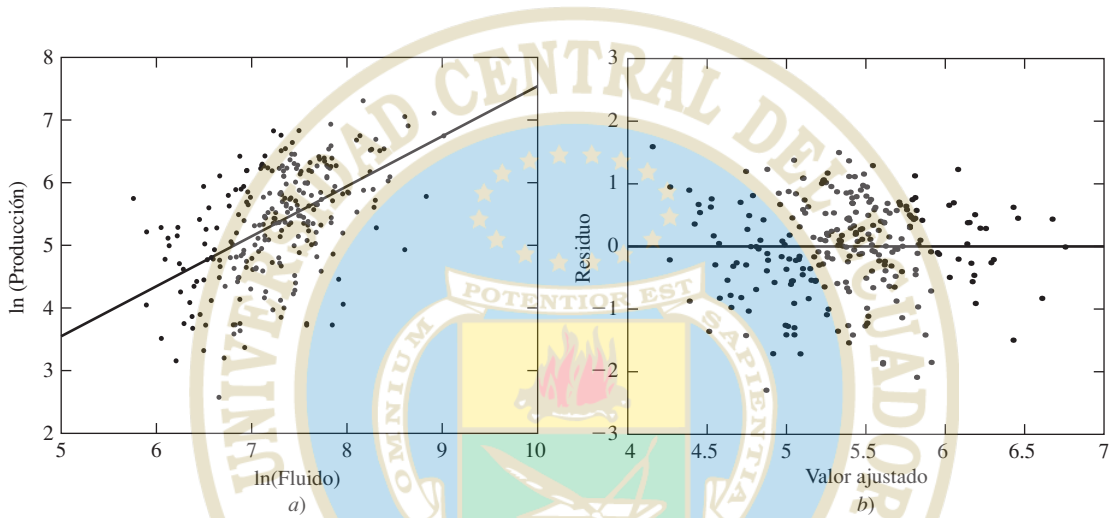


FIGURA 7.22 a) Gráfica del logaritmo de la producción contra el logaritmo del volumen de fluido fracturado para 255 pozos de gas, con la recta de mínimos cuadrados sobrepuesta. b) Gráfica de residuos contra valores ajustados. No hay patrón importante de los residuos. El modelo lineal queda bien.

Datos atípicos y puntos influyentes

Los datos atípicos son puntos separados del volumen de los datos. Tanto el diagrama de dispersión como la gráfica de residuos se deben revisar para encontrar datos atípicos. Lo primero respecto de un dato atípico es tratar de determinar por qué es diferente del resto de los puntos. A veces los datos atípicos los causan errores en el registro de datos o el funcionamiento defectuoso del equipo. En estos casos, los datos atípicos se pueden eliminar del conjunto de datos. Pero muchas veces la causa no se puede determinar con seguridad en un dato atípico. Eliminar el dato atípico resulta poco inteligente, porque da como resultado una subestimación de la variabilidad del proceso que generó los datos.

Con frecuencia los datos atípicos se pueden identificar mediante inspección visual. Mucho software enlista los puntos que tienen residuos inusualmente grandes; esta lista contendrá la mayor parte de los datos atípicos (y a veces también de algunos puntos inocuos). A veces al transformar las variables se eliminarán los datos atípicos cambiándolos de lugar más

cerca de la mayor parte de los datos. Cuando las transformaciones no ayudan, y cuando no hay justificación para eliminar los datos atípicos, un enfoque conveniente es ajustar la recta para todo el conjunto de datos, y eliminar cada dato atípico en turno, ajustando la recta al conjunto de datos con un dato atípico eliminado. Si ninguno de los datos atípicos eliminados hace una notable diferencia a la recta de mínimos cuadrados o a las desviaciones estándar estimadas de la pendiente y del intercepto, entonces utilice el ajuste con los datos atípicos incluidos. Si uno o más de los datos atípicos hacen una diferencia cuando se eliminan, entonces se debe reportar el rango de valores para los coeficientes de mínimos cuadrados. En estos casos se deben evitar los cálculos de los intervalos de confianza o de los de predicción, o la realización de las pruebas de hipótesis.

Un dato atípico que hace una diferencia considerable en la recta de mínimos cuadrados cuando se elimina se llama **punto influyente**. La figura 7.23 presenta un ejemplo de un dato atípico influyente, junto con uno que no es influyente. En general, los datos atípicos con valores de x inusuales tienen mayor probabilidad de ser influyentes que aquellos con valores de y inusuales, pero cada dato atípico se debe verificar. Mucho software identifica puntos potencialmente influyentes. Se puede encontrar más información acerca del tratamiento de datos atípicos y puntos influyentes en Draper y Smith (1998), Belsey, Kuh, Welsch (1980), y Cook y Weisberg (1994).

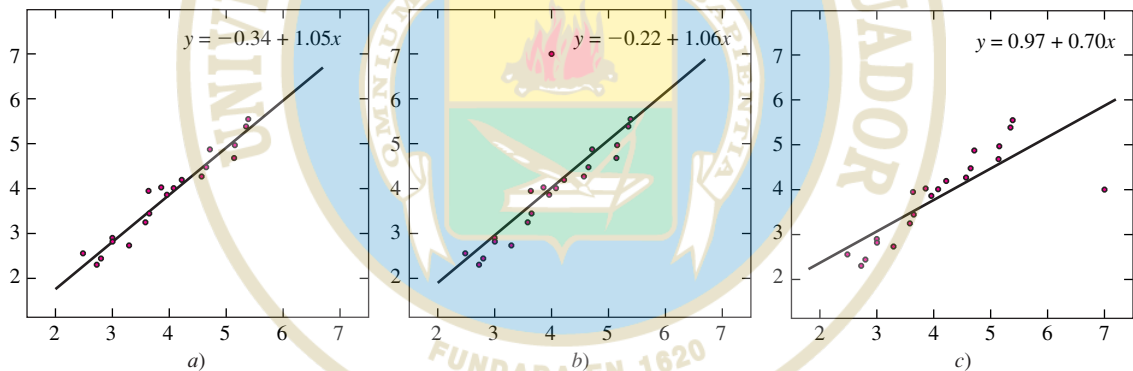


FIGURA 7.23 a) Diagrama de dispersión sin datos atípicos. b) Se agrega un dato atípico a la gráfica. Hay poco cambio en la recta de mínimos cuadrados, por lo que este punto no es influyente. c) Se agrega un dato atípico a la gráfica. Hay un cambio considerable en la recta de mínimos cuadrados, por lo que este punto es influyente.

Por último, se menciona que algunos escritores restringen la definición de datos atípicos a puntos con residuos inusualmente grandes. Según esta definición, un punto lejano de la mayor parte de los datos, aun cerca de la recta de mínimos cuadrados, no es un dato atípico. Ese punto podría ser o no influyente.

El siguiente ejemplo caracteriza un conjunto de datos que contiene dos datos atípicos. En un estudio para determinar si la frecuencia de cierto gene mutante aumenta con la edad, se contó el número de genes mutantes en un microgramo de ADN para cada uno de 30 hombres. Dos de los hombres tenían recuentos sumamente grandes; sus puntos son datos atípicos. La

recta de mínimos cuadrados fue ajustada para los 30 puntos, para cada conjunto de 29 puntos obtenidos al eliminar un atípico, y después para los 28 puntos que quedaban después de eliminar los dos datos atípicos. La figura 7.24 presenta diagramas de dispersión de frecuencia contra edad para todo el conjunto de datos y para los conjuntos con uno y con dos datos atípicos eliminados. Las rectas de mínimos cuadrados están sobrepuestas. Con los datos atípicos incluidos, la ecuación de la recta de mínimos cuadrados es $y = -137.76 + 4.54x$. Con los datos atípicos eliminados, la ecuación de la recta de mínimos cuadrados es $y = 31.86 + 1.23x$. Estos resultados son bastante diferentes, por lo que ambos se deben notificar.

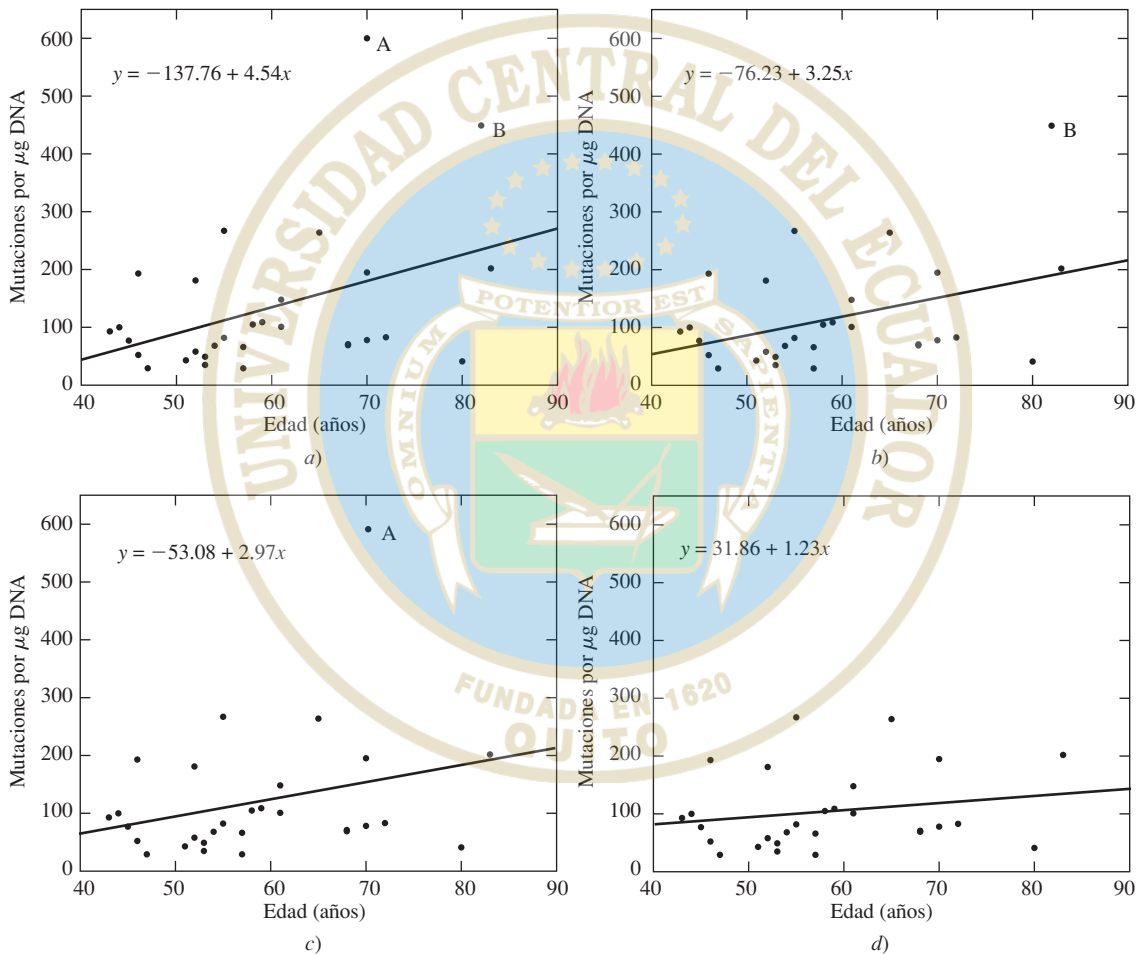


FIGURA 7.24 Frecuencia de mutación contra la edad. *a)* La gráfica contiene dos datos atípicos, A y B. *b)* El dato atípico A se elimina. El cambio en la recta de mínimos cuadrados es perceptible, aunque no extremo; este punto es un poco influyente. *c)* El dato atípico B se elimina. El cambio en la recta de mínimos cuadrados es otra vez perceptible, pero no extremo; también este punto es poco influyente. Observe que el dato atípico B es un poco más influyente que el dato atípico A, aunque visualmente B está cerca de la mayor parte de los datos. *d)* Ambos datos atípicos son eliminados. El efecto combinado en la recta de mínimos cuadrados es importante.

Otros métodos para transformar variables

Transformar variables no es el único método para analizar los datos cuando la gráfica de residuos indica un problema. Cuando ésta es heteroscedástica, algunas veces se utiliza una técnica llamada **mínimos cuadrados ponderados**. En este método, las coordenadas x y y de cada punto se multiplican por una cantidad conocida como **peso**. Los puntos en las regiones donde la dispersión vertical es grande se multiplican por pesos más pequeños, mientras que los puntos en regiones con dispersión vertical menor se multiplican por pesos más grandes. El efecto es lograr que los puntos cuya varianza del error es más pequeña tengan influencias mayores en el cálculo de la recta de mínimos cuadrados.

Cuando la gráfica de residuos muestra una tendencia, indica que se necesita más de una variable independiente para explicar la variación en la variable dependiente. En estos casos se agregan más variables independientes al modelo, y se utiliza la regresión múltiple. Por último, algunas relaciones son inherentemente no lineales. En éstas se puede aplicar un método llamado **regresión no lineal**. La regresión múltiple se trata en el capítulo 8. Los otros dos métodos están más allá del alcance de este libro. Una buena referencia sobre estos temas es Draper y Smith (1998).

Para resumir, en la figura 7.25 se presentan algunos ejemplos genéricos de gráficas de residuos. Para cada una se presentan un diagnóstico y una receta.

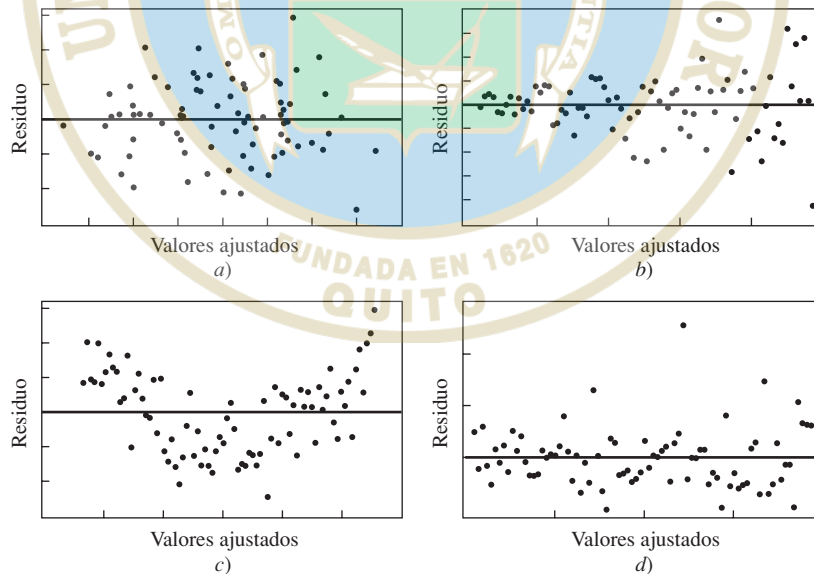


FIGURA 7.25 a) Ningún patrón importante, la gráfica es homoscedástica. El modelo lineal está bien. b) Heteroscedástica. Pruebe una transformación de potencia. c) Tendencia perceptible de los residuos. Pruebe una transformación de potencia o utilice la regresión múltiple. d) Dato atípico. Examine el punto de datos infractor para ver si hay un error. Si no, calcule la recta de mínimos cuadrados tanto con y sin el dato atípico para ver si hay una diferencia perceptible.

Comprobación de independencia y normalidad

Si la gráfica de residuos contra valores ajustados queda bien, podría ser aconsejable realizar diagnósticos adicionales para verificar más el ajuste del modelo lineal. En particular, cuando las observaciones ocurren en un orden de tiempo definido, es deseable graficar los residuos contra el orden con el cual las observaciones fueron hechas. Si hay tendencias en la gráfica, indica que la relación entre x y y puede estar variando con el tiempo. En estos casos una variable que representa el tiempo, u otras variables relacionadas con el tiempo, se deben incluir en el modelo como variables independientes adicionales, y se debe realizar una regresión múltiple.

En los datos de contaminación del aire del ejemplo 7.19, con y representando la concentración de ozono y x la concentración de NO_x , la gráfica de residuos (figura 7.15) para el modelo $\ln y = \beta_0 + \beta_1 x + \varepsilon$ es homoscedástica, sin patrón o tendencia perceptible. Se reunieron estos datos durante el transcurso de 359 días durante un año especial. La figura 7.26 presenta la gráfica de residuos contra el tiempo para estos datos. Hay un patrón claro. Los residuos positivos en medio de los datos corresponden al verano, y los negativos en los extremos de los datos, al invierno. Cada residuo es igual al logaritmo de la concentración de ozono observada en ese día, menos el logaritmo del valor pronosticado por el modelo. Se concluye que los valores pronosticados por el modelo son demasiado bajos en verano y demasiado altos en invierno. Es obvio que conocer la época del año puede mejorar el pronóstico de la concentración de ozono sobre el proporcionado por el modelo con la concentración de NO_x como la única variable independiente. Por tanto, se ajustaría bien un modelo de regresión múltiple que contendría tanto al tiempo como al NO_x como variables independientes. Dependiendo de los resultados de ese ajuste, se podrían hacer los ajustes adicionales al modelo. Estas ideas serán tratadas más profundamente en el capítulo 8.

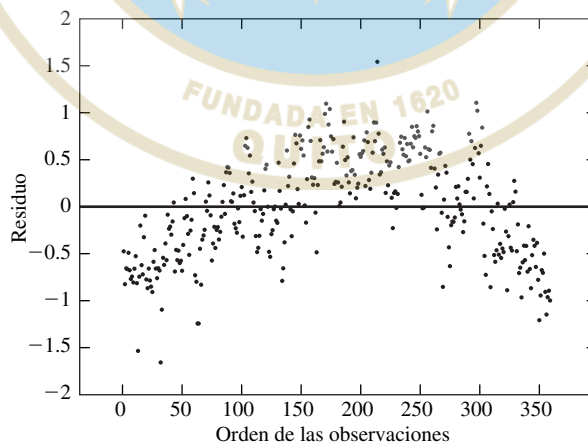


FIGURA 7.26 Gráfica de residuos contra el orden de las observaciones (tiempo) para el ozono contra los datos de NO_x . Se ajustó el modelo $\ln \text{ozono} = \beta_0 + \beta_1 \text{NO}_x + \varepsilon$. Los residuos muestran un patrón claro con el tiempo que indica que se debe ajustar a un modelo de regresión múltiple, con el tiempo como una variable adicional.

A veces una gráfica de residuos contra el tiempo muestra que aquéllos oscilan con el tiempo. Esto último indica que el valor de cada error está influenciado por los errores en las observaciones anteriores, por lo que los errores no son independientes. Cuando esta característica es grave no se debe usar la regresión lineal, y en su lugar se deben usar los métodos de análisis de serie de tiempo. Una buena referencia sobre el análisis de serie de tiempo es Brockwell y Davis (2002).

Para comprobar que los errores están normalmente distribuidos, se puede hacer una gráfica de probabilidad normal de los residuos. Si esta última tiene bruscamente la apariencia de una línea recta, ello significa que los residuos están casi normalmente distribuidos. Puede ser una buena idea hacer una gráfica de probabilidad cuando se transforman las variables, ya que una señal de una buena transformación es que los residuos estén casi normalmente distribuidos. Como ya se mencionó, el supuesto de la normalidad no es tan importante cuando el número de puntos de datos es grande. Desafortunadamente, cuando el número de puntos de datos es pequeño, puede ser difícil detectar las desviaciones de la normalidad.

Modelos empíricos y leyes físicas

¿Cómo se sabe si la relación entre dos variables es lineal? En algunos casos, las leyes físicas, como la ley de Hooke, proporcionan la garantía de que un modelo lineal es correcto. En otros, como en la relación entre el logaritmo del volumen de fluido fracturado bombeado en un pozo de gas y el logaritmo de su producción mensual, no existe una ley física conocida. En estos casos se utiliza un modelo lineal simplemente porque parece ajustarse a los datos del pozo. Un modelo que se elige porque parece ajustar a los datos, en ausencia de una teoría física, se llama **modelo empírico**. En la vida real, la mayoría de los análisis de datos están basados en modelos empíricos. Es menos frecuente cuando se aplica una ley física conocida. Por supuesto, muchas leyes físicas comenzaron como modelos empíricos. Si un modelo empírico se prueba en muchas ocasiones diferentes, bajo diversas circunstancias, y se encuentra que es válido sin excepción, puede obtener el estatus de una ley física.

Hay una diferencia importante entre la interpretación de los resultados basados en leyes físicas y la de los resultados basados en modelos empíricos. Una ley física se puede considerar *verdadera*, mientras que lo mejor que se puede esperar de un modelo empírico es que sea *útil*. Por ejemplo, en los datos de ley de Hooke se puede estar seguro que la relación entre la carga sobre el resorte y su longitud es realmente lineal. Se tiene la seguridad de que cuando se coloca otro peso en el resorte, la longitud del resorte se puede predecir con exactitud del modelo lineal. Por otro lado, para los datos del pozo de gas, mientras que la relación lineal describe los datos del pozo, no se puede asegurar que capta la verdadera relación entre el volumen del fluido fracturado y la producción.

He aquí un ejemplo simple que ilustra el punto. La figura 7.27 de la página 542 presenta 20 triángulos de formas diferentes. Suponga que no se conoce la fórmula para el área de un triángulo. Sin embargo, observe que los triángulos con perímetros mayores parecen tener áreas más grandes, así que se ajusta un modelo lineal:

$$\text{Área} = \beta_0 + \beta_1 (\text{Perímetro}) + \varepsilon$$

El diagrama de dispersión del área contra el perímetro, con la recta de mínimos cuadrados sobrepuesta, se muestra a la derecha de la figura 7.27. La ecuación de la recta de mínimos cuadrados es

$$\text{Área} = -1.232 + 1.373 (\text{Perímetro})$$

Las unidades en esta ecuación son arbitrarias. La correlación entre área y perímetro es $r = 0.88$, que es fuertemente positiva. El modelo lineal parece ajustar bien. Se podría utilizar este modelo para predecir, por ejemplo, que un triángulo con perímetro igual a 5 tendrá una área de 5.633.

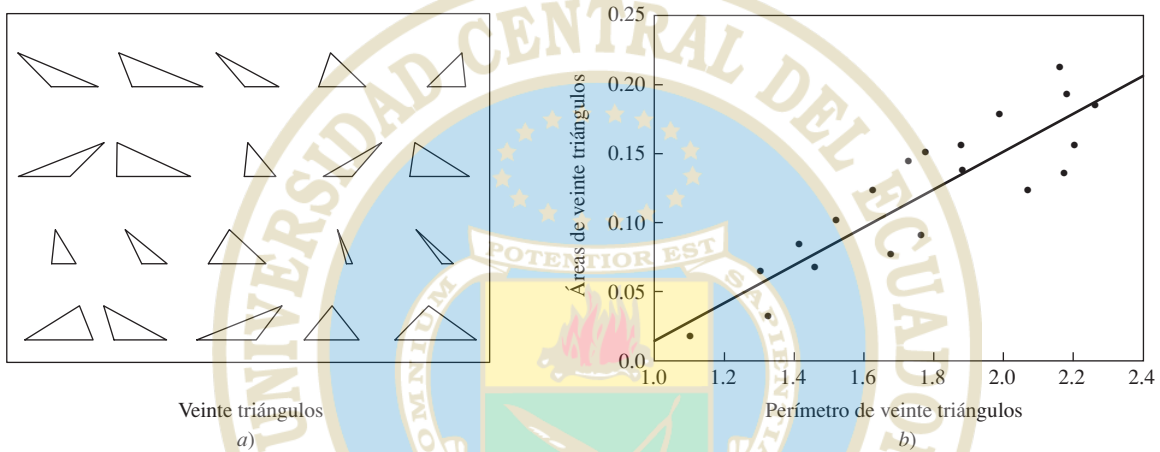


FIGURA 7.27 a) Veinte triángulos. b) Área contra perímetro para 20 triángulos. La correlación entre perímetro y área es 0.88.

Ahora aunque este modelo lineal podría ser útil, no es verdadero. El modelo lineal muestra que hay una fuerte tendencia de que triángulos con perímetros mayores tengan áreas más grandes. En ausencia de un mejor método, éste podría ser calcular algunas de las áreas de los triángulos. Pero no ayuda para mostrar el verdadero mecanismo detrás de la determinación del área. El mecanismo verdadero, por supuesto, está dado por la ley

$$\text{Área} = 0.5 \times \text{base} \times \text{altura}$$

Los resultados pronosticados por un modelo empírico podrían no ser válidos bajo la reproducción. Por ejemplo, un conjunto de triángulos podría estar diseñado de tal manera que algunos con perímetros mayores tuvieran áreas más pequeñas. En otro conjunto, el área podría parecer ser proporcional al cuadrado del perímetro, o su logaritmo. No se puede determinar con el análisis estadístico de los datos de los triángulos qué tan bien será aplicable el modelo empírico a un triángulo que aún no se observa. Determinar si es adecuado aplicar los resultados de un modelo empírico a las futuras observaciones es un tema de juicio científico más que estadístico.

Resumen

- Las leyes físicas son aplicables a todas las futuras observaciones.
- Un modelo empírico es válido solamente para los datos a los que se ajusta. Esto último puede o no ser útil para predecir los resultados para las siguientes observaciones.
- Determinar si se aplica un modelo empírico a una futura observación requiere juicio científico más que análisis estadístico.

Ejercicios de la sección 7.4

1. El siguiente resultado de MINITAB corresponde a la recta de mínimos cuadrados $\ln y = \beta_0 + \beta_1 \ln x + \varepsilon$, donde y representa la producción mensual de un pozo de gas y x el volumen de fluido fracturado bombeado dentro. (Un diagrama de dispersión de estos datos se presentó en la figura 7.22.)

Regression Analysis: LN PROD versus LN FLUID

The regression equation is

$$\text{LN PROD} = -0.444 + 0.798 \text{ LN FLUID}$$

Predictor	Coef	SE Coef	T	P
Constant	-0.4442	0.5853	-0.76	0.449
LN FLUID	0.79833	0.08010	9.97	0.000

S = 0.7459 R-Sq = 28.2% R-Sq(adj) = 27.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	55.268	55.268	99.34	0.000
Residual Error	253	140.756	0.556		
Total	254	196.024			

Predicted Values for New Observations

New Obs	Fit	SE Fit	95.0% CI	95.0% PI
1	5.4457	0.0473	(5.3526, 5.5389)	(3.9738, 6.9176)

Values of Predictors for New Observations

New Obs	LN FLUID
1	7.3778

- ¿Cuál es la ecuación de la recta de mínimos cuadrados para pronosticar $\ln y$ a partir de $\ln x$?
- Pronostique la producción de un pozo en el cual se han bombeado 2 500 gal/pie de fluido.
- Pronostique la producción de un pozo en el que se han bombeado 1 600 gal/pie de fluido.
- Determine un intervalo de predicción de 95% para la producción de un pozo en el que se han bombeado 1 600 gal/pie de fluido. (Nota: $\ln 1\,600 = 7.3778$)

2. El procesamiento de carbón sin tratar implica el “lavado”, en donde se elimina la ceniza de carbón (no orgánico, material incombustible). El artículo “Quantifying Sampling Precision for Coal Ash Using Gy’s Discrete Model of the Fundamental Error” (*Journal of Coal Quality*, 1989:33-39) proporciona los datos relacionados con los porcentajes de ceniza con el volumen de una partícula de carbón. Se midieron los porcentajes promedio de ceniza para seis volúmenes de partículas de carbón. Los datos son los siguientes:

Volumen (cm ³)	0.01	0.06	0.58	2.24	15.55	276.02
Porcentaje de ceniza	3.32	4.05	5.69	7.06	8.17	9.36

- Calcule la recta de mínimos cuadrados para predecir el porcentaje de ceniza (y) a partir del volumen (x). Grafique los residuos contra los valores ajustados. ¿El modelo lineal parece adecuado? Explique.
 - Calcule la recta de mínimos cuadrados para predecir el porcentaje de ceniza a partir del $\ln$ volumen. Grafique los residuos contra los valores ajustados. ¿El modelo lineal parece adecuado? Explique.
 - Calcule la recta de mínimos cuadrados para predecir el porcentaje de ceniza a partir del $\sqrt{\text{volumen}}$. Grafique los residuos contra los valores ajustados. ¿El modelo lineal parece adecuado? Explique.
 - Utilizando el modelo más adecuado, pronostique el porcentaje de ceniza para partículas con un volumen de 50 m³.
 - Utilizando el modelo más adecuado, construya un intervalo de confianza de 95% para la media de un porcentaje de ceniza para partículas con un volumen de 50 m³.
3. Para determinar el efecto de la temperatura sobre la producción de cierto proceso químico, el proceso se opera 24 veces en diferentes temperaturas. La temperatura (en °C) y la producción (expresada como un porcentaje de un máximo teórico) para cada operación está dada en la tabla siguiente. Los resultados se presentan en el orden en el que se operaron, de los primeros a los últimos.

Orden	Temp	Producción	Orden	Temp	Producción	Orden	Temp	Producción
1	30	49.2	9	25	59.3	17	34	65.9
2	32	55.3	10	38	64.5	18	43	75.2
3	35	53.4	11	39	68.2	19	34	69.5
4	39	59.9	12	30	53.0	20	41	80.8
5	31	51.4	13	30	58.3	21	36	78.6
6	27	52.1	14	39	64.3	22	37	77.2
7	33	60.2	15	40	71.6	23	42	80.3
8	34	60.5	16	44	73.0	24	28	69.5

- Calcule la recta de mínimos cuadrados para predecir la producción (y) de temperatura (x).
 - Grafique los residuos contra los valores ajustados. ¿El modelo lineal parece adecuado? Explique.
 - Grafique los residuos contra el orden en el que se hicieron las observaciones. ¿Hay una tendencia en los residuos con el tiempo? ¿El modelo lineal parece adecuado? Explique.
4. En una roca detonada con explosivos, la velocidad máxima de la partícula (VMP) depende de la distancia de la explosión y de la cantidad de carga. El artículo “Prediction of Particle Velocity Caused by Blasting for an Infrastructure Excavation Covering Granite Bedrock” (A. Kahrman, *Mineral Resources Engineering*, 2001:205-218) sugiere que predecir la VMP (y) a partir de la distancia escalada (x), que es igual a la distancia dividida entre la raíz cuadrada de la carga. Los resultados para 15 explosiones se presentan en la tabla siguiente.

VMP (mm/s)	Distancia escalada (m/kg ^{0.5})
1.4	47.33
15.7	9.6
2.54	15.8
1.14	24.3
0.889	23.0
1.65	12.7
1.4	39.3
26.8	8.0
1.02	29.94
4.57	10.9
6.6	8.63
1.02	28.64
3.94	18.21
1.4	33.0
1.4	34.0

- a) Grafique la VMP contra la distancia escalada. ¿La relación parece ser lineal?
- b) Calcule la recta de mínimos cuadrados para el modelo $\ln \text{VMP} = \beta_0 + \beta_1 \ln \text{distancia escalada} + \varepsilon$. Grafique los residuos contra los valores ajustados. ¿Este modelo lineal parece adecuado?
- c) Utilice la recta de mínimos cuadrados calculada en el inciso b) para predecir la VMP cuando la distancia escalada es 20. Determine un intervalo de predicción de 95%.
5. La buena previsión y el control de las actividades de preconstrucción conducen al uso más eficiente del tiempo y recursos en proyectos de construcción de autopistas. Los datos acerca de los costos de construcción (en miles de dólares) y las horas-persona de trabajo requeridas para varios proyectos se presentan en la tabla siguiente y fueron tomados del artículo “Forecasting Engineering Manpower Requirements for Highway Preconstruction Activities” (K. Persad, J. O’Connor, y K. Varghese, *Journal of Management Engineering*, 1995:41-47). Cada valor representa un promedio de algunos proyectos, y se han eliminado dos datos atípicos.

Horas-persona (x)	Costo (y)	Horas-persona (x)	Costo (y)
939	251	1 069	355
5 796	4 690	6 945	5 253
289	124	4 159	1 177
283	294	1 266	802
138	138	1 481	945
2 698	1 385	4 716	2 327
663	345		

- a) Calcule la recta de mínimos cuadrados para predecir y a partir de x .
- b) Grafique los residuos contra los valores ajustados. ¿El modelo parece adecuado?
- c) Calcule la recta de mínimos cuadrados para predecir $\ln y$ a partir de $\ln x$.
- d) Grafique los residuos contra los valores ajustados. ¿El modelo parece adecuado?
- e) Usando el modelo más adecuado, construya un intervalo de predicción de 95% para el costo de un proyecto que requiere mil horas-persona de trabajo.
6. El artículo “Oxidation State and Activities of Chromium Oxides in CaO-SiO₂-CrO_x Slag System” (Y. Xiao, L. Holappa y M. Reuter, *Metallurgical and Materials Transactions B*, 2002:595-603) presenta la cantidad x (en porcentaje molar) y el coefi-

cienta de actividad y de $\text{CrO}_{1.5}$ para varios elementos. Los datos, extraídos de una tabla más grande, se presentan en la tabla siguiente.

x	y	x	y	x	y
10.20	2.6	7.13	5.8	5.33	13.1
5.03	19.9	3.40	29.4	16.70	0.6
8.84	0.8	5.57	2.2	9.75	2.2
6.62	5.3	7.23	5.5	2.74	16.9
2.89	20.3	2.12	33.1	2.58	35.5
2.31	39.4	1.67	44.2	1.50	48.0

- Calcule la recta de mínimos cuadrados para predecir y a partir de x .
 - Grafique los residuos contra los valores ajustados.
 - Calcule la recta de mínimos cuadrados para predecir y a partir de $1/x$.
 - Grafique residuos contra valores ajustados.
 - Usando la mejor línea ajustada, encuentre un intervalo de confianza de 95% para la media del valor de y cuando $x = 5.0$.
7. Se utiliza un molino de viento para generar corriente continua. Se reúnen datos en 45 días diferentes para determinar la relación entre la velocidad del viento en mi/h (x) y la corriente en kA (y). Los datos se presentan en la tabla siguiente.

Día	Velocidad del viento	Corriente	Día	Velocidad del viento	Corriente	Día	Velocidad del viento	Corriente
1	4.2	1.9	16	3.7	2.1	31	2.6	1.4
2	1.4	0.7	17	5.9	2.2	32	7.7	2.8
3	6.6	2.2	18	6.0	2.6	33	6.1	2.4
4	4.7	2.0	19	10.7	3.2	34	5.5	2.2
5	2.6	1.1	20	5.3	2.3	35	4.7	2.3
6	5.8	2.6	21	5.1	1.9	36	4.0	2.0
7	1.8	0.3	22	4.9	2.3	37	2.3	1.2
8	5.8	2.3	23	8.3	3.1	38	11.9	3.0
9	7.3	2.6	24	7.1	2.3	39	8.6	2.5
10	7.1	2.7	25	9.2	2.9	40	5.6	2.1
11	6.4	2.4	26	4.4	1.8	41	4.2	1.7
12	4.6	2.2	27	8.0	2.6	42	6.2	2.3
13	1.6	1.1	28	10.5	3.0	43	7.7	2.6
14	2.3	1.5	29	5.1	2.1	44	6.6	2.9
15	4.2	1.5	30	5.8	2.5	45	6.9	2.6

- Calcule la recta de mínimos cuadrados para predecir y a partir de x . Realice una gráfica de residuos contra valores ajustados.
- Calcule la recta de mínimos cuadrados para predecir y a partir de $\ln x$. Realice una gráfica de residuos contra valores ajustados.
- Calcule la recta de mínimos cuadrados para predecir $\ln y$ a partir de x . Realice una gráfica de residuos contra valores ajustados.
- Calcule la recta de mínimos cuadrados para predecir $\sqrt{y}$, a partir de x . Realice una gráfica de residuos contra valores ajustados.
- ¿Cuál de los cuatro modelos de la a) a la d) se ajusta mejor? Explique.
- Para el modelo que ajusta mejor, grafique los residuos contra el orden en el cual se realizaron las observaciones. ¿Los residuos parecen variar con el tiempo?
- Utilizando el mejor modelo, pronostique la corriente cuando la velocidad del viento es 5.0 mi/h.

h) Utilizando el mejor modelo, determine un intervalo de predicción de 95% para la corriente en una día específico cuando la velocidad del viento es 5.0 mi/h.

8. Dos detectores de radón se colocaron en ubicaciones diferentes en el sótano de una casa. Cada uno proporcionó mediciones horarias de la concentración de radón, en unidades de pCi/L. Los datos se presentan en la tabla siguiente.

R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
1.2	1.2	3.4	2.0	4.0	2.6	5.5	3.6
1.3	1.5	3.5	2.0	4.0	2.7	5.8	3.6
1.3	1.6	3.6	2.1	4.3	2.7	5.9	3.9
1.3	1.7	3.6	2.1	4.3	2.8	6.0	4.0
1.5	1.7	3.7	2.1	4.4	2.9	6.0	4.2
1.5	1.7	3.8	2.2	4.4	3.0	6.1	4.4
1.6	1.8	3.8	2.2	4.7	3.1	6.2	4.4
2.0	1.8	3.8	2.3	4.7	3.2	6.5	4.4
2.0	1.9	3.9	2.3	4.8	3.2	6.6	4.4
2.4	1.9	3.9	2.4	4.8	3.5	6.9	4.7
2.9	1.9	3.9	2.4	4.9	3.5	7.0	4.8
3.0	2.0	3.9	2.4	5.4	3.5		

- a) Calcule la recta de mínimos cuadrados para predecir la concentración de radón en la posición 2 a partir de la concentración en la posición 1.
- b) Grafique los residuos contra los valores ajustados. ¿El modelo lineal parece adecuado?
- c) Divida los datos en dos grupos: los puntos donde $R_1 < 4$ en un grupo, donde $R_1 \geq 4$ en los demás. Calcule la recta de mínimos cuadrados y la gráfica de residuos para cada grupo. ¿La recta describe bien cualquier grupo? ¿Cuál?
- d) Explique por qué podría ser una buena idea ajustar un modelo lineal a una parte de estos datos, y un modelo no lineal a la otra.
9. El artículo “The Equilibrium Partitioning of Titanium Between Ti^{3+} and Ti^{4+} Valency States in $CaO-SiO_2-TiO_x$ Slags” (G. Trannell, O. Ostrovski y S. Jahanshahi, en *Metallurgical and Materials Transactions B*, 2002:61-66) analiza la relación entre el Ti^{3+}/Ti^{4+} de proporción de redox y la presión parcial de oxígeno p_{O_2} en $CaO-SiO_2-TiO_x$. Algunas mediciones independientes del cociente redox se hicieron en cada uno de cinco presiones parciales diferentes: 10^{-7} , 10^{-8} , 10^{-9} , 10^{-10} y 10^{-12} atmósferas. Los resultados para las operaciones en 14% de la masa de TiO_x se presentan en la tabla siguiente.

Presión parcial de oxígeno	Mediciones del cociente redox
10^{-7}	0.011, 0.017, 0.034, 0.039
10^{-8}	0.018, 0.011, 0.026, 0.050, 0.034, 0.068, 0.061
10^{-9}	0.027, 0.038, 0.076, 0.088
10^{-10}	0.047, 0.069, 0.123, 0.162
10^{-12}	0.160, 0.220, 0.399, 0.469

- a) Denotando el cociente redox por y y la presión parcial por x , la teoría establece que y debe ser proporcional a x^β para alguna β . Expresa esta relación teórica como un modelo lineal.
- b) Calcule la recta de mínimos cuadrados para este modelo lineal. Grafique los residuos contra los valores ajustados. ¿El modelo lineal es válido?
- c) Consideraciones teóricas adicionales sugieren que bajo las condiciones de este experimento, y debe ser proporcional a $x^{-1/4}$, ¿los datos de la tabla anterior son consistentes con esta teoría? Explique.
10. El artículo “The Selection of Yeast Strains for the Production of Premium Quality South African Brandy Base Products” (C. Steger y M. Lambrechts, en *Journal of Industrial Microbiology and Biotechnology*, 2000:431-440) se presenta información de-

tallada sobre la composición compuesta volátil de la base de vino realizada para cada una de 16 variedades de levadura seleccionadas. Debajo están las concentraciones de ésteres totales y ácidos volátiles totales (en mg/l) en cada uno de los vinos.

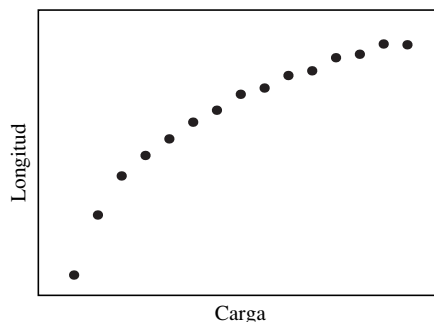
Ésteres	Ácidos	Ésteres	Ácidos	Ésteres	Ácidos	Ésteres	Ácidos
284.34	445.70	173.01	265.43	229.55	210.58	312.95	203.62
215.34	332.59	188.72	166.73	144.39	254.82	172.79	342.21
139.38	356.88	197.81	291.72	303.28	215.83	256.02	152.38
658.38	192.59	105.14	412.42	295.24	442.55	170.41	391.30

- Construya un diagrama de dispersión de la concentración de ácido contra la concentración de ésteres. Indique el dato atípico.
- Calcule los coeficientes de la recta de mínimos cuadrados para predecir el nivel de ácido (y) a partir del nivel de ésteres (x), junto con sus desviaciones estándar estimadas.
- Calcule el P -valor de que la prueba de hipótesis nula $H_0: \beta_1 = 0$.
- Elimine el dato atípico, y calcule nuevamente los coeficientes de la recta de mínimos cuadrados, junto con sus desviaciones estándar estimadas.
- Calcule el P -valor de la prueba de la hipótesis nula $H_0: \beta_1 = 0$ para los datos con el dato atípico eliminado.
- ¿Un modelo lineal parece útil para predecir la concentración de ácido de la concentración de ésteres? Explique.

11. El artículo “Mathematical Modeling of the Argon-Oxygen Decarburization Refining Process of Stainless Steel: Part II. Application of the Model to Industrial Practice” (J. Wei y D. Zhu, en *Metallurgical and Materials Transactions B*, 2001:212-217) presenta el contenido de carbono (en % de masa) y la temperatura del baño (en K) para 32 calores de acero inoxidable austenítico. Estos datos se muestran en la tabla siguiente.

% de carbono	Temperatura	% de carbono	Temperatura	% de carbono	Temperatura	% de carbono	Temperatura
19	1 975	17	1 984	18	1 962	17	1 983
23	1 947	20	1 991	19	1 985	20	1 966
22	1 954	19	1 965	19	1 946	21	1 972
16	1 992	22	1 963	15	1 986	17	1 989
17	1 965	18	1 949	20	1 946	18	1 984
18	1 971	22	1 960	22	1 950	23	1 967
12	2 046	20	1 960	15	1 979	13	1 954
24	1 945	19	1 953	15	1 989	15	1 977

- Calcule la recta de mínimos cuadrados para predecir la temperatura del baño (y) del contenido de carbono (x).
 - Identifique dos datos atípicos. Calcule las dos rectas de mínimos cuadrados que resultan de la eliminación de cada dato atípico por separado, y la recta de mínimos cuadrados que da como resultado la eliminación de ambos datos atípicos.
 - ¿Las rectas de los mínimos cuadrados calculados en los incisos a) y b) son similares? Si es así, reporte la recta que ajustó a todo el conjunto de datos, junto con los intervalos de confianza de 95% para la pendiente y el intercepto. Si no, informe el rango de pendientes, sin un intervalo de confianza.
12. El artículo “Mechanistic-Empirical Design of Bituminous Roads: An Indian Perspective” (A. Das y B. Pandey, en *Journal of Transportation Engineering*, 1999:463-471) presenta una ecuación de la forma $y = a(1/x_1)^b(1/x_2)^c$ para predecir el número de las repeticiones para fallas por fatiga de laboratorio (y) en relación con la fuerza de tensión en la parte baja de la viga bituminosa (x_1), y con el módulo de resiliencia (x_2). Transforme esta ecuación en un modelo lineal, y exprese los coeficientes del modelo lineal en función de a , b y c .
13. Un ingeniero quiere determinar la constante del resorte para un resorte especial. Cuelga cada vez varias pesas al final del resorte y mide la longitud de éste. En la siguiente figura se muestra un diagrama de dispersión de longitud (y) contra carga (x).



- ¿El modelo $y = \beta_0 + \beta_1 x$ es un modelo empírico o una ley física?
- ¿El ingeniero debe transformar las variables para hacer la relación más lineal, o sería mejor repetir el experimento? Explique.

Ejercicios adicionales para el capítulo 7

- La ley de Beer-Lambert relaciona la absorción A de una solución con la concentración C de una especie en solución por $A = MLC$, donde L es la longitud de la trayectoria y M el coeficiente de absorción molar. Suponga que $L = 1$ cm. Las mediciones de A se realizan a diferentes concentraciones. Los datos se presentan en la tabla siguiente.

Concentración (mol/cm ³)	1.00	1.20	1.50	1.70	2.00
Absorción	0.99	1.13	1.52	1.73	1.96

- Sea $A = \hat{\beta}_0 + \hat{\beta}_1 C$ la ecuación de la recta de mínimos cuadrados para predecir la absorción (A) a partir de la concentración (C). Calcule los valores de $\hat{\beta}_0$ y $\hat{\beta}_1$.
 - ¿Qué valor le asigna la ley de Beer-Lambert a β_0 ?
 - ¿Cuál es la cantidad física que estima $\hat{\beta}_1$?
 - Pruebe $H_0 : \beta_0 = 0$. ¿Este resultado es consistente con la ley de Beer-Lambert?
- En una prueba de material de guerra se dejó caer un gran número de bombas sobre un objetivo desde diferentes alturas. La velocidad inicial de las bombas en dirección a la tierra fue 0. Sea y la altura en metros a partir de cual se deja caer una bomba, x el tiempo en segundos para que la bomba pegue en el suelo, sea $w = x^2$, y $v = \sqrt{y}$. La relación entre x y y está dada por $y = 4.9x^2$. Para cada uno de los siguientes pares de variables, establezca si el coeficiente de correlación es un resumen adecuado.
 - x y y
 - w y y
 - x y v
 - w y v
 - $\ln x$ y $\ln y$
 - Las erupciones del géiser Old Faithful en el Parque Nacional de Yellowstone generalmente duran de 1.5 a 5 minutos. Entre las erupciones hay periodos de inactividad, que generalmente duran de 50 a 100 minutos. Un periodo de inactividad también se puede considerar como el tiempo de espera entre las erupciones. Las duraciones en minutos para 60 periodos de inactividad consecutivos se muestran en la tabla siguiente. Se desea predecir la duración de un periodo de inactividad a partir de la duración del que le precede inmediatamente. Para expresar esto en forma simbólica, denote la secuencia de periodos de inactividad $T_1, \dots, T_{60}$. Se quiere predecir T_{i+1} a partir de T_i .

i	T_i	i	T_i	i	T_i	i	T_i	i	T_i
1	80	11	56	21	82	31	88	41	72
2	84	12	80	22	51	32	51	42	75
3	50	13	69	23	76	33	80	43	75
4	93	14	57	24	82	34	49	44	66
5	55	15	90	25	84	35	82	45	84
6	76	16	42	26	53	36	75	46	70
7	58	17	91	27	86	37	73	47	79
8	74	18	51	28	51	38	67	48	60
9	75	19	79	29	85	39	68	49	86
10	80	20	53	30	45	40	86	50	71
								60	57

- a) Construya un diagrama de dispersión de los puntos (T_i, T_{i+1}) para $i = 1, \dots, 59$.
- b) Calcule la recta de mínimos cuadrados para predecir T_{i+1} a partir de T_i . (Sugerencia: los valores de la variable (x) independiente son $T_1, \dots, T_{59}$, y los valores de la variable dependiente (y) son $T_2, \dots, T_{60}$).
- c) Determine un intervalo de confianza del 95% para la pendiente β_1 .
- d) Si los tiempos de espera antes de la última erupción eran de 70 minutos, ¿cuál es el tiempo de espera pronosticado antes de la próxima erupción?
- e) Determine un intervalo de confianza del 98% para la media del tiempo de espera antes de la próxima erupción cuando el tiempo de espera de la última erupción fue de 70 minutos.
- f) Determine un intervalo de predicción del 99% para el tiempo de espera antes de la próxima erupción, si el tiempo de espera antes de la última erupción fue de 70 minutos.
4. Con referencia al ejercicio 3.
- a) Grafique los residuos contra los valores ajustados. ¿La gráfica indica incumplimientos serios de los supuestos usuales?
- b) Grafique los residuos contra el orden de los datos. ¿La gráfica indica violaciones serias de los supuestos usuales?
5. Un químico está calibrando un espectrómetro que se utilizará para medir la concentración de monóxido de carbono (CO) en muestras atmosféricas. Para comprobar la calibración, se miden muestras de concentración conocida. Las concentraciones verdaderas (x) y las medidas (y) están dadas en la tabla siguiente. Debido al error aleatorio, mediciones repetidas en la misma muestra variarán. Se considera que la máquina está calibrada si su media de respuesta es igual a la concentración verdadera.

Concentración verdadera (ppm)	Concentración medida (ppm)
0	1
10	11
20	21
30	28
40	37
50	48
60	56
70	68
80	75
90	86
100	96

Para comprobar la calibración se ajusta al modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$. Idealmente, el valor de β_0 debe ser 0 y el valor β_1 debe ser 1.

- Calcule los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ de mínimos cuadrados.
 - ¿Se puede rechazar la hipótesis nula $H_0 : \beta_0 = 0$?
 - ¿Se puede rechazar la hipótesis nula $H_1 : \beta_1 = 1$?
 - ¿Los datos proporcionan suficiente evidencia para concluir que la máquina está fuera de calibración?
 - Calcule un intervalo de confianza de 95% para la media de la medición cuando la concentración verdadera es 20 ppm.
 - Calcule un intervalo de confianza de 95% para la media de la medición cuando la concentración verdadera es 80 ppm.
 - Alguien afirma que la máquina está calibrada para concentraciones cerca de 20 ppm, ¿estos datos proporcionan suficiente evidencia para rechazar esta afirmación? Explique.
6. El artículo “Experimental Measurement of Radiative Heat Transfer in Gas-Solid Suspension Flow System” (G. Han, K. Tuzla y J. Chen, en *AiChE Journal*, 2002: 1910-1916) analiza la calibración de un radiómetro. Se hicieron algunas mediciones de la fuerza electromotriz en lecturas del radiómetro (en volts) y el flujo de radiación (en kilowatts por metro cuadrado). Los resultados (leídos de una gráfica) se presentan en la tabla siguiente.

Flujo de calor (y)	15	31	51	55	67	89
Señal de salida (x)	1.08	2.42	4.17	4.46	5.17	6.92

- Calcule la recta de mínimos cuadrados para predecir el flujo de calor a partir de la señal de salida.
 - Si el radiómetro lee 3.00 V, pronostique el flujo de calor.
 - Si el radiómetro lee 8.00 V, ¿se puede predecir el flujo de calor? Si es así, predígalos. Si no, explique.
7. El artículo “Effect of Temperature on the Marine Immersion Corrosion of Carbon Steels” (R. Melchers, en *Corrosion*, 2002:768-781) presenta mediciones de pérdida por corrosión (en mm) que soportan elementos de acero y de cobre sumergidos en agua de mar en 14 ubicaciones diferentes. Para cada ubicación se registró la media de la pérdida por corrosión (en mm), junto con la media de la temperatura del agua (en °C). Los resultados, después de un año de inmersión, se presentan en la tabla siguiente.

Corrosión	Media de la temperatura	Corrosión	Media de la temperatura
0.2655	23.5	0.2200	26.5
0.1680	18.5	0.0845	15.0
0.1130	23.5	0.1860	18.0
0.1060	21.0	0.1075	9.0
0.2390	17.5	0.1295	11.0
0.1410	20.0	0.0900	11.0
0.3505	26.0	0.2515	13.5

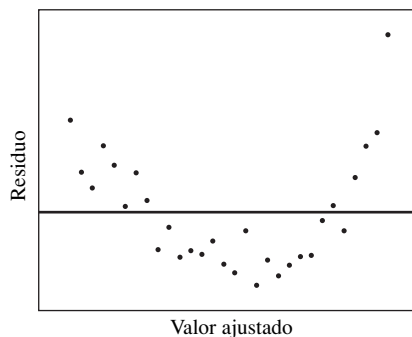
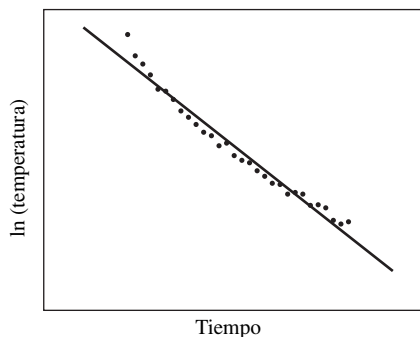
- Calcule la recta de mínimos cuadrados para predecir la pérdida por corrosión (y) para la media de la temperatura (x).
 - Determine un intervalo de confianza de 95% para la pendiente.
 - Determine un intervalo de confianza de 95% para la media de la pérdida de corrosión a una media de temperatura de 20°C.
 - Determine un intervalo de predicción de 90% para la pérdida por corrosión de un elemento inmerso a una media de la temperatura de 20°C.
8. El artículo “Measurements of the Thermal Conductivity and Thermal Diffusivity of Polymer Melts with the Short-Hot-Wire Method” (X. Zhang, W. Hendro y cols., en *International Journal of Thermophysics*, 2002: 1077–1090) presenta mediciones de la conductividad térmica y difusividad de algunos polímeros a diferentes temperaturas. Los resultados para la difusividad térmica de policarbonato (en $10^{-7} \text{ m}^2\text{s}^{-1}$) se presentan en la tabla siguiente.

Temperatura		Temperatura		Temperatura		Temperatura	
Difusividad	(°C)	Difusividad	(°C)	Difusividad	(°C)	Difusividad	(°C)
1.43	28	1.36	107	1.05	159	1.26	215
1.53	38	1.34	119	1.13	169	0.86	225
1.43	61	1.29	130	1.03	181	1.01	237
1.34	83	1.36	146	1.06	204	0.98	248

- a) Calcule la recta de mínimos cuadrados para predecir la difusividad (y) a partir de la temperatura (x).
- b) Determine un intervalo de confianza de 95% para la pendiente.
- c) Determine un intervalo de confianza de 95% para la difusividad de policarbonato a 100°C.
- d) Determine un intervalo de predicción de 95% para la difusividad de policarbonato a 100°C.
- e) ¿Cuál es más útil, el intervalo de confianza o el de predicción? Explique.
9. El artículo “Copper Oxide Mounted on Activated Carbon as Catalyst for Wet Air Oxidation of Aqueous Phenol. 1. Kinetic and Mechanistic Approaches” (P. Alvarez, D. McLurgh y P. Plucinski, en *Industrial Engineering and Chemistry Research*, 2002:2147-2152) notifica resultados de experimentos para describir el mecanismo de oxidación catalítica húmeda aérea de fenol acuoso. En un conjunto de experimentos se midieron la tasa de oxidación inicial (en kilogramos de fenol por kilogramo de catalizador por hora) y la concentración de oxígeno (en mol/m³). Los resultados (leídos de una gráfica) se presentan en la tabla siguiente.
- | | | | | | |
|---|------|------|------|------|------|
| Tasa (y) | 0.44 | 0.49 | 0.60 | 0.64 | 0.72 |
| Concentración de O ₂ (x) | 3.84 | 4.76 | 6.08 | 7.06 | 8.28 |
- a) Se sabe que x y y están relacionados por una ecuación de la forma $y = kx^r$, donde r es el orden de reacción de oxígeno. Realice las transformaciones apropiadas para expresar esto último como una ecuación lineal.
- b) Estime los valores de k y r calculando la recta de mínimos cuadrados.
- c) Con base en estos datos, ¿es creíble que el orden de reacción del oxígeno es igual a 0.5? Explique.
10. Una científica especialista en materiales está experimentando con un nuevo material para fabricar latas de bebidas. Llena latas con líquido a temperatura ambiente, y después las refrigera para ver qué tan rápido se enfrían. De acuerdo con la ley de enfriamiento de Newton, si t es el tiempo de refrigeración y y la disminución de la temperatura al tiempo t , entonces y está relacionado con t por una ecuación de la forma

$$\ln y = \beta_0 + \beta_1 t,$$

donde β_0 representa una constante que depende de la temperatura inicial de la lata y de la temperatura ambiente del refrigerador, y β_1 , una constante que depende de las propiedades físicas de la lata. La científica mide la temperatura a intervalos regulares, y después ajusta este modelo a los datos. Los resultados se muestran en la siguiente figura. Un diagrama de dispersión, con la recta de mínimos cuadrados sobrepuesta, está a la izquierda, y la gráfica de residuos, a la derecha.



¿Qué debe hacer después la científica?

- Tratar de encontrar una transformación que haga la relación más lineal.
- Utilizar el modelo tal como está, ya que la ley del enfriamiento de Newton es una ley física.
- Utilizar el modelo tal como está, porque ajusta bastante bien.
- Revisar cuidadosamente la configuración experimental para ver qué puede estar mal.

11. La supervisión de la producción de una reacción química dada a diferentes temperaturas del recipiente de reacción produce los resultados que se muestran en tabla siguiente.

Temperatura (°C)	Producción (%)	Temperatura (°C)	Producción (%)
150	77.4	250	88.9
150	76.7	250	89.2
150	78.2	250	89.7
200	84.1	300	94.8
200	84.5	300	94.7
200	83.7	300	95.9

- Determine los estimadores de mínimos cuadrados para β_0 , β_1 y σ^2 para el modelo lineal simple $= \beta_0 + \beta_1 \text{Temperatura} + \varepsilon$.
- ¿Puede concluir que β_0 no es igual a 0?
- ¿Puede concluir que β_1 no es igual a 0?
- Realice una gráfica de residuos. ¿El modelo lineal parece adecuado?
- Determine un intervalo de confianza de 95% para la pendiente.
- Determine un intervalo de confianza de 95% para la media de la producción a 225°C.
- Determine un intervalo de predicción de 95% para una producción a 225°C.

12. El artículo "Approach to Confidence Interval Estimation for Curve Numbers" (R. McCuen, en *Journal of Hydrologic Engineering*, 2002:43-48) analiza la relación entre la profundidad de precipitación y la profundidad de corrientes en algunos lugares. En un lugar especial se registraron la profundidad de precipitación y la de corrientes para 13 lluvias torrenciales. El siguiente corresponde al resultado de MINITAB para un ajuste de la recta de mínimos cuadrados para predecir la profundidad de las corrientes de la profundidad de precipitación (ambas medidas en pulgadas).

The regression equation is

Runoff = -0.23 + 0.73 Rainfall

Predictor	Coef	SE Coef	T	P
Constant	-0.23429	0.23996	-0.98	0.350
Rainfall	0.72868	0.06353	11.47	0.000

S = 0.40229 R-Sq = 92.3% R-Sq(adj) = 91.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	21.290	21.290	131.55	0.000
Residual Error	11	1.780	0.16184		
Total	12	23.070			

- Pronostique la corriente para una tormenta con 2.5 pulg de precipitación.
- Alguien afirma que si dos tormentas difieren en su precipitación en 1 pulg, entonces sus corrientes serán diferentes, también en promedio por 1 pulg. ¿Esta afirmación es creíble? Explique.
- Es un hecho que si la precipitación es 0 la corriente es 0. ¿La recta de mínimos cuadrados es consistente con este hecho? Explique.

13. Con referencia al ejercicio 12, alguien quiere calcular un intervalo de confianza de 95% para la media de la corriente cuando la precipitación es 3. ¿Se puede calcular a partir de la información en el resultado de MINITAB que se muestra en el ejercicio 12? ¿O se necesita más información? Elija la mejor respuesta.

- Sí, se puede calcular del resultado de MINITAB.
- No, también se necesita conocer los valores de la precipitación que se utilizaron para calcular la recta de mínimos cuadrados.
- No, también se necesita conocer los valores de la corriente que se utilizaron para calcular la recta de mínimos cuadrados.
- No, también se necesita conocer tanto los valores de la precipitación como de la corriente que se utilizaron para calcular la recta de mínimos cuadrados.

14. Durante la producción de un plato de caldera, las piezas de prueba están sujetas a una carga, y se miden sus elongaciones. En un experimento especial, se realizarán cinco pruebas, con cargas (en MPa) de 11, 37, 54, 70 y 93. Se calculará la recta de cuadrados mínimos para predecir la elongación de la carga. Se calcularán los intervalos de confianza para la media de la elongación respecto a algunas cargas diferentes. ¿Cuál de los siguientes intervalos será el más ancho? ¿Cuál será el más angosto?

- El intervalo de confianza de 95% para la media de la elongación bajo una carga de 53 MPa.
- El intervalo de confianza de 95% para la media de la elongación bajo una carga de 72 MPa.
- El intervalo de confianza de 95% para la media de la elongación bajo una carga de 35 MPa.

15. El artículo “Low-Temperature Heat Capacity and Thermodynamic Properties of 1,1,1-trifluoro-2,2-dichloroethane” (R. Varushchenko y A. Druzhinina, en *Fluid Phase Equilibria*, 2002: 109-119) describe un experimento en el cual muestras de freón R-123 fueron derretidas en un calorímetro. Se proporcionaron diferentes cantidades de energía al calorímetro para derretirlas. Se midieron las temperaturas de equilibrio de fusión (t) y las fracciones derretidas (f). La recta de cuadrados mínimos se ajustó con el modelo $t = \beta_0 + \beta_1(1/f) + \varepsilon$, donde $1/f$ es el recíproco de la fracción. Los resultados del ajuste son los siguientes.

The regression equation is

Temperature = 145.74 - 0.052 Reciprocal Frac

Predictor	Coef	SE Coef	T	P
Constant	145.736	0.00848	17190.1	0.000
Recip Frac	-0.05180	0.00226	-22.906	0.000

S = 0.019516

R-Sq = 97.6%

R-Sq(adj) = 97.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.200	0.200	524.70	0.000
Residual Error	13	0.00495	0.000381		
Total	14	0.205			

- Estime la temperatura con la cual la mitad de la muestra se ha derretido (es decir, $f = 1/2$).
- ¿Puede determinar el coeficiente de correlación entre la temperatura de equilibrio y el recíproco de la fracción derretida en este resultado? Si es así, determínelo. Si no, explique qué información adicional se necesita.
- La temperatura de punto triple es la más baja a la cual toda la muestra se derrite (es decir, $f = 1$). Calcule la temperatura del punto triple.

16. El artículo “Polyhedral Distortions in Tourmaline” (A. Ertl, J. Hughes y cols., en *The Canadian Mineralogist*, 2002: 153-162) presenta un modelo para calcular la distorsión que soporta la longitud de enlace turmalina-vanadio. Para comprobar la exactitud del modelo se calcularon diferentes valores (x) que se compararon directamente con los valores observados (y). Los resultados (leídos de una gráfica) se presentan en la tabla siguiente.

Valor observado	Valor calculado	Valor observado	Valor calculado
0.33	0.36	0.74	0.78
0.36	0.36	0.79	0.86
0.54	0.58	0.97	0.97
0.56	0.64	1.03	1.11
0.66	0.64	1.10	1.06
0.66	0.67	1.13	1.08
0.74	0.58	1.14	1.17

- a) Suponga que el valor observado y es una medición sin sesgo del valor verdadero. Muestre que si el valor calculado x es exacto (es decir, es igual al valor verdadero), entonces $y = x + \mathcal{E}$, donde $\mathcal{E}$ es el error de medición.
- b) Calcule la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$.
- c) Muestre que si los valores calculados son precisos, entonces los coeficientes verdaderos son $\beta_0 = 0$ y $\beta_1 = 1$.
- d) Pruebe la hipótesis nula $\beta_0 = 0$ y $\beta_1 = 1$.
- e) ¿Es creíble que el valor calculado es exacto? ¿O puede concluir que no lo es? Explique.
17. Considere el modelo $y = \beta x + \mathcal{E}$, donde se sabe que el intercepto de la recta es cero. Suponga que se observan los valores $(x_1, y_1), \dots, (x_n, y_n)$, y se calcula el estimador de mínimos cuadrados $\hat{\beta}$ de β .
- a) Deduzca el estimador de mínimos cuadrados $\hat{\beta}$ en función de x_i y y_i .
- b) Sea σ^2 la varianza de $\mathcal{E}$ (que es también la varianza de y). Deduzca el estimador de la varianza $\sigma_{\hat{\beta}}^2$ de los mínimos cuadrados en función de σ^2 y de las x_i .
18. Utilice la ecuación (7.33) (pág. 510) para mostrar que $\mu_{\hat{\beta}_1} = \beta_1$.
19. Utilice la ecuación (7.34) (pág. 510) para mostrar que $\mu_{\hat{\beta}_0} = \beta_0$.
20. Utilice la ecuación (7.33) (pág. 510) para deducir la fórmula $\sigma_{\hat{\beta}_1}^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$.
21. Utilice la ecuación (7.34) (pág. 510) para deducir la fórmula $\sigma_{\hat{\beta}_0}^2 = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)$.

Capítulo 8

Regresión múltiple

Introducción

Los métodos de regresión lineal simple, analizados en el capítulo 7, son aplicables cuando se desea ajustar un modelo lineal al relacionar el valor de una variable independiente y con el valor de una sola variable dependiente x . Sin embargo, hay muchos casos en los que una sola variable independiente no es suficiente. Por ejemplo, el grado de desgaste de un cojinete lubricado en una máquina puede depender tanto de la carga en el cojinete como de las propiedades físicas del lubricante. Una ecuación que exprese el desgaste como una función sólo de la carga o de las propiedades del lubricante fallará como predictor. En situaciones como éstas hay varias variables independientes, $x_1, x_2, \dots, x_p$, relacionadas con una variable dependiente y . Si la relación entre las variables dependiente e independiente es lineal, se puede usar la técnica de **regresión múltiple**.

8.1 El modelo de regresión múltiple

A continuación se describirá el modelo de regresión múltiple. Suponga que se tiene una muestra de n elementos, y para cada uno se ha medido una variable dependiente y y p variables independientes $x_1, \dots, x_p$. El i -ésimo elemento de la muestra, por tanto, tiene el conjunto ordenado $(y_i, x_{1i}, \dots, x_{pi})$. En consecuencia, se puede ajustar al **modelo de regresión múltiple**

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i \quad (8.1)$$

Hay algunos casos especiales del modelo de regresión múltiple (8.1) que con frecuencia se utilizan en la práctica. Uno es el **modelo de regresión polinomial**, en el cual las variables independientes son potencias de una sola variable. El modelo de regresión polinomial de grado p es

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_p x_i^p + \varepsilon_i \quad (8.2)$$

Los modelos de regresión múltiple también se pueden hacer con potencias de diversas variables. Por ejemplo, un modelo de regresión polinomial de grado 2, también llamado **modelo cuadrático**, en dos variables x_1 y x_2 está dado por

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \beta_4 x_{1i}^2 + \beta_5 x_{2i}^2 + \varepsilon_i \quad (8.3)$$

Una variable producto de las otras dos variables es llama **interacción**. En el modelo (8.3), la variable $x_{1i} x_{2i}$ es la **interacción** entre x_1 y x_2 .

Los modelos (8.2) y (8.3) se consideran lineales, aunque contengan términos no lineales de las variables independientes. La razón de que continúen siendo modelos lineales es que *son lineales en los coeficientes β_i* .

Estimando los coeficientes

En cualquier modelo de regresión múltiple, los estimadores $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, se calculan por medio de mínimos cuadrados, exactamente de la misma forma como en la regresión lineal simple. La ecuación

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_p x_p \quad (8.4)$$

se llama **ecuación de mínimos cuadrados** o **ecuación de regresión ajustada**. Ahora se define $\hat{y}_i$ como la coordenada y de la ecuación de mínimos cuadrados correspondiente a los valores $x(x_{1i}, \dots, x_{pi})$. Los residuos representan las cantidades $e_i = y_i - \hat{y}_i$, que constituyen las diferencias entre los valores observados y y los valores y que proporciona la ecuación. Se quieren calcular $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, para minimizar la suma de los cuadrados residuales $\sum_{i=1}^n e_i^2$. Con ese propósito se expresa e_i como función de $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$:

$$e_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_p x_{pi} \quad (8.5)$$

Por consiguiente, se desea minimizar la suma

$$\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_p x_{pi})^2 \quad (8.6)$$

Hacer esto último requiere tomar derivadas parciales de (8.6) con respecto a $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, e igualarlas a 0, así como resolver las $p + 1$ ecuaciones resultantes con $p + 1$ incógnitas. Las expresiones que se obtienen para $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$, son complicadas. Afortunadamente, han sido codificadas en muchos software para que pueda calcularlos en la computadora. Para cada coeficiente estimado $\hat{\beta}_i$, hay una desviación estándar estimada $s_{\hat{\beta}_i}$. También las expresiones para estas cantidades son complicadas, por lo que en la actualidad las personas dependen de las computadoras para calcularlas.

Sumas de cuadrados

Gran parte del análisis en la regresión múltiple se basa en tres cantidades fundamentales. Éstas son **suma de los cuadrados de la regresión** (SSR, por sus siglas en inglés), **suma de los cuadrados del error** (SSE, por sus siglas en inglés), y **suma total de los cuadrados** (SST,

por sus siglas en inglés). Estas cantidades se definieron en la sección 7.2, en el análisis de la regresión lineal simple. Las definiciones también valen para la regresión múltiple. Aquí se repiten.

Definición: Sumas de los cuadrados

En el modelo de regresión múltiple

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon_i,$$

se definen las siguientes sumas de cuadrados:

- Suma de los cuadrados de la regresión: $SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$
- Suma de los cuadrados del error: $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$
- Suma total de los cuadrados: $SST = \sum_{i=1}^n (y_i - \bar{y})^2$

Se puede mostrar que

$$SST = SSR + SSE \quad (8.7)$$

La ecuación (8.7) se llama **identidad del análisis de la varianza**. Ésta se obtiene para la regresión lineal simple al final de la sección 7.2.

Ahora se verá cómo se utilizan estas sumas de los cuadrados para deducir los estadísticos que se emplean en la regresión múltiple. Como se hizo en la regresión lineal simple, el análisis se restringirá al caso más simple, en el cual se satisfacen los cuatro supuestos acerca de los errores ε_i . Aquí se repiten estos supuestos.

Supuestos para los errores en modelos lineales

En la situación más simple se satisfacen los siguientes supuestos:

1. Los errores $\varepsilon_1, \dots, \varepsilon_n$ son aleatorios e independientes. En particular, la magnitud de cualquier error ε_i no influye en el valor del siguiente error ε_{i+1} .
2. Los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen media 0.
3. Los errores $\varepsilon_1, \dots, \varepsilon_n$ tienen la misma varianza, que se denota por medio de σ^2 .
4. Los errores $\varepsilon_1, \dots, \varepsilon_n$ están distribuidos normalmente.

Al igual que en la regresión lineal simple, estos supuestos implican que las observaciones y_i sean variables aleatorias independientes. Para ser más específico, cada y_i tiene una distribución normal con media $\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}$ y varianza σ^2 . Cada coeficiente β_i representa el cambio en la media de y relacionado con un aumento de una unidad en el valor de x_i , cuando las otras variables x se mantienen constantes.

Resumen

En el modelo de regresión múltiple $y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon_i$, bajo los supuestos 1 a 4, las observaciones $y_1, \dots, y_n$ son variables aleatorias independientes que siguen la distribución normal. La media y la varianza de y_i están dadas por

$$\mu_{y_i} = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}$$

$$\sigma_{y_i}^2 = \sigma^2$$

Cada coeficiente β_i representa el cambio en la media de y relacionada con un aumento de una unidad en el valor de x_i , cuando las otras variables x se mantienen constantes.

Los estadísticos s^2 , R^2 y F

Los tres estadísticos más utilizados en la regresión múltiple son el estimador de la varianza del error s^2 , el coeficiente de determinación R^2 y el estadístico F . Cada uno de éstos tiene un análogo en la regresión lineal simple. En su momento se analizará a cada uno.

En la regresión lineal simple, el estimador de la varianza de error es $\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - 2)$. Se divide entre $n - 2$ en lugar de n porque los residuos ($e_i = y_i - \hat{y}_i$) tienden a ser un poco más pequeños que los errores ε_i . La razón de que los residuos sean un poco más pequeños que los dos coeficientes (β_0 y β_1) es que se han elegido para minimizar $\sum_{i=1}^n (y_i - \hat{y}_i)^2$. Ahora, para el caso de la regresión múltiple, se están estimando $p + 1$ coeficientes en lugar de sólo dos. Por tanto, los residuos tienden a ser aún más pequeños, por lo que se debe dividir $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ entre un denominador aún más pequeño. En este sentido, el denominador adecuado es igual al número de las observaciones (n) menos el número de parámetros en el modelo ($p + 1$). Por tanto, el estimador de la varianza del error está dado por

$$s^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p - 1} = \frac{\text{SSE}}{n - p - 1} \quad (8.8)$$

El estimador de la varianza $s_{\hat{\beta}_i}^2$ de cada coeficiente de mínimos cuadrados $\hat{\beta}_i$ se calcula al multiplicar s^2 por una función bastante complicada de las variables x_{ij} . En la práctica, los valores de $s_{\hat{\beta}_i}^2$ se calculan en una computadora. Cuando se satisfacen los supuestos 1 a 4, la cantidad

$$\frac{\hat{\beta}_i - \beta_i}{s_{\hat{\beta}_i}}$$

tiene una distribución t de Student con $n - p - 1$ grados de libertad. El número de grados de libertad es igual al denominador con que se calculó el estimador de varianza del error s^2 (ecuación 8.8). Con este estadístico se calculan los intervalos de confianza y se realizan pruebas de hipótesis sobre los valores β_i , de igual forma que en la regresión lineal simple.

En la regresión lineal simple, el coeficiente de determinación, r^2 , mide la bondad del ajuste del modelo lineal. El estadístico de bondad del ajuste en la regresión múltiple represen-

ta una cantidad que se denota mediante R^2 , también denominado coeficiente de determinación. El valor de R^2 se calcula del mismo modo que r^2 en la regresión lineal simple (ecuación 7.21, en la sección 7.2). Éste es

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{SST - SSE}{SST} = \frac{SSR}{SST} \quad (8.9)$$

En la regresión lineal simple casi siempre se prueba la hipótesis nula $\beta_1 = 0$. Si ésta se aprueba, entonces el modelo lineal no podría ser útil. La hipótesis nula análoga en la regresión múltiple es $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$. Ésta es una hipótesis muy fuerte. Establece que ninguna de las variables independientes tiene alguna relación lineal con la variable dependiente. En la práctica, los datos por lo general proporcionan evidencia suficiente para rechazar esta hipótesis. El estadístico de prueba para esta hipótesis es

$$F = \frac{[\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2] / p}{[\sum_{i=1}^n (y_i - \hat{y}_i)^2] / (n - p - 1)} = \frac{[SST - SSE] / p}{SSE / (n - p - 1)} = \frac{SSR / p}{SSE / (n - p - 1)} \quad (8.10)$$

Éste es un estadístico F ; su distribución nula es $F_{p, n-p-1}$. Observe que el denominador del estadístico F es s^2 (ecuación 8.8). Los subíndices p y $n - p - 1$ son los **grados de libertad** del estadístico F .

Las diferentes versiones del estadístico F se pueden utilizar para probar la hipótesis nula más débil. En particular, dado un modelo con variables independientes $x_1, \dots, x_p$, a veces se quiere probar la hipótesis nula de que alguna de ellas (por ejemplo, $x_{k+1}, \dots, x_p$) no se relaciona linealmente con la variable dependiente. Con este propósito se puede construir una versión del estadístico F para probar la hipótesis nula $H_0 : \beta_{k+1} = \dots = \beta_p = 0$. Lo anterior se analizará con mayor profundidad en la sección 8.3.

Un ejemplo

Ahora se analizará un ejemplo de regresión múltiple. Primero se describen los datos. Una red de computadoras móviles *ad hoc* consta de varias computadoras (nodos) que se mueven dentro de un área de la red. Con frecuencia los mensajes se envían de uno a otro nodo. Cuando el nodo receptor está fuera del alcance, se debe enviar el mensaje a un nodo cercano, que después lo envía hacia su destino a lo largo de una ruta de direccionamiento. Ésta se determina por medio de una rutina conocida como un protocolo de direccionamiento. El porcentaje de mensajes que se distribuye exitosamente se llama *goodput* (caudal útil), y lo afecta la velocidad promedio del nodo y la pausa en los nodos en cada destino. La tabla 8.1 presenta la velocidad de nodo promedio, el promedio de pausa y el *goodput* para 25 redes *ad hoc* móviles simuladas. Estos datos se generaron para un estudio descrito en el artículo “Metrics to Enable Adaptive Protocols for Mobile Ad Hoc Networks” (J. Boleng, W. Navidi y T. Camp, en *Proceedings of the 2002 International Conference on Wireless Networks*, 2002:293-298).

TABLA 8.1 Velocidad de nodo promedio, tiempo de pausa y *goodput* para redes de computadoras

Velocidad (m/s)	Tiempo de pausa (s)	Goodput (%)	Velocidad (m/s)	Tiempo de pausa (s)	Goodput (%)
5	10	95.111	20	40	87.800
5	20	94.577	20	50	89.941
5	30	94.734	30	10	62.963
5	40	94.317	30	20	76.126
5	50	94.644	30	30	84.855
10	10	90.800	30	40	87.694
10	20	90.183	30	50	90.556
10	30	91.341	40	10	55.298
10	40	91.321	40	20	78.262
10	50	92.104	40	30	84.624
20	10	72.422	40	40	87.078
20	20	82.089	40	50	90.101
20	30	84.937			

El siguiente resultado de MINITAB presenta los resultados del ajuste al modelo

$$\text{Goodput} = \beta_0 + \beta_1 \text{Velocidad} + \beta_2 \text{Pausa} + \beta_3 \text{Velocidad} \cdot \text{Pausa} + \beta_4 \text{Velocidad}^2 + \beta_5 \text{Pausa}^2 + \varepsilon$$

The regression equation is

$$\text{Goodput} = 96.0 - 1.82 \text{ Speed} + 0.565 \text{ Pause} + 0.0247 \text{ Speed*Pause} + 0.0140 \text{ Speed}^2 - 0.0118 \text{ Pause}^2$$

Predictor	Coef	SE Coef	T	P
Constant	96.024	3.946	24.34	0.000
Speed	-1.8245	0.2376	-7.68	0.000
Pause	0.5652	0.2256	2.51	0.022
Speed*Pa	0.024731	0.003249	7.61	0.000
Speed^2	0.014020	0.004745	2.95	0.008
Pause^2	-0.011793	0.003516	-3.35	0.003

S = 2.942

R-Sq = 93.2%

R-Sq(adj) = 91.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	2240.49	448.10	51.77	0.000
Residual Error	19	164.46	8.66		
Total	24	2404.95			

Predicted Values for New Observations

New				
Obs	Fit	SE Fit	95% CI	95% PI
1	74.272	1.175	(71.812, 76.732)	(67.641, 80.903)

Values of Predictors for New Observations

New					
Obs	Speed	Pause	Speed*Pause	Speed^2	Pause^2
1	25.0	15.0	375	625	225

Gran parte del resultado es análogo al de la regresión lineal simple. La ecuación de regresión ajustada se presenta cerca de la parte superior del resultado. Debajo de ésta se muestran el estimador del coeficiente $\hat{\beta}_i$ y las estimaciones de sus desviaciones estándar $s_{\hat{\beta}_i}$. Después de cada desviación estándar está el estadístico t de Student para probar la hipótesis nula de que el valor verdadero del coeficiente es igual a 0. Este estadístico es igual al cociente del estimador del coeficiente y su desviación estándar. Debido a que hay $n = 25$ observaciones y $p = 5$ variables independientes, el número de grados de libertad para el estadístico t de Student es $25 - 5 - 1 = 19$. Los P -valores para las pruebas se presentan en la columna siguiente. Todos los P -valores son pequeños, por lo que sería razonable concluir que cada una de las variables independientes en el modelo es útil para pronosticar el *goodput*.

La cantidad “S” es s , la estimación de la desviación estándar del error, y “R-sq” representa el coeficiente de determinación de R^2 . El ajuste R^2 , “R-sq(adj)”, se utiliza principalmente en la selección de modelo. En la sección 8.3 se analizará este estadístico.

El análisis de la tabla de varianza es análogo al que se encuentra en la regresión lineal simple. Se recorrerá columna por columna. En la columna correspondiente, “DF”, los grados de libertad para la regresión son igual al número de variables independientes (5). Observe que Speed^2 (Velocidad²), Pause^2 (Pausa²), y $\text{Speed} \cdot \text{Pause}$ (Velocidad $\cdot$ Pausa) se consideran como variables independientes distintas, aun cuando se pueden calcular a partir de Speed (Velocidad) y Pause (Pausa). En el siguiente renglón, etiquetado como “Residual Error”, se encuentra que el número de grados de libertad es 19, que representa el número de observaciones (25) menos el número de los parámetros estimados (6: el intercepto, y los coeficientes para cinco variables independientes). Por último, el “Total” de grados de libertad es uno menos que el tamaño de muestra de 25. Observe que el total de grados de libertad constituye la suma de los grados de libertad de regresión más los grados de libertad para el error. Yendo hacia abajo en la columna “SS”, se encuentra la suma de los cuadrados de la regresión SSR, la suma de los cuadrados del error SSE, y la suma total de los cuadrados SST. Note que $\text{SST} = \text{SSR} + \text{SSE}$. La columna “MS” presenta la **media de los cuadrados**, que son las sumas de cuadrados divididas entre sus respectivos grados de libertad. Observe que la media del cuadrado del error es igual a s^2 , la estimación para la varianza de error: ($s^2 = S^2 = 2.942^2 = 8.66$). La columna etiquetada como “F” presenta los cuadrados medios para la regresión dividida entre los cuadrados medios para el error ($448.10/8.66 = 51.77$, lo que permite el error de redondeo). Éste es el estadístico F que se muestra en la ecuación (8.10), y con el que se prueba la hipótesis nula de que ninguna de las variables independientes están relacionadas linealmente con las variables dependientes. El P -valor para esta prueba es 0.

El resultado debajo del encabezado “Predicted Values for New Observations” presenta los intervalos de confianza de la respuesta media y los intervalos de predicción para los valo-

res de las variables dependientes que haya especificado el usuario. Los valores de las variables dependientes que se han especificado están listadas debajo del encabezado “Values of Predictors for New Observations”. Los valores de las variables independientes en este resultado son Speed = 25 y Pause = 15. La cantidad 74.242, etiquetada como “Fit”, es el valor de $\hat{y}$, obtenido al sustituir estos valores en la ecuación de regresión ajustada. La cantidad etiquetada “SE Fit” representa el estimador de la desviación estándar de $\hat{y}$, que se utiliza para calcular el intervalo de confianza de 95%, etiquetado como “95% CI”. La cantidad “95% PI” significa el intervalo de predicción de 95% para una futura observación de la variable dependiente cuando se establecen las variables independientes para algunos valores dados. De la misma manera que con el intervalo de confianza, este intervalo se centra en $\hat{y}$, pero es más ancho, igual que en la regresión lineal simple.

Ejemplo

8.1

Utilice el modelo de la regresión múltiple con el propósito de pronosticar el *goodput* para una red con velocidad de 12 m/s y tiempo de pausa de 25 s.

Solución

Del resultado de MINITAB, el modelo ajustado es

$$\text{Goodput} = 96.0 - 1.82 \text{ Velocidad} + 0.565 \text{ Pausa} + 0.0247 \text{ Velocidad} \cdot \text{Pausa} \\ + 0.0140 \text{ Velocidad}^2 - 0.0118 \text{ Pausa}^2$$

Al sustituir 12 para velocidad y 25 para pausa, se encuentra que el *goodput* pronosticado es 90.336.

Ejemplo

8.2

Para los datos *goodput*, determine el residuo para el punto Velocidad = 20, Pausa = 30.

Solución

El valor observado *goodput* (tabla 8.1) es $y = 84.937$. El valor pronosticado $\hat{y}$ se encuentra sustituyendo Velocidad = 20 y Pausa = 30 en el modelo ajustado que se presenta en la solución del ejemplo 8.1. Se obtiene un valor pronosticado para *goodput* de $\hat{y} = 86.350$. El residuo está dado por $y - \hat{y} = 84.937 - 86.350 = -1.413$.

Es sencillo calcular, mediante el resultado computacional, los intervalos de confianza y probar las hipótesis considerando los coeficientes de mínimos cuadrados. Los ejemplos 8.3 a 8.5 proporcionan casos ilustrativos.

Ejemplo

8.3

Determine un intervalo de confianza de 95% para el coeficiente de la Velocidad en el modelo de regresión múltiple.

Solución

Del resultado, el coeficiente estimado es -1.8245 , con desviación estándar de 0.2376. Un intervalo de confianza se determina utilizando la distribución t de Student con 19 grados de li-

bertad. Los grados de libertad para el estadístico t son iguales a los de libertad del error. El valor t para un intervalo de confianza de 95% es $t_{19, .025} = 2.093$. El intervalo de confianza de 95% es

$$-1.8245 \pm (2.093)(0.2376) = -1.8245 \pm 0.4973 = (-2.3218, -1.3272)$$

Ejemplo

8.4

Pruebe la hipótesis nula de que el coeficiente de Pausa es menor o igual que 0.3.

Solución

El estimador del coeficiente de Pausa es $\hat{\beta}_2 = 0.5652$, con desviación estándar $s_{\hat{\beta}_2} = 0.2256$. La hipótesis nula es $\beta_2 \leq 0.3$. Bajo H_0 se tiene que $\beta_2 = 0.3$, por lo que la cantidad

$$t = \frac{\hat{\beta}_2 - 0.3}{0.2256}$$

tiene una distribución t de Student con 19 grados de libertad. Observe que el concepto grados de libertad para el estadístico t es igual a los grados de libertad para el error. El valor del estadístico t es $(0.5652 - 0.3)/0.2256 = 1.1755$. El P -valor se encuentra entre 0.10 y 0.25. Es creíble que $\beta_2 \leq 0.3$.

Ejemplo

8.5

Determine un intervalo de confianza de 95% para la respuesta media μ_y , y un intervalo de predicción de 95% para una futura observación cuando Velocidad = 25 y Pausa = 15.

Solución

Del resultado, bajo el encabezado “Predicted Values for New Observations”, el intervalo de confianza de 95% es (71.812, 76.732) y el intervalo de predicción de 95% es (67.641, 80.903).

Comprobación de supuestos en la regresión múltiple

En la regresión múltiple, como en la regresión lineal simple, es importante probar la validez de los supuestos para los errores en modelos lineales (presentados al principio de esta sección). Los diagnósticos para estos supuestos empleados en el caso de la regresión lineal simple también son útiles en la regresión múltiple. Estos diagnósticos son las gráficas de residuos contra valores ajustados, las de probabilidad normal de residuos y las de residuos contra el orden en que se hacen las observaciones. También es una buena idea hacer gráficas de residuos contra cada una de las variables independientes. Si las gráficas de residuos indican incumplimiento de los supuestos, es posible intentar arreglar el problema al transformar las variables, como en la regresión lineal simple.

La figura 8.1 presenta una gráfica de residuos contra valores ajustados para los datos *goodput*. Las figuras 8.2 y 8.3 presentan gráficas de residuos contra velocidad y pausa, res-

pectivamente. La gráfica de los residuos contra valores ajustados da la impresión de curvatura, que causan principalmente algunos puntos en los extremos. Las gráficas de los residuos contra variables independientes no indican ningún incumplimiento grave de los supuestos del modelo. En la práctica se podría aceptar que este modelo ajusta bastante bien, o se podrían aplicar técnicas de selección de modelos (analizadas en la sección 8.3) para analizar modelos alternativos.

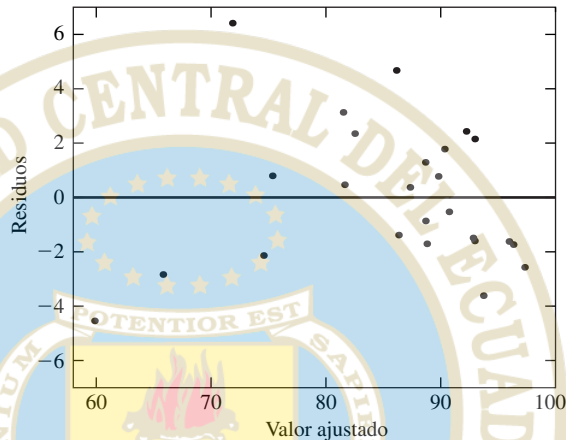


FIGURA 8.1 Gráfica de residuos contra valores ajustados para los datos *goodput*.

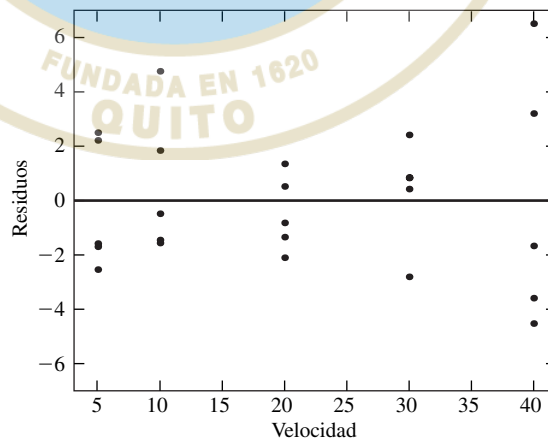


FIGURA 8.2 Gráfica de residuos contra la velocidad para los datos *goodput*.

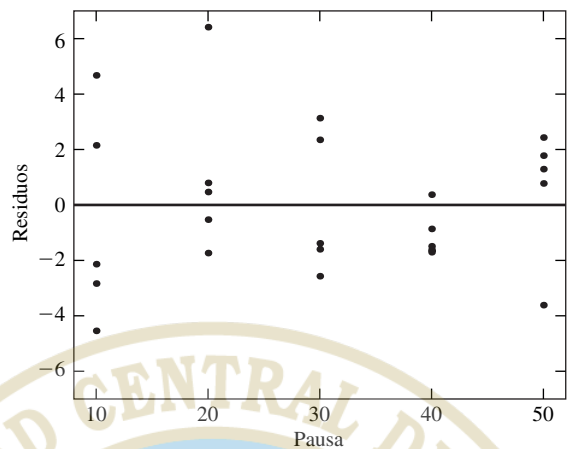


FIGURA 8.3 Gráfica de residuos contra pausa para los datos *goodput*.

Ejercicios para la sección 8.1

1. En un experimento para determinar los factores que afectan el ahorro de combustible en camiones se midió el consumo de combustible (mi/gal), el peso (toneladas) y la lectura de odómetro (miles de millas) en 15 camiones. El siguiente resultado de MINITAB presenta los resultados del ajuste del modelo millas por galón = $\beta_0 + \beta_1$ peso + β_2 odómetro.

The regression equation is
MPG = 8.24 - 0.108 Weight - 0.00392 Odometer

Predictor	Coef	SE Coef	T	P
Constant	8.2407	0.2871	28.70	0.000
Weight	-0.10826	0.01194	-9.06	0.000
Odometer	-0.0039249	0.001406	-2.79	0.016

S = 0.3182 R-Sq = 87.8% R-Sq(adj) = 85.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	8.720	4.360	43.05	0.000
Residual Error	12	1.215	0.101		
Total	14	9.935			

- Pronostique las millas por galón para un camión que pesa diez toneladas y tiene una lectura del odómetro de 50 000 millas.
- Si los dos camiones tienen el mismo peso y uno tiene 10 000 millas más en el odómetro, ¿qué tanto se predice que difieran sus millas por galón?
- Si dos camiones tienen la misma lectura del odómetro, y uno pesa cinco toneladas más que el otro, ¿qué tanto se predice que difieran sus millas por galón?

2. Considerando el ejercicio 1.

- Determine un intervalo de confianza de 95% para el coeficiente del peso.
- Determine un intervalo de confianza de 99% para el coeficiente del odómetro.
- ¿Puede concluir que $\beta_1 < -0.05$? Realice la prueba de hipótesis adecuada.
- ¿Puede concluir que $\beta_2 > -0.005$? Realice la prueba de hipótesis adecuada.

3. En la tabla siguiente se presentan los datos con que se ajusta el modelo del ejercicio 1, junto con los residuos y los valores ajustados. Grafique los residuos contra los valores ajustados. ¿La gráfica indica que ese modelo lineal es razonable? Explique.

MPG	Peso	Odometro	Residuo	Valor ajustado
7.28	10.5	15	0.235	7.045
5.63	23.0	71	0.158	5.472
5.26	27.5	36	0.138	5.122
6.58	14.5	113	0.353	6.227
5.01	30.5	39	0.224	4.786
6.73	14.0	97	0.386	6.344
5.37	21.0	195	0.168	5.202
7.28	8.5	8	-0.009	7.289
4.85	26.0	84	-0.246	5.096
5.08	26.5	25	-0.194	5.274
5.51	15.0	124	-0.620	6.130
4.75	30.0	25	-0.145	4.895
6.03	15.0	75	-0.292	6.322
5.26	22.5	192	0.209	5.051
5.60	16.0	139	-0.363	5.963

4. El artículo “Application of Analysis of Variance to Wet Clutch Engagement” (M. Mansouri, M. Khonsari y cols., en *Proceedings of the Institution of Mechanical Engineers*, 2002:117-125) presenta el siguiente modelo ajustado para pronosticar el tiempo que se presiona el clutch en segundos (y) a partir de la velocidad en que se inicia la presión en m/s (x_1), el momento de torsión máximo de conducción en $\text{N} \cdot \text{m}$ (x_2), la inercia del sistema en $\text{kg} \cdot \text{m}^2$ (x_3), y la tasa de la fuerza aplicada en kN/s (x_4):

$$y = -0.83 + 0.017x_1 + 0.0895x_2 + 42.771x_3 + 0.027x_4 - 0.0043x_2x_4$$

La suma de los cuadrados de la regresión fue $\text{SSR} = 1.08613$ y la suma de los cuadrados del error $\text{SSE} = 0.036310$. Había 44 grados de libertad del error.

- Pronostique el tiempo de presión del clutch cuando la velocidad inicial es 20 m/s, el momento de torsión máximo de conducción es $17 \text{ N} \cdot \text{m}$, la inercia de sistema es $0.006 \text{ kg} \cdot \text{m}^2$, y la tasa de la fuerza aplicada es 10 kN/s.
- ¿Es posible pronosticar el cambio en el tiempo de presión relacionado con un aumento de 2 m/s en la velocidad inicial? Si es así, encuentre el cambio pronosticado. Si no, explique.

- c) ¿Es posible predecir el cambio en el tiempo de presión relacionado con un aumento de $2 \text{ N} \cdot \text{m}$ en el momento de torsión máximo? Si es así, encuentre el cambio pronosticado. Si no, explique por qué.
- d) Calcule el coeficiente de determinación R^2 .
- e) Calcule el estadístico F para probar la hipótesis nula de que todos los coeficientes son iguales a 0. ¿Esta hipótesis puede ser rechazada?
5. En el artículo “Application of Statistical Design in the Leaching Study of Low-Grade Manganese Ore Using Aqueous Sulfur Dioxide” (P. Naik, L. Sukla y S. Das, en *Separation Science and Technology*, 2002:1375-1389), un modelo ajustado para predecir la extracción de manganeso en % (y) a partir del tamaño de partícula en mm (x_1), la cantidad de dióxido de azufre en múltiplos de la cantidad estequiométrica necesaria para la disolución de manganeso (x_2) y la duración de la filtración en minutos (x_3) están dadas como

$$y = 56.145 - 9.046x_1 - 33.421x_2 + 0.243x_3 - 0.5963x_1x_2 - 0.0394x_1x_3 + 0.6022x_2x_3 \\ + 0.6901x_1^2 + 11.7244x_2^2 - 0.0097x_3^2$$

Se tomaron $n = 27$ observaciones, con $\text{SSE} = 209.55$ y $\text{SST} = 6\,777.5$.

- a) Pronostique el porcentaje de extracción cuando el tamaño de partícula es 3 mm, la cantidad de dióxido de azufre 1.5, y la duración de la filtración es de 20 minutos.
- b) ¿Es posible predecir un cambio en el porcentaje de extracción cuando la duración de la filtración aumenta en un minuto? Si es así, encuentre el cambio pronosticado. Si no, explique.
- c) Calcule el coeficiente de determinación R^2 .
- d) Calcule el estadístico F para probar la hipótesis nula de que todos los coeficientes son iguales a 0. ¿Esta hipótesis es rechazable?
6. El artículo “Earthmoving Productivity Estimation Using Linear Regression Techniques” (S. Smith, en *Journal of Construction Engineering and Management*, 1999:133-141) presenta el siguiente modelo lineal para pronosticar la productividad del movimiento de tierra (en m^3 movido por hora):

$$\text{Productividad} = -297.877 + 84.787x_1 + 36.806x_2 + 151.680x_3 - 0.081x_4 - 110.517x_5 \\ - 0.267x_6 - 0.016x_1x_4 + 0.107x_4x_5 + 0.0009448x_4x_6 - 0.244x_5x_6$$

donde x_1 = cantidad de camiones,

x_2 = cantidad de cubos por carga,

x_3 = volumen del cubo, en m^3

x_4 = longitud de arrastre, en m

x_5 = factor de acoplamiento (cociente de la capacidad de arrastre entre la capacidad de carga),

x_6 = tiempo de viaje en camión, en s

- a) Si el volumen del cubo aumenta en 1 m^3 mientras que las otras variables independientes permanecen iguales, ¿puede determinar el cambio pronosticado en la productividad? Si es así, determínelo. Si no, diga qué otra información necesitaría para determinarlo.
- b) Si la longitud de arrastre aumenta en 1 m, ¿puede determinar el cambio pronosticado en la productividad? Si es así, determínelo. Si no, diga qué otra información necesitaría para determinarlo.
7. En un estudio de la función pulmonar de niños, el volumen de aire exhalado por la fuerza en un segundo se llama FEV_1 . (FEV_1 es el volumen de espiración forzada en un segundo.) Se hicieron mediciones en un grupo de niños cada año durante dos años. Se ajustó a un modelo lineal para pronosticar los FEV_1 de estos años como una función del FEV_1 (en litros) del último año, el sexo del niño (0 = masculino, 1 = femenino), la estatura del niño (en m), y la presión atmosférica ambiental (en mm). El siguiente resultado de MINITAB presenta los resultados de ajuste del modelo

$$FEV_1 = \beta_0 + \beta_1 \text{ Último FEV}_1 + \beta_2 \text{ Sexo} + \beta_3 \text{ Estatura} + \beta_4 \text{ Presión} + \varepsilon$$

The regression equation is

$$FEV_1 = -0.219 + 0.779 \text{ Last FEV} - 0.108 \text{ Gender} + 1.354 \text{ Height} - 0.00134 \text{ Pressure}$$

Predictor	Coef	SE Coef	T	P
Constant	-0.21947	0.4503	-0.49	0.627
Last FEV	0.779	0.04909	15.87	0.000
Gender	-0.10827	0.0352	-3.08	0.002
Height	1.3536	0.2880	4.70	0.000
Pressure	-0.0013431	0.0004722	-2.84	0.005

S = 0.22039

R-Sq = 93.5%

R-Sq(adj) = 93.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	111.31	27.826	572.89	0.000
Residual Error	160	7.7716	0.048572		
Total	164	119.08			

- Pronostique el FEV_1 para un niño con estatura de 1.4 m, si la medida se tomó a presión de 730 mm y la medición del último año fue 2.113 L.
- Si dos niñas difieren en estatura por 5 cm, ¿qué tanto esperaría que sus mediciones de FEV_1 difieran; los otros conceptos siguen igual?
- Se estima que el término constante β_0 es negativo, pero el FEV_1 debe ser siempre positivo. ¿Algo está erróneo? Explique.
- El responsable de este experimento quiere rediseñar el algoritmo que registra las mediciones electrónicamente con el fin de ajustar la presión atmosférica automáticamente. Se fija un barómetro al dispositivo para registrar la presión. Utilice el resultado anterior de MINITAB para determinar cómo calcular un valor FEV_1 ajustado como función del valor FEV_1 medido y de la presión.

8. En relación con el ejercicio 7.

- Determine un intervalo de confianza de 95% para el coeficiente del último FEV.
- Determine un intervalo de confianza de 98% para el coeficiente de la estatura.
- ¿Puede concluir que $\beta_2 < -0.08$? Realice la prueba de hipótesis adecuada.
- ¿Puede concluir que $\beta_3 > 0.5$? Realice la prueba de hipótesis adecuada.

- El artículo “Drying of Pulps in Sprouted Bed: Effect of Composition on Dryer Performance” (M. Medeiros, S. Rocha y cols., *Drying Technology*, 2002:865-881) presenta mediciones de pH, de viscosidad ($\text{kg/m} \cdot \text{s}$), la densidad (g/cm^3) y BRIX (%). El siguiente resultado de MINITAB presenta los resultados del modelo de ajuste

$$\text{pH} = \beta_0 + \beta_1 \text{ Viscosidad} + \beta_2 \text{ Densidad} + \beta_3 \text{ BRIX} + \varepsilon$$

The regression equation is

$$\text{pH} = -1.79 + 0.000266 \text{ Viscosity} + 9.82 \text{ Density} - 0.300 \text{ BRIX}$$

Predictor	Coef	SE Coef	T	P
Constant	-1.7914	6.2339	-0.29	0.778
Viscosity	0.00026626	0.00011517	2.31	0.034
Density	9.8184	5.7173	1.72	0.105
BRIX	-0.29982	0.099039	-3.03	0.008

S = 0.379578

R-Sq = 50.0%

R-Sq(adj) = 40.6%

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	3.0875	0.1351	(2.8010, 3.3740)	(2.2333, 3.9416)
2	3.7351	0.1483	(3.4207, 4.0496)	(2.8712, 4.5990)
3	2.8576	0.2510	(2.3255, 3.3896)	(1.8929, 3.8222)

Values of Predictors for New Observations

New Obs	Viscosity	Density	BRIX
1	1000	1.05	19.0
2	1200	1.08	18.0
3	2000	1.03	20.0

- a) Pronostique el pH para una pulpa con viscosidad de 1 500 kg/m · s, densidad de 1.04 g/cm³ y BRIX de 17.5%.
- b) Si dos pulpas difieren en densidad por 0.01 g/cm³, ¿qué tanto esperarías que difirieran en pH; los otros conceptos siguen igual?
- c) Se estima que el término constante β_0 es negativo, pero el pH de pulpa debe ser siempre positivo. ¿Algo está equivocado? Explique.
- d) Determine un intervalo de confianza de 95% para la media del pH de las pulpas con viscosidad 1 200 kg/m · s, densidad 1.08 g/cm³ y BRIX 18.0%.
- e) Determine un intervalo de predicción de 95% para el pH de una pulpa con viscosidad 1 000 kg/m · s, densidad 1.05 g/cm³ y BRIX 19.0%.
- f) La pulpa A tiene viscosidad 2 000, densidad 1.03 y BRIX 20.0. La pulpa B tiene viscosidad 1 000, densidad 1.05 y BRIX 19.0. ¿Qué pulpa tendrá su pH pronosticado con mayor precisión? Explique.
10. Una científica ha medido las cantidades, y , x_1 y x_2 . Cree que y está relacionado con x_1 y x_2 a través de la ecuación $y = \alpha e^{\beta_1 x_1 + \beta_2 x_2} \delta$, donde δ representa un error aleatorio siempre positivo. Encuentre una transformación de los datos que le permita utilizar un modelo lineal para estimar β_1 y β_2 .
11. El siguiente resultado de MINITAB es para una regresión múltiple. Algo salió mal en la impresión y faltan algunos números. Compléte los.

Predictor	Coef	SE Coef	T	P
Constant	-0.58762	0.2873	(a)	0.086
X1	1.5102	(b)	4.30	0.005
X2	(c)	0.3944	-0.62	0.560
X3	1.8233	0.3867	(d)	0.003

S = 0.869 R-Sq = 90.2% R-Sq(adj) = 85.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	41.76	(e)	(f)	0.000
Residual Error	6	(g)	0.76		
Total	(h)	46.30			

12. El siguiente resultado de MINITAB es para una regresión múltiple. Algunos de los números se mancharon y son ilegibles. Complete los números que faltan.

Predictor	Coef	SE Coef	T	P
Constant	(a)	1.4553	5.91	0.000
X1	1.2127	(b)	1.71	0.118
X2	7.8369	3.2109	(c)	0.035
X3	(d)	0.8943	-3.56	0.005

S = 0.82936 R-Sq = 78.0% R-Sq(adj) = 71.4%

Source	DF	SS	MS	F	P
Regression	(e)	(f)	8.1292	11.818	0.001
Residual Error	10	6.8784	(g)		
Total	13	(h)			

13. El artículo "Evaluating Vent Manifold Inerting Requirements: Flash Point Modeling for Organic Acid-Water Mixtures" (R. Garland y M. Malcolm, en *Process Safety Progress*, 2002:254-260) presenta un modelo para pronosticar el punto de inflamación (en °F) de una mezcla de agua, ácido acético, ácido propiónico y ácido butírico a partir de las concentraciones (en % de peso) de los tres ácidos. Los resultados son los siguientes. La variable "ácido butírico * ácido acético" es la interacción entre la concentración de ácido butírico y la de ácido acético.

Predictor	Coef	SE Coef	T	P
Constant	267.53	11.306	23.66	0.000
Acetic Acid	-1.5926	0.1295	-12.30	0.000
Propionic Acid	-1.3897	0.1260	-11.03	0.000
Butyric Acid	-1.0934	0.1164	-9.39	0.000
Butyric Acid*Acetic Acid	-0.002658	0.001145	-2.32	0.034

- a) Pronostique el punto de inflamación para una mezcla de 30% de ácido acético y 35% de ácido propiónico, y 30% de ácido butírico. (Nota: en el modelo, 30% se representa por 30, no por 0.30.)
- b) Alguien pregunta qué tanto cambiará el punto de inflamación pronosticado si la concentración de ácido acético se aumenta en 10% mientras que las otras concentraciones se conservan constantes. ¿Es posible responder esta pregunta? Si es así, respóndala. Si no, explique.
- c) Alguien pregunta qué tanto cambiará el punto de inflamación pronosticado si la concentración del ácido propiónico aumenta en 10% mientras que las otras concentraciones se mantienen constantes. ¿Es posible responder esta pregunta? Si es así, respóndala. Si no, explique.
14. En el artículo “Low-Temperature Heat Capacity and Thermodynamic Properties of 1,1,1-trifluoro-2, 2-dichloroethane” (R. Varushchenko y A. Druzhinina, en *Fluid Phase Equilibria*, 2002:109-119), la relación entre la presión de vapor (p) y la capacidad calorífica (t) está dada como $p = t^{\beta_1} \cdot e^{\beta_0 + \beta_1 t + \beta_2 t^2} \delta$, donde δ es un error aleatorio siempre positivo. Expresé esta relación como un modelo lineal usando una transformación adecuada.
15. Se reunieron los siguientes datos en un experimento para estudiar la relación entre la fuerza de corte en kPa (y) y la temperatura de endurecimiento en °C (x).

x	138	140	146	148	152	153
y	5 390	5 610	5 670	5 140	4 480	4 130

El modelo de mínimos cuadrados es $y = -291\,576.77 + 4\,168.6479x - 14.6133933x^2$.

- a) Utilizando esta ecuación, calcule los residuos.
- b) Calcule la suma de los cuadrados del error SSE y la suma total de los cuadrados SST.
- c) Calcule la estimación de la varianza del error s^2 .
- d) Calcule el coeficiente de determinación R^2 .
- e) Calcule el valor del estadístico F para la hipótesis $H_0: \beta_1 = \beta_2 = 0$. ¿Cuántos grados de libertad tiene este estadístico?
- f) ¿La hipótesis $H_0: \beta_1 = \beta_2 = 0$ puede ser rechazada al nivel de 5%? Explique.
16. Se reunieron los siguientes datos en un experimento para estudiar la relación entre el número de libras de fertilizante (x) y la cosecha de tomates en *busheles* (y).

x	5	10	15	20	25
y	16	21	27	25	21

El modelo de mínimos cuadrados es $y = 4.8000 + 2.508571x - 0.07428571x^2$.

- a) Utilizando esta ecuación, calcule los residuos.
- b) Calcule la suma de los cuadrados del error SSE y la suma total de los cuadrados SST.
- c) Calcule la estimación de la varianza del error s^2 .
- d) Calcule el coeficiente de determinación R^2 .
- e) Calcule el valor del estadístico F para la hipótesis $H_0: \beta_1 = \beta_2 = 0$. ¿Cuántos grados de libertad tiene este estadístico?
- f) ¿La hipótesis $H_0: \beta_1 = \beta_2 = 0$ puede rechazarse al nivel de 5%? Explique.
17. El 24 de noviembre de 2001, el diario *The Economist* publicó datos para 15 países industrializados. Se incluyen los cambios porcentuales del producto interno bruto (PIB), la producción industrial (PI), los precios al consumidor (PC), y los precios de producción (PP) de otoño de 2000 a otoño de 2001, y la tasa de desempleo en otoño de 2001 (DESEMPLEO). Un economista quiere construir un modelo para pronosticar el PIB a partir de las otras variables. Un ajuste del modelo

$$\text{PIB} = \beta_0 + \beta_1 \text{PI} + \beta_2 \text{DESEMPLEO} + \beta_3 \text{PC} + \beta_4 \text{PP} + \varepsilon$$

produce el siguiente resultado:

The regression equation is

$$\text{GDP} = 1.19 + 0.17 \text{ IP} + 0.18 \text{ UNEMP} + 0.18 \text{ CP} - 0.18 \text{ PP}$$

Predictor	Coef	SE Coef	T	P
Constant	1.18957	0.42180	2.82	0.018
IP	0.17326	0.041962	4.13	0.002
UNEMP	0.17918	0.045895	3.90	0.003
CP	0.17591	0.11365	1.55	0.153
PP	-0.18393	0.068808	-2.67	0.023

- Pronostique el cambio porcentual en el PIB para un país con $\text{PI} = 0.5$, $\text{DESEMPLEO} = 5.7$, $\text{PC} = 3.0$, y $\text{PP} = 4.1$.
- Si los dos países difieren en tasa de desempleo por 1%, ¿qué tanto predeciría que cambie el porcentaje del PIB; los otros conceptos siguen igual?
- PC y PP son medidas de la tasa de inflación. ¿Cuál es más útil para predecir el PIB? Explique.
- El índice de precios de producción para Suecia en septiembre de 2000 era 4.0, y para Austria, 6.0. Los otros conceptos siguen igual, ¿para qué país esperaría el mayor cambio porcentual en el PIB? Explique.

18. El artículo “Multiple Linear Regression for Lake Ice and Lake Temperature Characteristics” (S. Gao y H. Stefan, en *Journal of Cold Regions Engineering*, 1999:59-77) presenta datos del espesor máximo de hielo en mm (y), el número promedio de días por año cubiertos de hielo (x_1), y el número promedio de días en que la temperatura inferior es menor de 8°C (x_2), y la profundidad de nieve promedio en mm (x_3) para 13 lagos en Minnesota. Los datos se presentan en la tabla siguiente.

y	x_1	x_2	x_3	y	x_1	x_2	x_3
730	152	198	91	730	157	204	90
760	173	201	81	650	136	172	47
850	166	202	69	850	142	218	59
840	161	202	72	740	151	207	88
720	152	198	91	720	145	209	60
730	153	205	91	710	147	190	63
840	166	204	70				

- Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = \varepsilon$. Encuentre el P -valor de cada coeficiente para probar la hipótesis nula de que el coeficiente es igual a 0.
 - Si dos lagos difieren en dos en el número promedio de días por año cubiertos de hielo, con las otras variables permaneciendo constantes, ¿qué diferencia esperaría entre sus capas de hielo máximas?
 - ¿Los lagos con media de profundidad de nieve mayor tenderían a tener espesores de hielo máximo mayores o menores? Explique.
19. En un experimento para estimar la aceleración de un objeto hacia abajo de un plano inclinado, aquél se suelta y su distancia en metros (y) desde la parte superior del plano se mide cada 0.1 segundo a partir del tiempo $t = 0.1$ a $t = 1.0$. En la tabla siguiente se presentan los datos.

t	y
0.1	0.03
0.2	0.1
0.3	0.27
0.4	0.47
0.5	0.73
0.6	1.07
0.7	1.46
0.8	1.89
0.9	2.39
1.0	2.95

Los datos siguen el modelo cuadrático $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon$, donde β_0 representa la posición inicial del objeto, β_1 la velocidad inicial del objeto, y $\beta_2 = a/2$, donde a es la aceleración del objeto, que se supone constante. En un experimento perfecto, tanto la posición como la velocidad del objeto serían cero al tiempo 0. Sin embargo, debido al error experimental, es posible que la posición y la velocidad en $t = 0$ sean diferentes de cero.

- Ajuste el modelo cuadrático $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon$.
- Determine un intervalo de confianza de 95% para β_2 .
- Determine un intervalo de confianza de 95% para la aceleración a .
- Calcule el P -valor para cada coeficiente.
- ¿Puede concluir que la posición inicial no fue cero? Explique.
- ¿Puede concluir que la velocidad inicial no fue cero? Explique.

8.2 Confusión y colinealidad

El subtítulo de esta sección es “Ajustar modelos por separado para cada variable no es lo mismo que ajustar el modelo multivariado”. Para ilustrar de qué se está hablando, se revisan los datos del pozo de gas, descrito por vez primera en el ejercicio 15 de la sección 7.3. Un total de 255 pozos de gas recibieron “tratamiento de fractura” para incrementar la producción. En este tratamiento, el fluido fracturado, fluido mezclado con arena, se bombea al pozo. La arena abre grietas en la roca, lo que incrementa el flujo del gas. Las preguntas principales son éstas: ¿Al aumentar el volumen de fluido bombeado crece la producción del pozo? ¿Al aumentar el volumen de arena se aumenta la producción del pozo?

Los pozos de mayor profundidad producen más gas porque proporcionan más superficie a través de la cual penetra el gas. Por esa razón es adecuado expresar todas las variables en unidades por pie de profundidad del pozo. Por tanto, se mide la producción en unidades de pie³ de gas por pie de profundidad, el fluido en unidades de gal/pie, y la arena en unidades de lb/pie.

En la figura 7.17 (de la sección 7.4) se mostró que se necesitaba una transformación logaritmo para obtener homocedasticidad en la gráfica de producción contra fluido. Se indicó que también se requiere una transformación logaritmo para la variable arena. La figura 8.4 muestra los diagramas de dispersión de $\ln$ Producción contra $\ln$ Fluido y $\ln$ Producción contra $\ln$ Arena. Tanto el fluido como la arena parecen estar fuertemente relacionados con la producción.

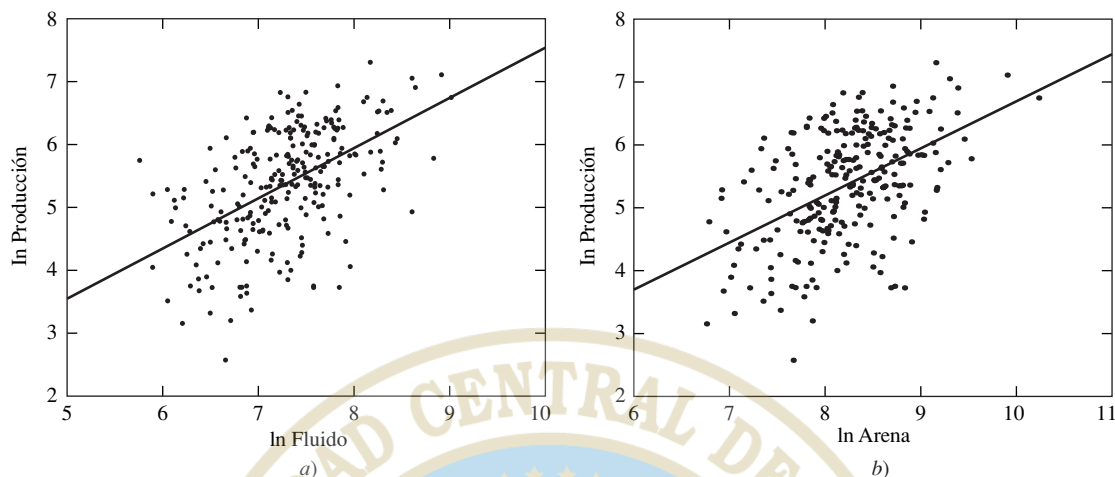


FIGURA 8.4 a) Gráfica del logaritmo de la producción contra el logaritmo del volumen de fluido fracturado para 255 pozos de gas, con la recta de mínimos cuadrados sobrepuesta. b) Gráfica del logaritmo de la producción contra el logaritmo del peso de arena para los mismos 255 pozos. Parecen existir fuertes relaciones lineales entre el logaritmo de la producción tanto con el logaritmo de fluido como con el logaritmo de arena.

Para confirmar el resultado aparente de las gráficas en la figura 8.4 se ajustan dos modelos de regresión lineal simple:

$$\ln \text{ Producción} = \beta_0 + \beta_1 \ln \text{ Fluido} + \varepsilon$$

$$\ln \text{ Producción} = \beta_0 + \beta_1 \ln \text{ Arena} + \varepsilon$$

El resultado MINITAB para estos modelos es el siguiente:

The regression equation is
 $\ln \text{ Prod} = -0.444 + 0.798 \ln \text{ Fluid}$

Predictor	Coef	SE Coef	T	P
Constant	-0.4442	0.5853	-0.76	0.449
ln Fluid	0.79833	0.08010	9.97	0.000

S = 0.7459 R-Sq = 28.2% R-Sq(adj) = 27.9%

The regression equation is
 $\ln \text{ Prod} = -0.778 + 0.748 \ln \text{ Sand}$

Predictor	Coef	SE Coef	T	P
Constant	-0.7784	0.6912	-1.13	0.261
ln Sand	0.74751	0.08381	8.92	0.000

S = 0.7678 R-Sq = 23.9% R-Sq(adj) = 23.6%

Tanto el fluido como la arena tienen coeficientes que son definitivamente diferentes de 0 (los P -valores para ambos son ≈ 0). Por tanto, se podría intentar concluir inmediatamente que al aumentar el volumen de fluido o el volumen de arena bombeada en un pozo se aumentará la producción, pero primero se debe considerar la posibilidad de confusión.

La cuestión de confusión surge de esta manera. Fluido y arena se bombean dentro juntos en una sola mezcla. Es lógico esperar que los pozos que tienen más fluido también tienden a tener más arena. Si esto último es verdadero, entonces la confusión es una posibilidad. La figura 8.5 presenta el diagrama de dispersión de $\ln$ Fluido contra $\ln$ Arena. Se tiene bastante seguridad de que la cantidad de fluido bombeada en un pozo está muy correlacionada con la cantidad de arena bombeada. Por tanto, es muy posible que cualquiera de los dos resultados univariados que se presentaron antes pudieran representar confusión en lugar de una relación real. Si la producción depende solamente del volumen de fluido, aún habrá una relación entre los datos entre la producción y la arena. Si la producción depende solamente del volumen de arena, aún habrá una relación en los datos entre la producción y el fluido.

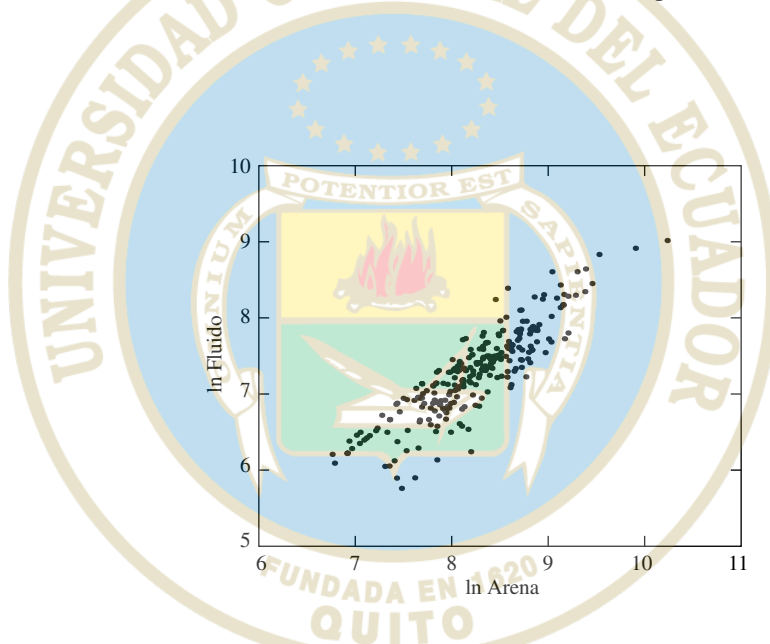


FIGURA 8.5 Diagrama de dispersión de $\ln$ Fluido contra $\ln$ Arena para 255 pozos de gas. Evidentemente hay una fuerte relación lineal. Por tanto, las relaciones aparentes entre fluido o arena y producción podrían representar un confusión en lugar de una causalidad.

La regresión múltiple proporciona una manera de resolver la cuestión. El siguiente resultado de MINITAB corresponde al modelo

$$\ln \text{ Producción} = \beta_0 + \beta_1 \ln \text{ Fluido} + \beta_2 \ln \text{ Arena} + \varepsilon \quad (8.11)$$

The regression equation is

$$\ln \text{Prod} = -0.729 + 0.670 \ln \text{Fluid} + 0.148 \ln \text{Sand}$$

Predictor	Coef	SE Coef	T	P
Constant	-0.7288	0.6719	-1.08	0.279
ln Fluid	0.6701	0.1687	3.97	0.000
ln Sand	0.1481	0.1714	0.86	0.389

S = 0.7463

R-Sq = 28.4%

R-Sq(adj) = 27.8%

Se puede ver que el coeficiente de ln Fluido es significativamente diferente de 0, pero el coeficiente de ln Arena no lo es. Si se supone que no hay otra confusión (por ejemplo, con la posición de los pozos), se puede concluir que el aumento en la cantidad de fluido tiende a aumentar la producción, pero no está claro que el aumento en la cantidad de arena tiene un efecto. Por tanto, se podría aumentar la cantidad de fluido, pero podría ser innecesario agregarle más arena.

Una observación final: ninguno de los modelos tiene un valor particularmente alto de R^2 . Lo anterior indica que hay otros factores importantes que afectan la producción y que no se han incluido en los modelos. En un análisis más completo, se intentaría identificar y medir algunos de estos factores para construir un modelo con el potencial de predicción mayor.

Colinealidad

Cuando dos variables independientes están muy fuertemente correlacionadas la regresión múltiple no puede ser capaz de determinar cuál es la importante. En este caso, se dice que las variables son **colineales**. La palabra *colineal* significa que se encuentra en la misma línea, y cuando dos variables están muy correlacionadas su diagrama de dispersión es casi una línea recta. También a veces se utiliza la palabra *multicolinealidad*. Cuando la colinealidad está presente, se dice a veces que el conjunto de variables independientes está **mal condicionado**. La tabla 8.2 de la página 578 presenta algunos datos hipotéticos que muestran el fenómeno de colinealidad.

Primero se ajustan los modelos lineales simples

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

$$y = \beta_0 + \beta_1 x_2 + \varepsilon$$

El siguiente resultado de MINITAB muestra que tanto x_1 como x_2 tienen una fuerte relación lineal con y . Los valores de r^2 están ambos alrededor de 0.96, por lo que las correlaciones r entre x_1 y y y entre x_2 y y están ambos alrededor de 0.98.

TABLA 8.2 Datos colineales

x_1	x_2	y
0.1	0.3	3.6
0.2	0.2	0.3
0.6	1.4	6.0
1.4	3.4	10.6
2.0	5.2	8.4
2.0	5.5	11.8
2.1	5.5	12.7
2.1	5.3	6.8
2.8	7.4	9.9
3.6	9.4	16.7
4.2	10.3	16.3
4.5	11.4	19.9
4.7	11.3	20.2
5.3	13.6	22.9
6.1	15.3	26.6
6.8	17.4	28.1
7.5	18.5	31.0
8.2	20.4	28.8
8.5	21.3	32.4
9.4	23.3	35.0

The regression equation is
 $Y = 2.90 + 3.53 X_1$

Predictor	Coef	SE Coef	T	P
Constant	2.8988	0.8224	3.52	0.002
X1	3.5326	0.1652	21.38	0.000

S = 2.080 R-Sq = 96.2% R-Sq(adj) = 96.0%

The regression equation is
 $Y = 2.74 + 1.42 X_2$

Predictor	Coef	SE Coef	T	P
Constant	2.7431	0.8090	3.39	0.003
X2	1.42024	0.06485	21.90	0.000

S = 2.033 R-Sq = 96.4% R-Sq(adj) = 96.2%

La figura 8.6 presenta el diagrama de dispersión de x_2 contra x_1 . Evidentemente hay una fuerte relación lineal, por lo que se sospecha que y podría tener realmente una relación con sólo una de estas variables, con lo demás enredado.

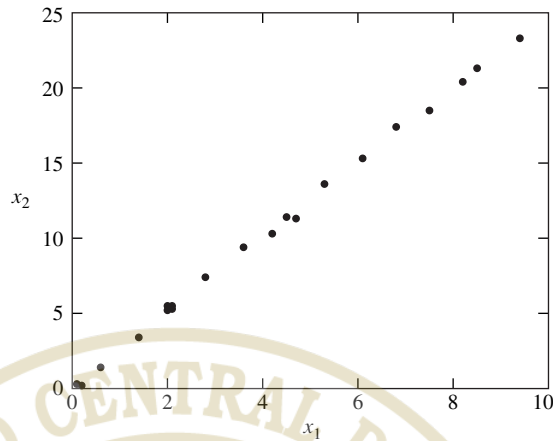


FIGURA 8.6 Las variables independientes x_1 y x_2 son *colineales* porque tienen una fuerte relación lineal.

Por tanto, se ajusta el modelo de regresión múltiple

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

El resultado MINITAB es el siguiente.

The regression equation is

$$Y = 2.72 - 0.49 X1 + 1.62 X2$$

Predictor	Coef	SE Coef	T	P
Constant	2.7248	0.8488	3.21	0.005
X1	-0.490	4.460	-0.11	0.914
X2	1.617	1.791	0.90	0.379

S = 2.091

R-Sq = 96.4%

R-Sq(adj) = 96.0%

Sorprendentemente, el resultado parece indicar que ni x_1 ni x_2 están linealmente relacionados con y , ya que ambos tienen P -valores grandes. Lo que está ocurriendo es que la relación lineal entre x_1 y x_2 es tan fuerte que resulta imposible determinar cuál de los dos es responsable de la relación lineal con y . Visto de esta manera, los P -valores grandes tienen sentido. Es creíble que el coeficiente de x_1 sea 0 y que solamente x_2 tenga una relación real con y . Por tanto, el P -valor para x_1 debe ser grande. Igual, es creíble que el coeficiente de x_2 es 0 y que sólo x_1 tiene una relación real con y . Por tanto, el P -valor para x_2 también debe ser grande.

En general, no hay mucho que se pueda hacer cuando las variables son colineales. La única buena manera de arreglar la situación es reunir más datos, incluyendo algunos valores para las variables independientes que no están en la misma línea recta. Entonces la regresión múltiple será capaz de determinar cuáles de las variables son realmente importantes.

Ejercicios para la sección 8.2

1. En un experimento para determinar los factores relacionados con la dureza de las soldaduras se midió la dureza de impacto de la muesca Charpy V pies · lb (y) para 22 soldaduras a 0°C, junto con la expansión lateral en la muesca en % (x_1), y la superficie frágil en % (x_2). En la siguiente tabla se presentan los datos.

y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
32	20.0	28	27	16.0	29	25	14.6	36
39	23.0	28	43	26.2	27	25	10.4	29
20	12.8	32	22	9.6	32	20	11.6	30
21	16.0	29	22	15.2	32	20	12.6	31
25	10.2	31	18	8.8	43	24	16.2	36
20	11.6	28	32	20.4	24	18	9.2	34
32	17.6	25	22	12.2	36	28	16.8	30
29	17.8	28						

- a) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \varepsilon$. Para cada coeficiente, pruebe la hipótesis nula que es igual a 0.
- b) Ajuste el modelo $y = \beta_0 + \beta_1 x_2 + \varepsilon$. Para cada coeficiente, pruebe la hipótesis nula que es igual a 0.
- c) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Para que cada coeficiente pruebe que la hipótesis nula es igual a 0.
- d) ¿Cuál de los modelos de los incisos a) al c) es el mejor? ¿Cómo justifica su respuesta?
2. En una prueba de laboratorio de un nuevo diseño de motor se midió la tasa de emisiones (en mg/s de óxidos de nitrógeno, NO_x) como una función de la velocidad de motor (en rpm), del momento de torsión del motor (en pies · lb), y de los caballos de fuerza totales. (De “In-Use Emissions from Heavy-Duty Diesel Vehicles”, J. Yanowitz, tesis de doctorado, Escuela de Minas, de Colorado, 2001). El resultado de MINITAB se presenta para los siguientes tres modelos:

$$\text{NO}_x = \beta_0 + \beta_1 \text{Velocidad} + \beta_2 \text{Momento de torsión} + \varepsilon$$

$$\text{NO}_x = \beta_0 + \beta_1 \text{Velocidad} + \beta_2 \text{HP} + \varepsilon$$

$$\text{NO}_x = \beta_0 + \beta_1 \text{Velocidad} + \beta_2 \text{Momento de torsión} + \beta_3 \text{HP} + \varepsilon$$

The regression equation is

$$\text{NO}_x = -321 + 0.378 \text{ Speed} - 0.160 \text{ Torque}$$

Predictor	Coef	SE Coef	T	P
Constant	-320.59	98.14	-3.27	0.003
Speed	0.37820	0.06861	5.51	0.000
Torque	-0.16047	0.06082	-2.64	0.013

$$S = 67.13$$

$$R\text{-Sq} = 51.6\% \quad R\text{-Sq}(\text{adj}) = 48.3\%$$

The regression equation is

$$\text{NOx} = -380 + 0.416 \text{ Speed} - 0.520 \text{ HP}$$

Predictor	Coef	SE Coef	T	P
Constant	-380.1	104.8	-3.63	0.001
Speed	0.41641	0.07510	5.54	0.000
HP	-0.5198	0.1980	-2.63	0.014

S = 67.19 R-Sq = 51.5% R-Sq(adj) = 48.2%

The regression equation is

$$\text{NOx} = -302 + 0.366 \text{ Speed} - 0.211 \text{ Torque} + 0.16 \text{ HP}$$

Predictor	Coef	SE Coef	T	P
Constant	-301.8	347.3	-0.87	0.392
Speed	0.3660	0.2257	1.62	0.116
Torque	-0.2106	0.8884	-0.24	0.814
HP	0.164	2.889	0.06	0.955

S = 68.31 R-Sq = 51.6% R-Sq(adj) = 46.4%

De las variables velocidad, momento de torsión y HP, ¿cuáles dos son más cercanamente colineales? ¿Cómo justifica su respuesta?

3. Dos ingenieros químicos, A y B, están trabajando de manera independiente para desarrollar un modelo que pronostique la viscosidad de un producto (y) a partir del pH (x_1) y la concentración de cierto catalizador (x_2). Cada ingeniero ha ajustado al modelo lineal

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon.$$

Los ingenieros le han enviado el resultado MINITAB que resume sus resultados

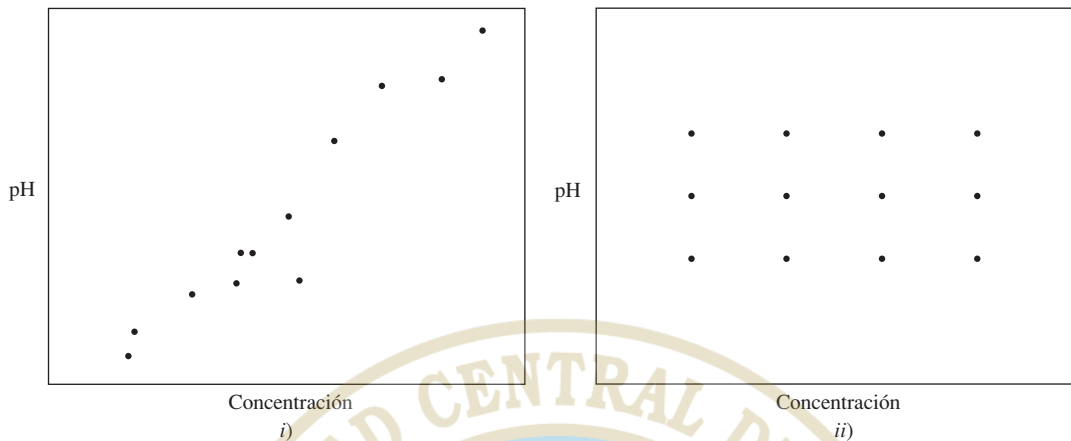
Engineer A

Predictor	Coef	SE Coef	T	P
Constant	199.2	0.5047	394.7	0.000
pH	-1.569	0.4558	-3.44	0.007
Concent.	-4.730	0.5857	-8.08	0.000

Engineer B

Predictor	Coef	SE Coef	T	P
Constant	199.0	0.548	363.1	0.000
pH	-1.256	1.983	-0.63	0.544
Concent.	-3.636	1.952	-1.86	0.112

Los ingenieros también le han enviado los siguientes diagramas de dispersión de pH contra la concentración, pero olvidaron poner sus nombres en ellos.



- a) ¿Qué gráfica proviene de cada ingeniero? ¿Cómo justifica su respuesta?
- b) ¿Cuál experimento produce los resultados más confiables? Explique.
4. El artículo “Influence of Freezing Temperature on Hydraulic Conductivity of Silty Clay” (J. Konrad y M. Samson, en *Journal of Geotechnical and Geoenvironmental Engineering*, 2000:180-187) describe un estudio de los factores que afectan la conductancia hidráulica de suelos. Las mediciones de esta última en unidades de 10^{-8} cm/s (y), cociente inicial vacío (x_1) y el cociente de deshielo vacío (x_2) para 12 elementos de arcilla cenagosa se presentan en la tabla siguiente.

y	1.01	1.12	1.04	1.30	1.01	1.04	0.955	1.15	1.23	1.28	1.23	1.30
x_1	0.84	0.88	0.85	0.95	0.88	0.86	0.85	0.89	0.90	0.94	0.88	0.90
x_2	0.81	0.85	0.87	0.92	0.84	0.85	0.85	0.86	0.85	0.92	0.88	0.92

- a) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \varepsilon$. Para cada coeficiente pruebe la hipótesis nula que es igual a 0.
- b) Ajuste el modelo $y = \beta_0 + \beta_1 x_2 + \varepsilon$. Para cada coeficiente, pruebe la hipótesis nula que es igual a 0.
- c) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Para cada coeficiente pruebe la hipótesis nula que es igual a 0.
- d) ¿Cuál de los modelos de los incisos a) al c) es el mejor? ¿Cómo justifica su respuesta?
5. En relación con el ejercicio 8 de la sección 7.4.
- a) Divida los datos en dos grupos: los puntos donde $R_1 < 4$ en un grupo, los puntos donde $R_1 \geq 4$ en los otros. Calcule la recta de mínimos cuadrados para pronosticar R_2 a partir de R_1 para cada grupo. (Usted ya hizo esto si realizó el ejercicio 8c en la sección 7.4.)
- b) Para uno de los dos grupos la relación es obviamente no lineal. Para este grupo, ajuste un modelo cuadrático (es decir, utilizando R_1 y R_1^2 como variables independientes), un modelo cúbico y otro cuártico. Calcule los P -valores para cada uno de los coeficientes en cada uno de los modelos.
- c) Grafique los residuos contra los valores ajustados para cada uno de los tres modelos del inciso b).
- d) Calcule el coeficiente de correlación entre R_1^3 y R_1^4 , y realice un diagrama de dispersión de los puntos (R_1^3 , R_1^4).
- e) Con base en el coeficiente de correlación y en el diagrama de dispersión, explique por qué son muy diferentes los P -valores para el modelo cuártico que para el cúbico.
- f) ¿Cuál de los tres modelos del inciso b) es más adecuado? ¿Por qué?

6. La tabla siguiente lista valores para tres variables medidas para 60 erupciones consecutivas del géiser Old Faithful en el Parque Nacional Yellowstone. Estas son las duraciones de la erupción (x_1), las duraciones del periodo de inactividad inmediatamente anterior a la erupción (x_2) y las duraciones del periodo de inactividad inmediatamente después de la erupción (y). Todos los tiempos están en minutos. (La variable x_2 se presentó en el ejercicio adicional 3 del capítulo 7.)

x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y
3.5	80	84	1.8	42	91	4.7	88	51	4.1	70	79
4.1	84	50	4.1	91	51	1.8	51	80	3.7	79	60
2.3	50	93	1.8	51	79	4.6	80	49	3.8	60	86
4.7	93	55	3.2	79	53	1.9	49	82	3.4	86	71
1.7	55	76	1.9	53	82	3.5	82	75	4.0	71	67
4.9	76	58	4.6	82	51	4.0	75	73	2.3	67	81
1.7	58	74	2.0	51	76	3.7	73	67	4.4	81	76
4.6	74	75	4.5	76	82	3.7	67	68	4.1	76	83
3.4	75	80	3.9	82	84	4.3	68	86	4.3	83	76
4.3	80	56	4.3	84	53	3.6	86	72	3.3	76	55
1.7	56	80	2.3	53	86	3.8	72	75	2.0	55	73
3.9	80	69	3.8	86	51	3.8	75	75	4.3	73	56
3.7	69	57	1.9	51	85	3.8	75	66	2.9	56	83
3.1	57	90	4.6	85	45	2.5	66	84	4.6	83	57
4.0	90	42	1.8	45	88	4.5	84	70	1.9	57	71

- Calcule la recta de mínimos cuadrados para predecir la duración del periodo de inactividad seguido de una erupción (y) a partir de la duración de la erupción (x_1). ¿Existe una relación lineal entre la duración de una erupción y el tiempo de espera hasta que ocurre la próxima erupción?
- Calcule la recta de mínimos cuadrados para pronosticar la duración del periodo de inactividad seguido de una erupción (y) a partir de la duración del que precede a la erupción (x_2). (Usted ya hizo esto si realizó el ejercicio adicional 3 del capítulo 7. Los resultados en este problema son ligeramente diferentes, debido a que hay 60 puntos en lugar de 59.) ¿Hay una relación lineal entre la duración del periodo de inactividad que precede a una erupción y el tiempo de espera hasta que ocurre la próxima erupción?
- Ajuste al modelo de regresión múltiple que incluya tanto la duración de la erupción x_1 como la duración del periodo de inactividad que precede a la erupción x_2 como variables independientes.
- Si usted pudiera encontrar el valor ya sea de x_1 o de x_2 pero no de ambos, ¿cuál querría conocer para pronosticar y ? ¿Ayudaría mucho conocer al otro también? Explique.

8.3 Selección de modelos

Hay muchas situaciones en las que se han medido bastantes variables independientes; se necesita decidir cuáles implican un modelo. Éste es el problema de **selección de modelos**, el cual es difícil. En la práctica con frecuencia se realiza la selección de modelos usando métodos *ad hoc*, con la guía de toda la intuición física disponible. No se intenta realizar un análisis completo de este extenso y difícil tema. En vez de eso, se limita a establecer algunos principios básicos y presentar ejemplos. Se puede consultar un texto avanzado como el de Miller (2002) para información acerca de métodos específicos.

La buena selección de modelos se basa en un principio básico conocido como navaja de Occam. Este principio se enuncia de la siguiente manera:

La navaja de Occam

El mejor modelo científico es el modelo más simple que explica los hechos observados.

En términos de modelos lineales, la navaja de Occam implica el **principio de parsimonia**:

El principio de parsimonia

Un modelo debe contener el menor número de variables necesario para ajustar los datos.

Existen algunas excepciones al principio de parsimonia:

1. Un modelo lineal siempre debe contener un intercepto, a menos que una teoría física indique otra cosa.
2. Si una potencia x^n de una variable se incluye en un modelo, también estarán incluidas todas las potencias inferiores $x^2, \dots, x^{n-1}$, a menos que una teoría física indique lo contrario.
3. Si un producto $x_i x_j$ de dos variables está incluido en un modelo, entonces las variables x_i y x_j también deben estar incluidas por separado, a menos que una teoría física indique algo distinto.

Los modelos que sólo contienen las variables necesarias para ajustar los datos se llaman **parsimoniosos**. Gran parte del trabajo práctico de la regresión múltiple implica su desarrollo.

Con el siguiente ejemplo se muestra el principio de parsimonia. Los datos en la tabla 8.3 se tomaron del artículo “Capacities and Performance Characteristics of Jaw Crushers” (S. Sastri, en *Minerals and Metallurgical Processing*, 1994:80-86). Se midieron tasas de alimentación y de potencia para varias trituradoras de quijada industriales.

TABLA 8.3 Tasas de alimentación y potencia para trituradoras de quijada industriales

Tasa de alimentación (100 t/h)	Potencia (kW)	Tasa de alimentación (100 t/h)	Potencia (kW)	Tasa de alimentación (100 t/h)	Potencia (kW)	Tasa de alimentación (100 t/h)	Potencia (kW)
0.10	11	0.20	15	0.91	45	1.36	58
1.55	60	2.91	84	0.59	12	2.36	45
3.00	40	0.36	30	0.27	24	2.95	75
3.64	150	0.14	16	0.55	49	1.09	44
0.38	69	0.91	30	0.68	45	0.91	58
1.59	77	4.27	150	4.27	150	2.91	149
4.73	83	4.36	144	3.64	100		

El siguiente resultado MINITAB presenta los resultados para ajustar el modelo

$$\text{Potencia} = \beta_0 + \beta_1 \text{Tasa de Alimentación} + \varepsilon \quad (8.12)$$

The regression equation is
Power = 21.0 + 24.6 FeedRate

Predictor	Coef	SE Coef	T	P
Constant	21.028	8.038	2.62	0.015
FeedRate	24.595	3.338	7.37	0.000

S = 26.20 R-Sq = 68.5% R-Sq(adj) = 67.2%

A partir del resultado se observa que el modelo ajustado es

$$\text{Potencia} = 21.028 + 24.595 \text{Tasa de Alimentación} \quad (8.13)$$

y que el coeficiente para la Tasa de Alimentación es significativamente diferente de 0 ($t = 7.37$, $P \approx 0$). Ahora la pregunta es si un modelo cuadrático pudiese ajustar mejor que este lineal. Así que se ajusta

$$\text{Potencia} = \beta_0 + \beta_1 \text{Tasa de Alimentación} + \beta_2 \text{Tasa de Alimentación}^2 + \varepsilon \quad (8.14)$$

Los resultados se presentan en el siguiente resultado de MINITAB. Observe que los valores para el intercepto y del coeficiente de la Tasa de Alimentación son diferentes que cuando estaban en el modelo lineal. Esto es típico. El agregar una nueva variable a un modelo podría cambiar considerablemente los coeficientes de las variables que ya estaban en el modelo.

The regression equation is
Power = 19.3 + 27.5 FeedRate - 0.64 FeedRate^2

Predictor	Coef	SE Coef	T	P
Constant	19.34	11.56	1.67	0.107
FeedRate	27.47	14.31	1.92	0.067
FeedRate^2	-0.6387	3.090	-0.21	0.838

S = 26.72 R-Sq = 68.5% R-Sq(adj) = 65.9%

El punto más importante que debe observarse es que el P -valor para el coeficiente de Tasa de Alimentación² es mayor de (0.838). Recuerde que este P -valor es para la prueba de la hipótesis nula de que el coeficiente es igual a 0. Por tanto, los datos no proporcionan evidencias de que el coeficiente de Tasa de Alimentación² sea diferente de 0. Observe también que incluyendo Tasa de Alimentación² en el modelo se aumenta sólo ligeramente el valor del estadístico de bondad del ajuste R^2 , en forma tan ligera que los primeros tres dígitos son iguales. En este sentido, no hay evidencias de que el modelo cuadrático ajuste mejor los datos que el modelo lineal, así que junto con el principio de parsimonia se debe preferir el modelo lineal.

La figura 8.7 proporciona una ilustración gráfica del principio de parsimonia. Se presenta el diagrama de dispersión de la potencia contra la tasa de alimentación, y ambas rectas de mínimos cuadrados (8.13) y el modelo cuadrático (8.14) están sobrepuestas. Aunque los coeficientes de los modelos son diferentes, se puede advertir que las dos curvas son casi idénticas. No hay razón para incluir el término cuadrático en el modelo. Esto último hace que el modelo sea más complicado, sin que mejore el ajuste.

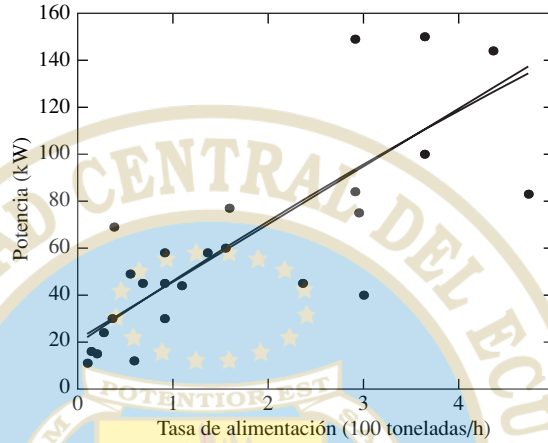


FIGURA 8.7 Diagrama de dispersión de potencia contra tasa de alimentación para 27 trituradoras de quijada industriales. La recta de mínimos cuadrados y la del mejor ajuste de modelo cuadrático están sobrepuestas. Las dos curvas son prácticamente idénticas, lo que refleja el hecho de que el coeficiente de Tasa de Alimentación² en el modelo cuadrático no es significativamente diferente de 0.

Determinando si se pueden eliminar variables de un modelo

Con frecuencia ocurre que se ha formado un modelo con muchas variables independientes, y que se desea determinar si un subconjunto en particular de ellas se puede eliminar del modelo sin reducir significativamente la precisión de éste. Para ser más específico, suponga que se conoce que el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \beta_{k+1} x_{k+1i} + \cdots + \beta_p x_{pi} + \varepsilon_i \quad (8.15)$$

es correcto, en lo que respecta a representar la relación verdadera entre las variables x y y . Se llamará a éste el modelo “completo”.

Se desea probar la hipótesis nula

$$H_0 : \beta_{k+1} = \cdots = \beta_p = 0$$

Si H_0 es verdadera, el modelo permanecerá correcto si se eliminan las variables $x_{k+1}, \dots, x_p$, por lo que se puede reemplazar el modelo completo con el siguiente modelo reducido:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \varepsilon_i \quad (8.16)$$

El desarrollo de un estadístico de prueba para H_0 se inicia calculando la suma de cuadrados del error tanto para el modelo completo como para el reducido. Se denotarán por SSE_{completo} y SSE_{reducido} . El número de grados de libertad para SSE_{completo} es $n - p - 1$, y el número de grados de libertad para SSE_{reducido} es $n - k - 1$.

Ahora, puesto que el modelo completo es correcto, se sabe que la cantidad $SSE_{\text{completo}}/(n - p - 1)$ es un estimador de la varianza del error σ^2 ; en efecto, es exactamente s^2 . Si H_0 es verdadera, entonces el modelo reducido también es correcto, así que la cantidad $SSE_{\text{reducido}}/(n - k - 1)$ es también un estimador de la varianza del error. Intuitivamente, SSE_{completo} está cerca de $(n - p - 1)\sigma^2$, y si H_0 es verdadera, SSE_{reducido} está cerca de $(n - k - 1)\sigma^2$. Por lo que si H_0 es verdadera, la diferencia $(SSE_{\text{reducido}} - SSE_{\text{completo}})$ es cercana a $(p - k)\sigma^2$, por lo que el $(SSE_{\text{reducido}} - SSE_{\text{completo}})/(p - k)$ está cerca de σ^2 . El estadístico de prueba es

$$f = \frac{(SSE_{\text{reducido}} - SSE_{\text{completo}})/(p - k)}{SSE_{\text{completo}}/(n - p - 1)} \quad (8.17)$$

Ahora, si H_0 es verdadera, tanto el numerador como el denominador de f son estimadores de σ^2 , por lo que f está probablemente cerca de 1. Si H_0 es falsa, la cantidad SSE_{reducido} tiende a ser más grande, por lo que el valor de f tiende a ser mayor. El estadístico f es un estadístico F ; su distribución nula es $F_{p-k, n-p-1}$.

El método recién descrito es muy útil en la práctica para desarrollar modelos parsimoniosos eliminando las variables superfluas. Sin embargo, las condiciones en las que esto es formalmente válido rara vez se encuentran en la práctica. Primero, es raro el caso que el modelo completo sea correcto; habrá cantidades no aleatorias que afectan el valor de la variable dependiente y que no se consideran para las variables independientes. Segundo, para que el método sea formalmente válido, el subconjunto de variables que se eliminará debe determinarse independientemente de los datos. Éste por lo general no es el caso. Más a menudo, cuando un modelo grande se ajusta, algunas de las variables se ve que tienen realmente P -valores grandes, y se utiliza la prueba F para determinar si se debe eliminar del modelo. Como se ha dicho, ésta es una técnica útil en la práctica, pero, de la misma manera que la mayoría de los métodos de la selección de modelos, debe verse como una herramienta informal en vez de un procedimiento riguroso basado en la teoría.

Se muestra el método con un ejemplo. En redes *ad hoc* de computadoras móviles, los mensajes deben enviarse de computadora a computadora hasta que alcanzan su destino. La sobrecarga de datos es el número de bytes de información que se deben transmitir junto con los mensajes para llegar a las posiciones correctas. Un protocolo exitoso generalmente tiene una sobrecarga de datos menor. La tabla 8.4 presenta la velocidad promedio, el tiempo de espera, la razón de cambio de enlace (LCR), y la sobrecarga de datos para 25 redes de computadoras simuladas. La razón de cambio de enlace para una computadora en particular es la razón en la que las otras computadoras en la red entran y dejan el rango de transmisión de la computadora dada. Estos datos fueron generados para un estudio publicado en el artículo “Metrics to Enable Adaptive Protocols for Mobile Ad Hoc Networks” (J. Boleng, W. Navidi y T. Camp, en *Proceedings of the 2002 International Conference On Wireless Networks*, 2002:293-298).

Se inicia ajustando un modelo realmente grande para estos datos, a saber,

$$\begin{aligned} \text{Sobrecarga} = & \beta_0 + \beta_1 \text{LCR} + \beta_2 \text{Velocidad} + \beta_3 \text{Pausa} + \beta_4 \text{Velocidad} \cdot \text{Pausa} \\ & + \beta_5 \text{LCR}^2 + \beta_6 \text{Velocidad}^2 + \beta_7 \text{Pausa}^2 + \varepsilon \end{aligned}$$

TABLA 8.4 Sobrecarga de datos, velocidad, tiempo de pausa, y razón de cambio de enlace por una red de computadoras móviles

Velocidad (m/s)	Tiempo de pausa (s)	LCR (100/s)	Sobrecarga de datos (kB)	Velocidad (m/s)	Tiempo de pausa (s)	LCR (100/s)	Sobrecarga de datos (kB)
5	10	9.426	428.90	20	40	12.117	501.48
5	20	8.318	443.68	20	50	10.284	519.20
5	30	7.366	452.38	30	10	33.009	445.45
5	40	6.744	461.24	30	20	22.125	489.02
5	50	6.059	475.07	30	30	16.695	506.23
10	10	16.456	446.06	30	40	13.257	516.27
10	20	13.281	465.89	30	50	11.107	508.18
10	30	11.155	477.07	40	10	37.823	444.41
10	40	9.506	488.73	40	20	24.140	490.58
10	50	8.310	498.77	40	30	17.700	511.35
20	10	26.314	452.24	40	40	14.064	523.12
20	20	19.013	475.97	40	50	11.691	523.36
20	30	14.725	499.67				

Los resultados del ajuste a este modelo son los siguientes.

The regression equation is

$$\text{Overhead} = 368 + 3.48 \text{ LCR} + 3.04 \text{ Speed} + 2.29 \text{ Pause} - 0.0122 \text{ Speed} \cdot \text{Pause} - 0.1041 \text{ LCR}^2 - 0.0313 \text{ Speed}^2 - 0.0132 \text{ Pause}^2$$

Predictor	Coef	SE Coef	T	P
Constant	367.96	19.40	18.96	0.000
LCR	3.477	2.129	1.63	0.121
Speed	3.044	1.591	1.91	0.073
Pause	2.2924	0.6984	3.28	0.004
Speed*Pa	-0.01222	0.01534	-0.80	0.437
LCR^2	-0.10412	0.03192	-3.26	0.005
Speed^2	-0.03131	0.01906	-1.64	0.119
Pause^2	-0.01318	0.01045	-1.26	0.224

S = 5.72344 R-Sq = 97.2% R-Sq(adj) = 96.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	19567.5	2795.4	85.33	0.000
Residual Error	17	556.9	32.8		
Total	24	20124.3			

Se puede ver que LCR, Velocidad · Pausa, Velocidad², y Pausa² tienen *P*-valores grandes. Se dejará a LCR en el modelo por ahora, porque LCR² tiene un *P*-valor muy pequeño, y, por tanto, debe quedarse en el modelo. Se utilizará la prueba *F* para determinar si el modelo reducido obtenido al eliminar Velocidad · Pausa, Velocidad² y Pausa² es razonable. Primero, del

resultado para el modelo completo, observe que $SSE_{\text{completo}} = 556.9$, y que tiene 17 grados de libertad. El número de variables independientes en el modelo completo es $p = 7$.

Queda bien el modelo reducido

$$\text{Sobrecarga} = \beta_0 + \beta_1 \text{LCR} + \beta_2 \text{Velocidad} + \beta_3 \text{Pausa} + \beta_5 \text{LCR}^2 + \varepsilon$$

Los resultados del ajuste de este modelo son:

The regression equation is

$$\text{Overhead} = 359 + 6.69 \text{ LCR} + 0.777 \text{ Speed} + 1.67 \text{ Pause} - 0.156 \text{ LCR}^2$$

Predictor	Coef	SE Coef	T	P
Constant	359.22	13.01	27.61	0.000
LCR	6.695	1.156	5.79	0.000
Speed	0.7766	0.2054	3.78	0.001
Pause	1.6729	0.1826	9.16	0.000
LCR^2	-1.5572	0.02144	-7.26	0.000

S = 6.44304

R-Sq = 95.9%

R-Sq(adj) = 95.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	19294.1	4823.5	116.19	0.000
Residual Error	20	830.3	41.5		
Total	24	20124.3			

Los P -valores para las variables en este modelo son bastante pequeños. Del resultado para este modelo reducido, se observa que $SSE_{\text{reducido}} = 830.3$. El número de variables en este modelo reducido es $k = 4$.

Ahora se puede calcular el estadístico F . Utilizando la ecuación (8.17) se calcula

$$f = \frac{(830.3 - 556.9)/(7 - 4)}{556.9/17} = 2.78$$

La distribución nula es $F_{3,17}$. De la tabla F (tabla A.7 en el apéndice A) se encuentra que $0.05 < P < 0.10$. De acuerdo con la regla general del 5%, puesto $P > 0.05$, el modelo reducido es creíble, pero sólo apenas es cierto. Más que establecer un modelo apenas creíble, es inteligente ir más lejos para buscar un modelo ligeramente menos reducido que tenga un P -valor mayor.

Con este propósito se observa que de las tres variables que se eliminan, la variable Velocidad² tenía el P -valor más pequeño en el modelo completo. Lo anterior se tomará como una indicación de que ésta podría ser la más importante de las variables que se elimina, y se pondrá en el modelo. Ahora se ajustará un segundo modelo reducido, que es

$$\text{Sobrecarga} = \beta_0 + \beta_1 \text{LCR} + \beta_2 \text{Velocidad} + \beta_3 \text{Pausa} + \beta_5 \text{LCR}^2 + \beta_6 \text{Velocidad}^2$$

Los resultados del ajuste a este modelo son:

The regression equation is
Overhead = 373 + 4.80 LCR + 1.99 Speed + 1.45 Pause - 0.123 LCR^2
-0.0212 Speed^2

Predictor	Coef	SE Coef	T	P
Constant	372.60	16.93	22.00	0.000
LCR	4.799	1.935	2.48	0.023
Speed	1.993	1.023	1.95	0.066
Pause	1.4479	0.2587	5.60	0.000
LCR^2	-0.12345	0.03400	-3.63	0.002
Speed^2	-0.02120	0.01746	-1.21	0.240

S = 6.36809 R-Sq = 96.2% R-Sq(adj) = 95.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	19353.8	3870.8	95.45	0.000
Residual Error	19	770.5	40.6		
Total	24	20124.3			

Observe que el P -valor para Velocidad² en este modelo es grande (0.240). Esto no es bueno. En general no se quiere agregar una variable cuyo coeficiente podría ser igual a 0. Por lo que tal vez no se querrá seguir con este modelo. De todos modos, sólo por práctica se calcula el valor del estadístico F . El valor de SSE_{reducido} en este modelo es 770.5. El número de variables independientes es $k = 5$. Por tanto, el valor del estadístico F , utilizando la ecuación (8.17), es

$$f = \frac{(770.5 - 556.9)/(7 - 5)}{556.9/17} = 3.26$$

La distribución nula es $F_{2,17}$. En la tabla F (tabla A.7) se encuentra nuevamente que $0.05 < P < 0.10$, por lo que el modelo reducido apenas es creíble como el mejor.

Se elige poner Velocidad² en el modelo porque tenía el menor P -valor entre las variables eliminadas originalmente. Pero, como se acaba de ver, esto último no garantiza que tiene un P -valor pequeño cuando se pone nuevamente en el modelo reducido. Quizás una de las otras variables que se eliminaron sería mejor. De las tres variables que se eliminaron originalmente, la que tenga el segundo P -valor más pequeño era Pausa². Se trata de sustituir Velocidad² con Pausa². En este contexto, ahora se ajusta a un tercer modelo:

$$\text{Sobrecarga} = \beta_0 + \beta_1 \text{LCR} + \beta_2 \text{Velocidad} + \beta_3 \text{Pausa} + \beta_5 \text{LCR}^2 + \beta_6 \text{Pausa}^2$$

Los resultados del ajuste a este modelo son:

The regression equation is

$$\text{Overhead} = 345 + 6.484 \text{ LCR} + 0.707 \text{ Speed} + 2.85 \text{ Pause} - 0.145 \text{ LCR}^2 - 0.0183 \text{ Pause}^2$$

Predictor	Coef	SE Coef	T	P
Constant	345.42	13.19	26.20	0.000
LCR	6.484	1.050	6.17	0.000
Speed	0.7072	0.1883	3.76	0.001
Pause	2.8537	0.5337	5.35	0.000
LCR^2	-0.14482	0.01996	-7.25	0.000
Pause^2	-0.018334	0.007879	-2.33	0.031

S = 5.83154 R-Sq = 96.8% R-Sq(adj) = 95.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	19478.2	3895.6	114.55	0.000
Residual Error	19	646.1	34.0		
Total	24	20124.3			

Este modelo tiene buen aspecto, al menos al principio. Todas las variables tienen P -valores pequeños. Se calculará el estadístico F para ver si este modelo es creíble. El valor de $\text{SSE}_{\text{reducido}}$ en este modelo es 646.1. El número de variables independientes es $k = 5$. El valor del estadístico F , usando la ecuación (8.17), es, por tanto,

$$f = \frac{(646.1 - 556.9)/(7 - 5)}{556.9/17} = 1.36$$

La distribución nula es $F_{2,17}$. Al consultar la tabla F (tabla A.7) se encuentra que el punto 0.10 en esta distribución F es 2.64. Por tanto, el P -valor es mucho mayor que 0.10. Este modelo es evidentemente creíble.

Se ha utilizado un método informal para encontrar un buen modelo parsimonioso. Es importante darse cuenta que este procedimiento informal se podría haber realizado de manera diferente, con elecciones diferentes de las variables eliminadas e incluidas en el modelo. Se podría haber tenido un modelo final diferente que pudiera haber sido tan bueno como el que se encontró. Con frecuencia, en la práctica hay muchos modelos que ajustan los datos casi igualmente bien; no hay un único modelo “correcto”.

Regresión con los mejores subconjuntos

Como se ha mencionado, con frecuencia los métodos de selección de modelos son más informales y adecuados. Sin embargo, hay algunas herramientas que pueden hacer el proceso más sistemático. Una de ellas es la **regresión con los mejores subconjuntos**. El concepto es muy simple. Suponga que hay p variables independientes, $x_1, \dots, x_p$, que están disponibles para

El procedimiento para encontrar los mejores subconjuntos es computacionalmente intenso. Cuando hay muchas potenciales variables independientes, existirán bastantes modelos para ajustar. Sin embargo, para la mayoría de los conjuntos de datos, las computadoras actuales pueden manejar 30 o más variables independientes, ello es suficiente para tratar muchas situaciones prácticas. El siguiente resultado de MINITAB es para la regresión con los mejores subconjuntos aplicados a los datos de la tabla 8.4. Se están considerando siete variables independientes: Velocidad, Pausa, LCR, Velocidad · Pausa, Velocidad², Pausa² y LCR².

En este resultado se presenta tanto el mejor y el segundo mejor subconjunto, para tamaños del 1 al 7. Se destaca que el término *mejor* sólo significa que el modelo tiene el mayor valor de R^2 , y no garantiza que éste sea el mejor en cualquier sentido práctico. Se explicará el resultado columna por columna. La primera, rotulada como “Vars”, presenta el número de va-

riables en el modelo. Por tanto, el primer renglón de la tabla describe el mejor modelo que se puede hacer con una variable independiente, y el segundo renglón describe el segundo mejor de este modelo. El tercero y cuarto renglones describen los mejores y los segundos mejores modelos que se pueden hacer con dos variables, y así sucesivamente. La segunda columna presenta el coeficiente de determinación, R^2 , para cada modelo. Observe que el valor de R^2 para el mejor subconjunto aumenta cuando el número de variables lo hace. Es un hecho matemático que el mejor subconjunto de $k + 1$ variables de k siempre tendrá por lo menos un R^2 tan grande como el mejor subconjunto de variables k . Por el momento no se considerarán las dos siguientes columnas. La columna rotulada como “s” presenta la estimación de la desviación estándar del error. Es la de raíz cuadrada de la estimación s^2 (ecuación 8.8 de la sección 8.1). Por último, las columnas de la derecha representan variables independientes que son candidatas para incluirse en el modelo. El nombre de cada variable se escribe verticalmente arriba de su columna. Una “X” en la columna significa que la variable está incluida en el modelo. Por tanto, el mejor modelo que contiene cuatro variables es el que tiene las variables Velocidad, Pausa, LCR y LCR².

Al observar los mejores subconjuntos del resultado de regresión, es importante fijarse que existe una pequeña diferencia en el ajuste entre el mejor y el segundo modelo de cada tamaño (excepto para el tamaño 1). También es importante darse cuenta de que el valor de R^2 es una cantidad aleatoria, que depende de los datos. Si el proceso se repitiera y se obtuvieran nuevos datos, los valores de R^2 para los diferentes modelos serían algo diferentes, y los modelos diferentes serían “mejores”. Por esta razón no se debe usar este procedimiento, o cualquier otro, para elegir sólo un modelo. En lugar de ello se debe dar cuenta de que habrá muchos modelos que ajustan los datos casi igualmente bien.

Sin embargo, existen métodos que se han desarrollado para elegir un solo modelo, presumiblemente el “mejor” de los “mejores”. Aquí se describen dos de ellos, con la advertencia de que no se consideren muy en serio. Se inicia al observar que si sólo se elige el modelo con el valor más alto de R^2 , siempre se elegirá el que contiene todas las variables, ya que el valor de R^2 aumenta necesariamente cuando el número de variables en el modelo se incrementa. Los métodos de selección de un modelo implican estadísticos que ajustan el valor de R^2 , para eliminar esta característica.

El primero es el R^2 ajustado. Sea n el número de observaciones y k el de variables independientes en el modelo. El R^2 ajustado se define así:

$$R^2 \text{ ajustado} = R^2 - \left(\frac{k}{n - k - 1} \right) (1 - R^2) \quad (8.18)$$

El R^2 ajustado es siempre menor que R^2 , ya que una cantidad positiva es restada de R^2 . Conforme el número de variables k aumenta, R^2 aumentará, pero la cantidad restada de éste también lo hará. El valor de k para el cual el valor de R^2 ajustado es un máximo se puede usar para determinar el número de variables en el modelo, y el mejor subconjunto de ese tamaño que se puede elegir como modelo. En el resultado anterior se puede ver que el R^2 ajustado alcanza su máximo (96.2%) en el modelo de seis variables que contienen las variables Velocidad, Pausa, LCR, Velocidad², Pausa² y LCR².

Otro estadístico comúnmente usado es C_p de Mallows. Para calcular esta cantidad, sea n el número de observaciones, p el número total de variables independientes en consideración y k el número de variables independientes en un subconjunto. Como antes, sea SSE_{completo} la

suma de los cuadrados del error para el modelo completo que contiene todas las variables p , y sea SSE_{reducido} la suma de los cuadrados del error para el modelo que contiene solamente el subconjunto de variables k . El C_p de Mallows se define como

$$C_p = \frac{(n - p - 1)SSE_{\text{reducido}}}{SSE_{\text{completo}}} - (n - 2k - 2) \quad (8.19)$$

Para modelos que contienen tantas variables independientes como se necesiten, se supone que el valor de C_p es casi igual al número de variables, incluyendo el intercepto, en el modelo. Para elegir sólo un modelo se puede elegir éste ya sea con el valor mínimo de C_p , o se puede escoger el modelo en el que el valor de C_p está más cerca del número de variables independientes en el modelo. En el resultado anterior, ambos criterios producen el mismo modelo de seis variables elegidas por el criterio del R^2 ajustado. El valor de C_p para este modelo es 6.6.

Por último, se indica que el procedimiento *ad hoc* es utilizar la prueba F que produjo el modelo de cinco variables que contiene las variables Velocidad, Pausa, LCR, Pausa² y LCR². El resultado muestra que este modelo es el mejor de cinco variables en función de R^2 . Su R^2 ajustado es 95.9%, y su valor C_p , 6.7, ambos están cercanos a sus valores óptimos. En la práctica no hay una razón clara para preferir el modelo elegido con R^2 ajustado o con C_p de Mallows, o viceversa.

Regresión stepwise

La **regresión stepwise** es quizás la técnica de selección de modelos más ampliamente usada. Su ventaja principal sobre la regresión con los mejores subconjuntos es que es menos intensiva computacionalmente, por lo que se puede utilizar en situaciones donde hay un número muy grande de variables candidatas independientes y demasiados subconjuntos posibles para revisarse cada uno. La versión de la regresión stepwise que se describirá está basada en los P -valores de los estadísticos t para las variables independientes. Una versión equivalente tiene como sustento el estadístico F (que es el cuadrado del estadístico t). Antes de operar el algoritmo, el usuario elige dos P -valores de umbral, α_{dentro} y α_{fuera} , con $\alpha_{\text{dentro}} \leq \alpha_{\text{fuera}}$. La regresión stepwise empieza con un paso de **selección hacia adelante**, en el cual se selecciona la variable independiente con el P -valor más pequeño, suponiendo que se satisface $P < \alpha_{\text{dentro}}$. Esta variable se introduce en el modelo, creando un modelo con una sola variable independiente. Esta variable se denotará por x_1 . En el siguiente paso, que también es un paso de selección hacia adelante, se revisan una a una las variables restantes como candidatas para la segunda variable en el modelo. La que tenga el P -valor más pequeño se agrega al modelo, suponiendo nuevamente que $P < \alpha_{\text{dentro}}$.

Ahora, es posible que al haber agregado la segunda variable al modelo se provoque un aumento en el P -valor de la primera variable. En el siguiente paso, llamado de **eliminación hacia atrás**, la primera variable se elimina del modelo si su P -valor es mayor que α_{fuera} . El algoritmo continúa alternando los pasos de selección hacia adelante con los de eliminación hacia atrás: en cada paso de selección hacia adelante se agrega la variable con el P -valor más pequeño si $P < \alpha_{\text{dentro}}$, y en cada paso de eliminación hacia atrás se elimina la variable con el P -valor más grande si $P > \alpha_{\text{fuera}}$. El algoritmo se termina cuando ninguna variable satisface los criterios para ser agregada o eliminada del modelo.

El siguiente resultado es el procedimiento de regresión stepwise de MINITAB, aplicado a los datos de la tabla 8.4. Los umbrales para los P -valores son $\alpha_{\text{dentro}} = \alpha_{\text{fuera}} = 0.15$. Hay un total de siete variables independientes que deben considerarse: Velocidad, Pausa, LCR, Velocidad $\cdot$ Pausa, Velocidad², Pausa² y LCR².

Alpha-to-Enter = 0.15 Alpha-to-Remove = 0.15					
Response is Overhead on 7 predictors, with N = 25					
Step	1	2	3	4	5
Constant	452.2	437.3	410.7	388.4	338.5
Speed*Pause	0.0470	0.0355	0.0355	0.0304	0.0146
T-Value	8.03	6.00	6.96	5.21	3.52
P-Value	0.000	0.000	0.000	0.000	0.002
Pause		0.74	3.02	3.75	3.24
T-Value		3.40	3.78	4.20	6.25
P-Value		0.003	0.001	0.000	0.000
Pause^2			-0.0380	-0.0442	-0.0256
T-Value			-2.94	-3.39	-3.19
P-Value			0.008	0.003	0.005
LCR				0.69	6.97
T-Value				1.62	6.95
P-Value				0.121	0.000
LCR^2					-0.139
T-Value					-6.46
P-Value					0.000
S	15.2	12.6	10.8	10.4	5.99
R-Sq	73.70	82.74	87.77	89.19	96.62
R-Sq(adj)	72.55	81.18	86.02	87.02	95.73
Mallows C-p	140.6	87.0	58.1	51.4	7.8

En el paso 1, la variable Velocidad $\cdot$ Pausa tenía el P -valor (0.000) más pequeño entre las siete, por lo que fue la primera variable en el modelo. En el paso 2, la Pausa tenía el P -valor (0.003) más pequeño entre las variables restantes, por lo que fue la segunda variable agregada en el modelo. El P -valor para la variable Velocidad $\cdot$ Pausa permanecía en 0.000 después de que se agregó Pausa al modelo; debido a que no subió a un valor mayor que $\alpha_{\text{fuera}} = 0.15$, no se elimina del modelo. En los pasos 3, 4 y 5 las variables Pausa², LCR y LCR² son agregadas una tras otra. En ningún punto el P -valor de una variable en el modelo superó el umbral $\alpha_{\text{fuera}} = 0.15$, por lo que no se eliminó ninguna variable. Después de cinco pasos, ninguna de las variables que permanecían tenía P -valores menores que $\alpha_{\text{dentro}} = 0.15$, por lo que el algoritmo termina. El modelo final contiene las variables Velocidad $\cdot$ Pausa, Pausa, Pausa², LCR y LCR².

El modelo elegido por la regresión stepwise es uno con cinco variables. En comparación con el resultado arrojado por la regresión con los mejores subconjuntos, muestra que no es uno de los mejores dos modelos con cinco variables en términos de R^2 . El modelo ajusta bien, y solamente en relación con el ajuste es razonable. Se indica que este modelo tiene una característica indeseable que contiene el término de interacción Velocidad · Pausa sin contener la variable Velocidad sola. Esto último indica una debilidad de todos los procedimientos automáticos de selección de variables, incluyendo la regresión stepwise y la regresión con los mejores subconjuntos. Operan sólo con base en la bondad del ajuste, y pueden no considerar las relaciones entre variables independientes que son importantes.

Los procedimientos de selección de modelos algunas veces encuentran modelos cuando no deben hacerlo

Cuando se construye un modelo para pronosticar el valor de una variable dependiente, podría parecer razonable tratar de empezar con tantas variables candidatas independientes como sea posible para que un procedimiento de selección de modelos tenga un número muy grande de modelos que elegir. Desgraciadamente, lo anterior no es una buena idea, como se mostrará a continuación.

Un coeficiente de correlación se puede calcular entre cualesquiera dos variables. A veces, dos variables sin ninguna relación real estarán correlacionadas fuertemente por probabilidad. Por ejemplo, George Udny Yule observó que la tasa de natalidad anual en Gran Bretaña estaba casi perfectamente correlacionada ($r = -0.98$) con la producción anual de hierro en lingotes en Estados Unidos durante 1875-1920. Aunque nadie sugeriría tratar de pronosticar alguna de estas variables por la otra. Ello ilustra una dificultad que comparten todos los procedimientos de selección de modelos. Lo más probable es que algunas de las variables independientes que se proporcionan sean mejores candidatas solamente por azar y presenten correlaciones sin sentido con la variable dependiente.

Este fenómeno se muestra con una simulación. Se genera una muestra aleatoria simple $y_1, \dots, y_{30}$ de tamaño 30 de una distribución $N(0, 1)$. Esta muestra se denota por y . Después se generan 20 muestras independientes más de tamaño 30 de una distribución $N(0, 1)$; estas muestras se denotarán por $x_1, \dots, x_{20}$. Para hacer la notación clara, la muestra x_i contiene 30 valores $x_{i1}, \dots, x_{i30}$. Después se aplican ambas, la regresión stepwise y la regresión con los mejores subconjuntos, a estos datos simulados. Ninguno de los x_i está relacionado con y ; todos fueron generados de manera independiente. Por tanto, el resultado ideal de un procedimiento de selección de modelos sería producir otro similar sin ninguna variable dependiente. El comportamiento verdadero era muy diferente. Los dos siguientes resultados de MINITAB muestran la regresión stepwise y la regresión con los mejores subconjuntos. El método de regresión stepwise recomienda un modelo que contiene seis variables, con un R^2 ajustado de 41.89%. El procedimiento de la regresión con los mejores subconjuntos produce el modelo de mejor ajuste para cada número de variables de 1 al 20. Usando el criterio de R^2 ajustado, el procedimiento de mejores subconjuntos recomienda un modelo de 12 variables, con un R^2 ajustado de 51.0%. De acuerdo con el criterio mínimo de C_p de Mallows, el “mejor” modelo es uno de cinco variables.

Alguien que considerara este resultado en sentido literal creería que algunas de las variables independientes podrían ser útiles para pronosticar la variable dependiente. Pero ninguna de ellas lo es. Todas las relaciones aparentes se deben completamente al azar.

Stepwise Regression: Y versus X1, X2, ...

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15

Response is Y on 20 predictors, with N = 30

Step	1	2	3	4	5	6
Constant	0.14173	0.11689	0.12016	0.13756	0.09070	0.03589
X15	-0.38	-0.38	-0.28	-0.32	-0.28	-0.30
T-Value	-2.08	-2.19	-1.60	-1.87	-1.69	-1.89
P-Value	0.047	0.037	0.122	0.073	0.105	0.071
X6		0.39	0.55	0.57	0.57	0.52
T-Value		2.04	2.76	2.99	3.15	2.87
P-Value		0.051	0.010	0.006	0.004	0.009
X16			-0.43	-0.43	-0.55	-0.73
T-Value			-1.98	-2.06	-2.60	-3.07
P-Value			0.058	0.050	0.016	0.005
X12				0.33	0.42	0.49
T-Value				1.79	2.29	2.66
P-Value				0.086	0.031	0.014
X3					-0.42	-0.52
T-Value					-1.83	-2.23
P-Value					0.080	0.035
X17						0.35
T-Value						1.53
P-Value						0.140
S	1.15	1.09	1.04	0.998	0.954	0.928
R-Sq	13.33	24.92	34.75	42.15	49.23	53.91
R-Sq(adj)	10.24	19.36	27.22	32.90	38.66	41.89
Mallows C-p	5.5	3.3	1.7	1.0	0.4	0.7

Best Subsets Regression: Y versus X1, X2, ...

Response is Y

				X X X X X X X X X 1 1 1 1 1 1 1 1 1 2																							
				X X X X X X X X X 1 1 1 1 1 1 1 1 1 2																							
Vars	R-Sq	R-Sq(adj)	C-p	S	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0			
1	13.3	10.2	5.5	1.1539																X							
2	28.3	23.0	2.0	1.0685						X											X						
3	34.8	27.2	1.7	1.0390						X										X	X						
4	43.2	34.1	0.6	0.98851			X			X						X				X							
5	49.2	38.7	0.4	0.95391			X		X							X				X	X						
6	53.9	41.9	0.7	0.92844			X		X							X				X	X	X					
7	57.7	44.3	1.3	0.90899	X	X			X							X				X	X		X				
8	61.2	46.4	2.1	0.89168			X		X							X	X			X	X	X	X				
9	65.0	49.3	2.7	0.86747			X		X	X						X	X			X	X	X	X				
10	67.6	50.5	3.8	0.85680			X		X	X						X	X			X	X	X	X				
11	69.2	50.4	5.2	0.85803	X	X	X		X	X						X	X			X	X	X	X				
12	71.3	51.0	6.4	0.85267	X	X	X		X	X						X	X			X	X	X	X				
13	72.4	49.9	8.0	0.86165	X	X	X		X	X						X	X	X		X	X	X	X				
14	73.0	47.8	9.8	0.87965	X	X	X		X	X	X					X	X	X		X	X	X	X				
15	74.2	46.5	11.4	0.89122	X	X	X		X	X	X					X	X	X		X	X	X	X				
16	74.5	43.1	13.3	0.91886	X	X	X		X	X	X	X				X	X	X		X	X	X	X				
17	74.8	39.2	15.1	0.94985	X	X	X		X	X	X	X				X	X	X	X	X	X	X	X				
18	75.1	34.2	17.1	0.98777	X	X	X		X	X	X	X				X	X	X	X	X	X	X	X				
19	75.1	27.9	19.0	1.0344	X	X	X		X	X	X	X	X			X	X	X	X	X	X	X	X				
20	75.2	20.1	21.0	1.0886	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X				

¿Cómo se puede determinar qué variables, si las hay, en el modelo seleccionado están muy relacionadas con la variable dependiente, y cuáles se seleccionaron solamente por azar? Los métodos estadísticos no son de mucha ayuda aquí. El método más confiable es repetir el experimento, reuniendo más datos de la variable dependiente y de las variables independientes que fueron seleccionadas para el modelo. Entonces las variables independientes sugeridas por el procedimiento de selección se pueden ajustar a la variable dependiente utilizando los nuevos datos. Si algunas de estas variables ajusta bien a los nuevos datos, la evidencia de una relación real será más convincente.

Enseguida se resume el análisis de la selección de modelos destacando cuatro puntos.

Resumen

Cuando seleccione un modelo de regresión considere lo siguiente:

- Cuando hay poca o ninguna teoría física para apoyarse, muchos modelos diferentes ajustarán los datos casi igualmente bien.
- Los métodos para elegir un modelo implican estadísticos (R^2 , el estadístico F , C_p) cuyos valores dependen de los datos. Por tanto, si el experimento se repite, estos estadísticos tendrán valores diferentes, y los modelos diferentes pueden parecer ser “mejores”.
- Algunas o todas las variables independientes en un modelo seleccionado pueden no estar realmente relacionadas con la variable dependiente. Siempre que sea posible, los experimentos se deben repetir para evaluar estas relaciones aparentes.
- La selección de modelos es un arte, no una ciencia.

Ejercicios para la sección 8.3

1. Verdadero o falso:

- a) Para algún conjunto de datos siempre hay un mejor modelo.
- b) Cuando no hay teoría física sobre la cual especificar un modelo, por lo general no hay ningún mejor modelo, pero muchos son casi igualmente buenos.
- c) Los métodos de selección de modelos como la regresión con los mejores subconjuntos o la regresión stepwise, cuando se usan adecuadamente, están diseñados en forma científica para encontrar el mejor modelo disponible.
- d) Los métodos de selección de modelos como la regresión con los mejores subconjuntos y regresión stepwise, cuando se usan adecuadamente, pueden sugerir modelos que ajustan los datos bien.

2. El artículo “Experimental Design Approach for the Optimization of the Separation of Enantiomers in Preparative Liquid Chromatography” (S. Lai y Z. Lin, en *Separation Science and Technology*, 2002:847-875) describe un experimento que implica un proceso químico diseñado para separar enantiómeros. Se ajustó un modelo para estimar el tiempo del ciclo (y) en función de la razón de flujo (x_1), la concentración de la muestra (x_2), y la composición de fase móvil (x_3). En la tabla siguiente se presentan los resultados de un ajuste de mínimos cuadrados. (El artículo no proporciona el valor del estadístico t para el término constante.)

Pronosticador	Coefficiente	T	P
Constante	1.603		
x_1	-0.619	-22.289	0.000
x_2	0.086	3.084	0.018
x_3	0.306	11.011	0.000
x_1^2	0.272	8.542	0.000
x_2^2	0.057	1.802	0.115
x_3^2	0.105	3.300	0.013
x_1x_2	-0.022	-0.630	0.549
x_1x_3	-0.036	-1.004	0.349
x_2x_3	0.036	1.018	0.343

De los siguientes, ¿cuál es el mejor paso siguiente en el análisis?

- i) No hay necesidad de hacer ninguno. Este modelo es excelente.

- ii) Eliminar x_1^2 , x_2^2 y x_3^2 del modelo, y después realizar una prueba F .
- iii) Eliminar x_1x_2 , x_1x_3 y x_2x_3 del modelo, y después realizar una prueba F .
- iv) Eliminar x_1 y x_1^2 del modelo, y después realizar una prueba F .
- v) Agregar los términos cúbicos x_1^3 , x_2^3 y x_3^3 al modelo para tratar de mejorar el ajuste.
3. En el artículo que se cita en el ejercicio 2 se ajustó un modelo para investigar la relación entre las variables independientes dadas en el ejercicio 2 y la cantidad del isómero S coleccionada. Los resultados del ajuste de mínimos cuadrados se presentan en la tabla siguiente. (El artículo no proporciona el valor del estadístico t para el término constante.)

Pronosticador	Coeficiente	T	P
Constante	3.367		
x_1	-0.018	-1.729	0.127
x_2	1.396	135.987	0.000
x_3	0.104	10.098	0.000
x_1^2	0.017	1.471	0.184
x_2^2	-0.023	-0.909	0.394
x_3^2	-0.030	-2.538	0.039
x_1x_2	-0.006	-0.466	0.655
x_1x_3	0.012	0.943	0.377
x_2x_3	0.055	4.194	0.004

De lo siguiente, ¿cuál es el mejor paso siguiente en el análisis?

- i) Elimine x_1^2 , x_2^2 y x_3^2 del modelo, y después realice una prueba F .
- ii) No se tiene que hacer nada. Este modelo es excelente.
- iii) Agregar términos cúbicos x_1^3 , x_2^3 y x_3^3 al modelo para tratar de mejorar el ajuste.
- iv) Eliminar x_1x_2 , x_1x_3 y x_2x_3 del modelo, después realice una prueba F .
- v) Elimine x_2^2 , x_1x_2 y x_1x_3 del modelo, y después realice una prueba F .
4. Un ingeniero mide una variable dependiente y y las variables independientes x_1 , x_2 , y x_3 . El resultado de MINITAB para el modelo $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \varepsilon$ se presenta a continuación.

The regression equation is

$$Y = 0.367 + 1.61 X_1 - 5.51 X_2 + 1.27 X_3$$

Predictor	Coef	SE Coef	T	P
Constant	0.3692	0.9231	0.40	0.698
X_1	1.6121	1.3395	1.21	0.254
X_2	5.5049	1.4959	3.68	0.004
X_3	1.2646	1.9760	0.64	0.537

De los siguientes, ¿cuál es el mejor paso siguiente en el análisis?

- i) Agregar términos de interacción x_1x_2 y x_2x_3 con el propósito de intentar determinar más variables para poner en el modelo.
- ii) Agregar el término de interacción x_1x_3 para intentar determinar otra variable para poner en el modelo.
- iii) No se tiene que hacer nada. Este modelo es excelente.
- iv) Eliminar x_1 y x_3 , y después realizar una prueba F .
- v) Eliminar x_2 , y después realizar una prueba F .
- vi) Eliminar el intercepto (constante), ya que tiene el P -valor más grande.

5. Un fisiólogo está tratando de pronosticar las tasas de respiración de las personas (en m^3/h) durante ejercicio pesado a partir de sus tasas de respiración en reposo, mientras están sentadas, y durante ejercicio ligero. El siguiente resultado de MINITAB es del modelo

$$\text{Pesado} = \beta_0 + \beta_1 \text{Ligero} + \beta_2 \text{Sentado} + \beta_3 \text{Reposo} + \beta_4 \text{Ligero} \cdot \text{Sentado} + \beta_5 \text{Ligero} \cdot \text{Reposo} + \beta_6 \text{Sentado} \cdot \text{Reposo} + \varepsilon$$

The regression equation is

$$\text{Heavy} = 265.6 + 1.48 \text{ Light} - 10.4 \text{ Sit} - 1.13 \text{ Rest} + 0.044 \text{ Light*Sit} + 0.0010 \text{ Light*Rest} + 0.026 \text{ Sit*Rest}$$

Predictor	Coef	SE Coef	T	P
Constant	265.64	10.59	25.08	0.000
Light	1.4755	0.2632	5.61	0.000
Sit	-10.410	1.660	-6.27	0.000
Rest	-1.1338	2.233	-0.51	0.622
Light*Sit	0.043728	0.01513	2.89	0.015
Light*Rest	0.00099612	0.01855	0.05	0.958
Sit*Rest	0.026348	0.02196	1.20	0.255

De los siguientes, ¿cuál es el mejor paso siguiente en el análisis?

- Eliminar Reposo, Ligero · Reposo, y Sentado · Reposo del modelo, y después realizar una prueba F .
 - Eliminar Ligero · Sentado del modelo, y después realizar una prueba F .
 - No se tiene que hacer nada. Este modelo es excelente.
 - Agregar las variables Ligero² y Sentado² para tratar de encontrar más variables que poner en el modelo.
 - Agregar la variable Reposo² para tratar de encontrar más variables que poner en el modelo.
6. El siguiente resultado de MINITAB es para una regresión con los mejores subconjuntos que implica cinco variables dependientes $X_1, \dots, X_5$. Los dos modelos de cada tamaño con los valores más altos de R^2 están listados.

Best Subsets Regression: Y versus X1, X2, X3, X4, X5

Response is Y

Vars	R-Sq	R-Sq(adj)	Mallows		S	X	X	X	X	X
			C-p			1	2	3	4	5
1	77.3	77.1	133.6	1.4051	X					
1	10.2	9.3	811.7	2.7940				X		
2	89.3	89.0	14.6	0.97126	X			X		
2	77.8	77.3	130.5	1.3966	X	X				
3	90.5	90.2	3.6	0.91630	X			X	X	
3	89.4	89.1	14.6	0.96763	X	X		X		
4	90.7	90.3	4.3	0.91446	X	X		X	X	
4	90.6	90.2	5.3	0.91942	X			X	X	X
5	90.7	90.2	6.0	0.91805	X	X		X	X	X

- a) ¿Qué variables están en el modelo seleccionado por el criterio del mínimo C_p ?
- b) ¿Qué variables están en el modelo seleccionado por el criterio del R^2 ajustado?
- c) ¿Existen algunos otros modelos buenos?
7. Supuestamente el siguiente es el resultado de la regresión con los mejores subconjuntos que implica las cinco variables independientes $X_1, \dots, X_5$. Los dos modelos de cada tamaño con los valores más altos de R^2 están listados. Algo está equivocado. ¿Qué es?

Best Subsets Regression

Response is Y

Vars	R-Sq	Adj. R-Sq	C-p	s	X 1	X 2	X 3	X 4	X 5
1	69.1	68.0	101.4	336.79					X
1	60.8	59.4	135.4	379.11		X			
2	80.6	79.2	55.9	271.60	X		X		
2	79.5	77.9	60.7	279.59	X	X			
3	93.8	92.8	13.4	184.27	X	X		X	
3	93.7	92.7	18.8	197.88	X	X	X		
4	91.4	90.4	5.5	159.59	X	X		X	X
4	90.1	88.9	5.6	159.81	X	X	X		X
5	94.2	93.0	6.0	157.88	X	X	X	X	X

8. Cierta proceso químico se realizó 12 veces, con las concentraciones diferentes de tres reactivos. Las concentraciones se denotan por x_1, x_2 y x_3 . El modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3 + \varepsilon$ se ajusta a los datos, y la suma de los cuadrados del error es $SSE = 4.6409$. Entonces se ajusta el modelo reducido $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$, y la suma de los cuadrados del error es $SSE = 11.5820$. ¿Es razonable utilizar el modelo reducido, en vez del que contiene todas las interacciones, para pronosticar la producción? Explique.
9. (Continúa el ejercicio 7 de la sección 8.1.) Para tratar de mejorar el pronóstico de FEV_1 , se incluyeron variables independientes adicionales en el modelo. Estas nuevas variables son el peso (en kg), el producto (interacción) de la estatura y el peso, y de la temperatura ambiente (en °C). El siguiente resultado de MINITAB presenta los resultados del ajuste del

$$FEV_1 = \beta_0 + \beta_1 \text{Último } FEV_1 + \beta_2 \text{Sexo} + \beta_3 \text{Altura} + \beta_4 \text{Peso} + \beta_5 \text{Altura} \cdot \text{Peso} + \beta_6 \text{Temperatura} + \beta_7 \text{Presión} + \varepsilon$$

The regression equation is

$$FEV1 = -0.257 + 0.778 \text{ Last FEV} - 0.105 \text{ Gender} + 1.213 \text{ Height} - 0.00624 \text{ Weight} + 0.00386 \text{ Height*Weight} - 0.00740 \text{ Temp} - 0.00148 \text{ Pressure}$$

Predictor	Coef	SE Coef	T	P
Constant	0.2565	0.7602	0.34	0.736
Last FEV	0.77818	0.05270	14.77	0.000
Gender	-0.10479	0.03647	-2.87	0.005
Height	1.2128	0.4270	2.84	0.005
Weight	-0.0062446	0.01351	-0.46	0.645
Height*Weight	0.0038642	0.008414	0.46	0.647

Temp - 0.007404 0.009313 - 0.79 0.428
 Pressure - 0.0014773 0.0005170 - 2.86 0.005
 S = 0.22189 R-Sq = 93.5% R-Sq(adj) = 93.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	111.35	15.907	323.06	0.000
Residual Error	157	7.7302	0.049237		
Total	164	119.08			

- a) El siguiente resultado de MINITAB, reproducido del ejercicio 7 de la sección 8.1, es para un modelo reducido en el cual Peso, Estatura · Peso y Temperatura se tienen que eliminar. Calcule el estadístico F para probar la credibilidad del modelo reducido.

The regression equation is
 FEV1 = -0.219 + 0.779 Last FEV - 0.108 Gender + 1.354 Height - 0.00134 Pressure

Predictor	Coef	SE Coef	T	P
Constant	-0.21947	0.4503	-0.49	0.627
Last FEV	0.779	0.04909	15.87	0.000
Gender	-0.10827	0.0352	-3.08	0.002
Height	1.3536	0.2880	4.70	0.000
Pressure	-0.0013431	0.0004722	-2.84	0.005

S = 0.22039 R-Sq = 93.5% R-Sq(adj) = 93.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	111.31	27.826	572.89	0.000
Residual Error	160	7.7716	0.048572		
Total	164	119.08			

- b) ¿Cuántos grados de libertad tiene el estadístico F ?
- c) Determine el P -valor para el estadístico F . ¿El modelo reducido es creíble?
- d) Alguien afirma que puesto que cada una de las variables que se ha eliminado tenía P -valores grandes, el modelo reducido debe ser creíble, y no era necesario realizar una prueba F . ¿Esto es correcto? Explique por qué.
- e) La suma total de los cuadrados es la misma en ambos modelos, aunque las variables independientes son diferentes. ¿Hay un error? Explique.

10. El artículo "Optimization of Enterocin P Production by Batch Fermentation of *Enterococcus faecium* P 13 at Constant pH" (C. Herran, J. Martínez y cols., en *Applied Microbiology and Biotechnology*, 2001:378-383) describe un estudio que implica la tasa de crecimiento de la bacteria *Enterococcus faecium* en medios con pH diferente. El logaritmo de la tasa máxima de crecimiento para diferentes valores de pH se presenta en la tabla siguiente:

ln tasa de crecimiento	-2.12	-1.51	-0.89	-0.33	-0.05	-0.11	0.39	-0.25
pH	4.7	5.0	5.3	5.7	6.0	6.2	7.0	8.5

- a) Ajuste el modelo lineal: $\ln \text{ tasa de crecimiento} = \beta_0 + \beta_1 \text{ pH} + \varepsilon$. Para cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0. Además, calcule al análisis de la varianza tabla (ANOVA).
- b) Ajuste el modelo cuadrático: $\ln \text{ tasa de crecimiento} = \beta_0 + \beta_1 \text{ pH} + \beta_2 \text{ pH}^2 + \varepsilon$. Para cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0. Además, calcule la tabla ANOVA.
- c) Ajuste el modelo cúbico: $\ln \text{ tasa de crecimiento} = \beta_0 + \beta_1 \text{ pH} + \beta_2 \text{ pH}^2 + \beta_3 \text{ pH}^3 + \varepsilon$. Para cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0. Además, calcule la tabla ANOVA.
- d) ¿Cuál de estos modelos prefiere y por qué?
11. En relación con el ejercicio 7 de la sección 7.4, un molino de viento se utiliza para generar corriente directa. Se reúnen datos para determinar la relación entre la velocidad del viento y la corriente. Sea y la corriente y x la velocidad del viento.
- a) Ajuste el modelo cuadrático $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$. Respecto a cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0.
- b) Ajuste el modelo $y = \beta_0 + \beta_1 \ln x + \varepsilon$. Respecto a cada coeficiente, determine el P -valor para la hipótesis nula de que el coeficiente es igual a 0.
- c) Realice una gráfica de residuos contra valores ajustados para cada modelo. ¿Los supuestos de cualquier modelo parece que hayan sido violados?
- d) Utilizando cada modelo, pronostique la corriente para las velocidades del viento de 3.0, 5.0, 7.0 y 9.0 mi/h.
- e) El modelo cuadrático tiene dos variables dependientes (x y x^2), mientras que el logaritmo del modelo sólo tiene uno. ¿La prueba F se puede utilizar para determinar qué modelo es más adecuado? Si es así, calcule el estadístico F y el P -valor. Si no, explique.
- f) ¿Puede decir una razón para preferir uno de estos modelos con respecto a los demás? Explique.
12. En una roca que se detona con explosivos la velocidad máxima de la partícula (PPV, por sus siglas en inglés) depende tanto de la distancia de la explosión como de la cantidad de carga. El artículo "Prediction of Particle Velocity Caused by Blasting for an Infrastructure Excavation Covering Granite Bedrock" (A. Kahrman, en *Mineral Resources Engineering*, 2001:205-218) presenta datos de PPV, la distancia escalada (que es igual a la dividida entre la raíz cuadrada de la carga), y la cantidad de carga. La tabla siguiente presenta los valores de PPV, la distancia escalada y la cantidad de carga para 15 explosiones. (Algunos de estos datos se presentaron en el ejercicio 4 de la sección 7.4.)

PPV (mm/s)	Distancia escalada (m/kg ^{0.5})	Cantidad de carga (kg)
1.4	47.33	4.2
15.7	9.6	92
2.54	15.8	40
1.14	24.3	48.7
0.889	23	95.7
1.65	12.7	67.7
1.4	39.3	13
26.8	8	70
1.02	29.94	13.5
4.57	10.9	41
6.6	8.63	108.8
1.02	28.64	27.43
3.94	18.21	59.1
1.4	33	11.5
1.4	34	175

- a) Ajuste el modelo $\ln \text{ PPV} = \beta_0 + \beta_1 \ln \text{ Distancia escalada} + \beta_2 \ln \text{ Carga} + \varepsilon$. Calcule el P -valor para probar $H_0 : \beta_i = 0$ respecto de β_0 , β_1 y β_2 .

- b) El artículo afirma que es apropiado el modelo $\ln \text{PPV} = \beta_0 + \beta_1 \ln \text{Distancia escalada} + \varepsilon$. Ajuste este modelo. Calcule el P -valor para probar $H_0: \beta_i = 0$ respecto de β_0, β_1 y β_2 .
- c) ¿Qué modelo prefiere usted? ¿Por qué?

13. El artículo “Ultimate Load Analysis of Plate Reinforced Concrete Beams” (N. Subedi y P. Baglin, en *Engineering Structures*, 2001:1068-1079) presenta fuerzas finales teóricas y medidas (en kN) para una muestra de vigas de concreto reforzado con acero. Los resultados se presentan en la tabla siguiente (se han eliminado dos datos atípicos).

Sea y la fuerza medida, x la fuerza teórica y t la fuerza verdadera, que no se conoce. Suponga que $y = t + \varepsilon$, donde ε es el error de medición. Está poco claro si t está relacionada con x mediante un modelo lineal $t = \beta_0 + \beta_1 x$ o por un modelo cuadrático $t = \beta_0 + \beta_1 x + \beta_2 x^2$.

Teórica	Medida	Teórica	Medida
991	1118	1516	1550
785	902	1071	1167
1195	1373	1480	1609
1021	1196	1622	1756
1285	1609	2032	2119
1167	1413	2032	2237
1519	1668	660	640
1314	1491	565	530
1743	1952	738	893
791	844	682	775

- a) Ajuste el modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$. Respecto de cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0.
- b) Ajuste el modelo lineal $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$. Respecto de cada coeficiente, encuentre el P -valor para la hipótesis nula de que el coeficiente es igual a 0.
- c) Grafique los residuos contra los valores ajustados para el modelo lineal.
- d) Grafique los residuos contra los valores ajustados para el modelo cuadrático.
- e) Con base en los resultados de los incisos a) al d), ¿qué modelo parece más adecuado? Explique.
- f) Utilizando el modelo más adecuado, calcule la fuerza verdadera si la fuerza teórica es 1 500.
- g) Utilizando el modelo más adecuado, encuentre un intervalo de confianza de 95% para la fuerza verdadera si la fuerza teórica es 1 500.

14. El artículo “Permanent Deformation Characterization of Subgrade Soils from RLT Test” (A. Puppala, L. Mohammad y cols., en *Journal of Materials in Civil Engineering*, 1999:274-282) presenta mediciones de tensiones plásticas (en %) sobre suelos para diferentes esfuerzos limitantes y de desviación en kPa. La tabla siguiente presenta las tensiones plásticas (y), los esfuerzos limitantes (x_1) y el esfuerzo de desviación (x_2) para pruebas sobre una tierra arenosa.

y	x_1	x_2	y	x_1	x_2
0.01	21	21	0.01	70	140
0.02	21	35	0.07	70	210
0.05	21	52.5	0.002	105	70
0.09	21	70	0.0003	105	105
0.003	35	35	0.0009	105	140
0.006	35	70	0.01	105	210
0.05	35	105	0.001	140	70
0.23	35	140	0.0003	140	105
0.003	70	35	0.0005	140	210
0.0008	70	70	0.03	140	280

- a) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Realice la gráfica de residuos contra valores ajustados. ¿Parece apropiado el modelo?
- b) Ajuste el modelo $\ln y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \varepsilon$. Realice la gráfica de residuos contra valores ajustados. ¿Parece apropiado el modelo?
- c) Utilice el más apropiado de los modelos de los incisos a) al b) para pronosticar y cuando $x_1 = 50$ y $x_2 = 100$.
- d) ¿Se mejorará el modelo utilizado en el inciso c) al incluir un término de interacción? Explique.
15. El artículo “Vehicle-Arrival Characteristics at Urban Uncontrolled Intersections” (V. Rengaraju y V. Rao, en *Journal of Transportation Engineering*, 1995:317-323) presenta los datos sobre las características de tránsito en diez cruceiros en Madrás, India. La siguiente tabla proporciona datos sobre el ancho del camino en m (x_1), el volumen de tránsito de vehículos por carril por hora (x_2) y la velocidad media en km/h (x_3).

y	x_1	x_2	y	x_1	x_2
35.0	76	370	26.5	75	842
37.5	88	475	27.5	92	723
26.5	76	507	28.0	90	923
33.0	80	654	23.5	86	1 039
22.5	65	917	24.5	80	1 120

- a) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Encuentre los P -valores para probar que los coeficientes sean iguales a 0.
- b) Ajuste el modelo $y = \beta_0 + \beta_1 x_1 + \varepsilon$. Encuentre los P -valores para probar que los coeficientes sean iguales a 0.
- c) Ajuste el modelo $y = \beta_0 + \beta_1 x_2 + \varepsilon$. Encuentre los P -valores para probar que los coeficientes sean iguales a 0.
- d) ¿Cuál de los modelos del a) al c) considera el mejor? ¿Por qué?
16. La tabla siguiente presenta mediciones de la media de los niveles de ruido en dBA (y), el ancho de la avenida en m (x_1) y media de la velocidad en km/h (x_2) de diez ubicaciones en Bangkok, Tailandia, como se informa en el artículo “Modeling of Urban Area Stop-and-Go Traffic Noise” (P. Pamanikabud y C. Tharasawatipat, en *Journal of Transportation Engineering* 1999:152-159).

y	x_1	x_2	y	x_1	x_2
78.1	6.0	30.61	78.1	12.0	28.26
78.1	10.0	36.55	78.6	6.5	30.28
79.6	12.0	36.22	78.5	6.5	30.25
81.0	6.0	38.73	78.4	9.0	29.03
78.7	6.5	29.07	79.6	6.5	33.17

Construya un buen modelo lineal para pronosticar la media de los niveles de ruido usando el ancho de la avenida, la media de la velocidad, o ambos, como indicadores. Proporcione las desviaciones estándar de los coeficientes estimados y los P -valores para probar que son diferentes de 0. Explique cómo escogió su modelo.

17. El artículo “Modeling Resilient Modulus and Temperature Correction for Saudi Roads” (H. Wahhab, I. Asi y R. Ramadhan, en *Journal of Materials in Civil Engineering*, 2001:298-305) describe un estudio diseñado para pronosticar los módulos de resiliencia del pavimento a partir de las propiedades físicas. La tabla siguiente presenta datos para los módulos de resiliencia a 40°C en 10^6 kPa (y), el área de superficie del agregado en m^2/kg (x_1), y el punto de suavidad del asfalto en °C (x_2).

y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1.48	5.77	60.5	3.06	6.89	65.3	1.88	5.93	63.2
1.70	7.45	74.2	2.44	8.64	66.2	1.90	8.17	62.1
2.03	8.14	67.6	1.29	6.58	64.1	1.76	9.84	68.9
2.86	8.73	70.0	3.53	9.10	68.6	2.82	7.17	72.2
2.43	7.12	64.6	1.04	8.06	58.8	1.00	7.78	54.1

El modelo cuadrático completo es $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 + \varepsilon$. ¿Qué submodelo de este modelo completo cree que es el más adecuado? Justifique su respuesta ajustando dos o más modelos y compare los resultados.

18. El artículo “Models for Assessing Hoisting Times of Tower Cranes” (A. Leung y C. Tam, en *Journal of Construction Engineering and Management*, 1999:385-391) presenta un modelo construido mediante un procedimiento de regresión stepwise para pronosticar el tiempo necesario para la operación con poleas de una grúa de torre. Se consideraron veinte variables y el procedimiento stepwise eligió un modelo de nueve variables. El R^2 ajustado para el modelo seleccionado fue 0.73. Verdadero o falso:
- El valor 0.73 es una medida confiable de la bondad del ajuste del modelo seleccionado.
 - El valor 0.73 puede exagerar la bondad del ajuste del modelo.
 - Un procedimiento de regresión stepwise selecciona solamente variables de cierta utilidad para el pronóstico del valor de la variable dependiente.
 - Es posible que una variable que no se utiliza en el pronóstico del valor de una variable dependiente sea parte de un modelo seleccionado por un procedimiento de regresión stepwise.

Ejercicios adicionales para el capítulo 8

1. El artículo “Advances in Oxygen Equivalence Equations for Predicting the Properties of Titanium Welds” (D. Harwig, W. Ittiwattana y H. Castner, en *The Welding Journal*, 2001:126s-136s) describe un experimento para pronosticar diferentes propiedades de las soldaduras de titanio. Entre otras propiedades, se midió la elongación (en %), junto con los contenidos de oxígeno y de nitrógeno (ambos en %). El siguiente resultado de MINITAB presenta los resultados del ajuste del modelo.

$$\text{Elongación} = \beta_0 + \beta_1 \text{Oxígeno} + \beta_2 \text{Nitrógeno} + \beta_3 \text{Oxígeno} \cdot \text{Nitrógeno}$$

The regression equation is

$$\text{Elongation} = 46.80 - 130.11 \text{ Oxygen} - 807.1 \text{ Nitrogen} + 3580.5 \text{ Oxy*Nit}$$

Predictor	Coef	SE Coef	T	P
Constant	46.802	3.702	12.64	0.000
Oxygen	-130.11	20.467	-6.36	0.000
Nitrogen	-807.10	158.03	-5.107	0.000
Oxy*Nit	3580.5	958.05	3.737	0.001

$$S = 2.809$$

$$R\text{-Sq} = 74.5\%$$

$$R\text{-Sq}(\text{adj}) = 72.3\%$$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	805.43	268.48	34.03	0.000
Residual Error	35	276.11	7.89		
Total	38	1081.54			

- a) Pronostique la elongación para una soldadura con un contenido de oxígeno de 0.15% y otro de nitrógeno de 0.01%.
- b) Si las dos soldaduras tienen un contenido de nitrógeno de 0.006%, y su contenido de oxígeno es diferente por 0.05%, ¿cuál predeciría que sea su diferencia en la elongación?
- c) Las dos soldaduras tienen contenidos de oxígeno idénticos, y contenidos de nitrógeno diferentes por 0.005%. ¿Esta es información para pronosticar su diferencia en la elongación? Si es así, pronostique la elongación. Si no, explique qué información adicional se necesita.
2. En relación con el ejercicio 1.
- a) Determine un intervalo de confianza de 95% para el coeficiente del oxígeno.
- b) Determine un intervalo de confianza de 99% para el coeficiente de nitrógeno.
- c) Determine un intervalo de confianza de 98% para el coeficiente del término de interacción del Oxígeno · Nitrógeno.
- d) ¿Puede concluir que $\beta_1 < -75$? Encuentre el P -valor.
- e) ¿Puede concluir que $\beta_2 > -1.000$? Encuentre el P -valor.
3. El siguiente resultado de MINITAB es para una regresión múltiple. Algunos de los números aparecieron borrosos y por lo tanto son ilegibles. Complete los números faltantes.

Predictor	Coef	SE Coef	T	P
Constant	(a)	0.3501	0.59	0.568
X1	1.8515	(b)	2.31	0.040
X2	2.7241	0.7124	(c)	0.002
S = (d)	R-Sq = 83.4%		R-Sq(adj) = 80.6%	

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	(e)	(f)	(g)	(h)	0.000
Residual Error	12	17.28	1.44		
Total	(i)	104.09			

4. Una ingeniera prueba tres métodos diferentes para seleccionar un modelo lineal. Primero utiliza un método informal basado en el estadístico F , como se describió en la sección 8.3. Después realiza la rutina de los mejores subconjuntos, y encuentra el modelo con el mejor ajuste R^2 y otro con el mejor C_p de Mallows. Ello da como resultado que los tres métodos seleccionen el mismo modelo. La ingeniera dice que puesto que los tres métodos concuerdan, este modelo debe ser el mejor. Uno de sus colegas dice que otros modelos podrían ser igualmente buenos. ¿Quién tiene razón? Explique.
5. En una simulación de 30 redes de computadoras móviles, se midieron la velocidad promedio, la pausa de tiempo y el número de vecinas. Una “vecina” es una computadora dentro del rango de transmisión de otra. En la tabla siguiente se presentan los datos.

Vecinas	Velocidad	Pausa	Vecinas	Velocidad	Pausa	Vecinas	Velocidad	Pausa
10.17	5	0	9.36	5	10	8.92	5	20
8.46	5	30	8.30	5	40	8.00	5	50
10.20	10	0	8.86	10	10	8.28	10	20
7.93	10	30	7.73	10	40	7.56	10	50
10.17	20	0	8.24	20	10	7.78	20	20
7.44	20	30	7.30	20	40	7.21	20	50
10.19	30	0	7.91	30	10	7.45	30	20
7.30	30	30	7.14	30	40	7.08	30	50
10.18	40	0	7.72	40	10	7.32	40	20
7.19	40	30	7.05	40	40	6.99	40	50

- Ajuste el modelo con vecinas como la variable dependiente, y las variables independientes Velocidad, Pausa, Velocidad · Pausa, Velocidad² y Pausa².
 - Construya un modelo reducido para eliminar algunas variables cuyos P -valores son grandes, y pruebe la credibilidad del modelo con una prueba F .
 - Realice la gráfica de residuos contra valores ajustados para el modelo reducido. ¿Existen algunos indicadores de que el modelo es inadecuado? Si es así, ¿cuáles son?
 - Alguien sugiere que un modelo que contenga Pausa y Pausa² como las únicas variables dependientes es adecuado. ¿Está de acuerdo? ¿Por qué sí o no?
 - Utilizando un mejor software para determinar los mejores subconjuntos, encuentre los dos modelos con el valor R^2 más alto para cada tamaño de modelo de una a cinco variables. Calcule C_p y R^2 ajustado para cada modelo.
 - ¿Qué modelo se selecciona para C_p mínimo? ¿Para R^2 ajustado? ¿Son los mismos?
6. Los datos en la tabla SE6 constan de las mediciones de producción de muchas operaciones de una reacción química. Las cantidades cambian con la temperatura en °C (x_1), la concentración del reactivo fundamental en % (x_2), y la duración de la reacción en horas (x_3). La variable dependiente (y) es la fracción convertida al producto deseado.
- Ajuste el modelo lineal $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$.
 - De las dos variables en este modelo se tienen coeficientes significativamente diferentes de 0 al nivel 15%. Ajuste un modelo de regresión lineal que contenga estas dos variables.
 - Calcule el producto (interacción) de las dos variables que se mencionan en el inciso b). Ajuste el modelo que contiene las dos variables junto con el término de interacción.
 - Con base en los resultados de los incisos a) al c), especifique un modelo que parezca bueno para pronosticar y a partir de x_1 , x_2 y x_3 .
 - ¿Podría ser posible construir un modelo igualmente bueno de otra manera?

TABLA SE6 Datos para el ejercicio 6

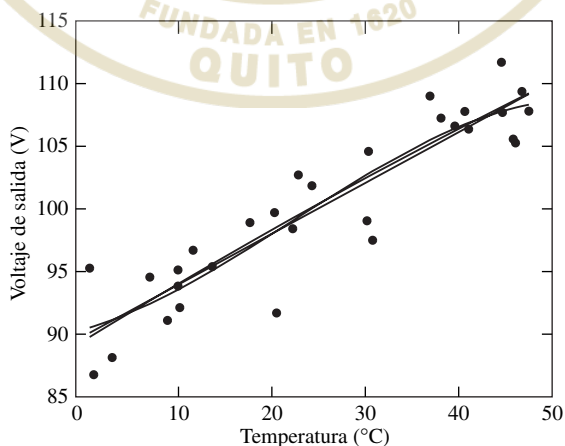
x_1	x_2	x_3	y	x_1	x_2	x_3	y	x_1	x_2	x_3	y
50	19	4.0	27.464	70	27	10.0	38.241	70	31	6.0	35.091
90	38	8.0	49.303	80	32	6.5	34.635	60	23	7.0	34.372
70	28	6.5	37.461	50	26	9.0	44.963	50	19	6.0	26.481
70	25	5.5	36.478	50	22	4.0	30.012	60	22	7.5	36.739
60	26	6.5	33.776	80	34	6.5	41.077	70	30	9.5	36.185
70	29	5.0	35.092	50	21	10.0	41.964	70	25	8.0	38.725
60	23	5.5	31.307	80	34	7.5	44.152	50	17	9.5	32.707
70	28	5.5	37.863	60	22	2.5	29.901	70	28	7.0	32.563
80	34	6.5	41.109	60	24	5.0	26.706	60	25	5.5	36.006
70	26	4.5	28.605	60	23	4.0	28.602	70	25	5.5	33.127
70	26	8.0	35.917	60	29	6.5	33.401	70	29	5.0	32.941
70	26	8.0	33.489	70	27	7.5	41.324	70	29	6.5	33.650
60	30	5.0	31.381	70	32	4.0	24.000	50	19	4.5	34.192
60	26	7.0	38.067	60	25	5.5	38.158	60	24	6.5	24.115
70	25	7.5	31.278	60	26	3.5	25.412	60	26	7.5	37.614
70	31	5.5	32.172	70	28	7.5	37.671	60	28	6.0	29.612
60	27	7.5	36.109	60	22	5.5	27.979	60	22	6.5	39.106
60	23	6.0	31.535	60	22	4.5	31.079	60	28	7.5	36.974
60	23	6.0	33.875	60	27	7.0	30.778	60	25	4.0	28.334
60	24	9.0	37.637	60	25	6.0	28.221	50	20	8.5	33.767
70	31	5.5	40.263	60	23	6.5	30.495	60	26	9.5	38.358
80	32	6.0	36.694	60	27	7.5	38.710	60	25	4.0	33.381
60	26	10.0	45.620	80	31	4.5	27.581	60	29	4.0	37.672
70	28	4.5	38.571	80	36	4.5	38.705	70	30	6.0	36.615
60	24	4.0	19.163	60	22	7.5	40.525	60	26	8.0	39.351
50	21	7.0	31.962	70	28	4.5	29.420	60	24	6.5	38.611
50	17	2.0	23.147	60	26	7.0	37.898	60	25	6.0	36.460
80	34	8.5	40.278	60	25	7.0	40.340	60	24	5.5	23.449
70	27	5.5	32.725	60	24	5.0	27.891	60	24	5.0	23.027
60	24	2.5	28.735	70	32	7.5	38.259	70	26	8.0	31.372

7. En un estudio para pronosticar la temperatura de la presión atmosférica en un dispositivo pistón-cilindro, se hicieron 19 mediciones de la temperatura en °F (y) y de la presión del aire en psi (x). Se ajustaron tres modelos: el modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$, el cuadrático $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$, y el cúbico $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon$. En la tabla siguiente se presentan los residuos y los valores ajustados para cada modelo. Realice una gráfica de los residuos contra los valores ajustados para cada modelo. Para cada uno de éstos establezca si es adecuado y explique.

Modelo lineal		Modelo cuadrático		Modelo cúbico	
Residuo	Ajuste	Residuo	Ajuste	Residuo	Ajuste
-56.2	125.6	11.2	58.2	3.3	66.1
-34.0	153.1	-7.4	126.5	-6.7	125.7
8.4	179.8	4.9	183.4	9.2	179.0
21.4	207.2	-3.6	232.2	0.9	227.8
28.6	234.7	-8.2	271.5	-5.9	269.2
46.9	260.9	8.1	299.7	7.3	300.5
47.2	288.1	15.7	319.6	12.1	323.2
8.5	314.4	-7.0	329.8	-11.7	334.6
-7.1	342.0	4.0	330.9	1.0	333.9
-47.1	139.3	-1.3	93.6	-4.1	96.4
-1.6	166.2	9.1	155.5	12.1	152.5
38.0	220.9	5.9	253.0	9.5	249.4
35.7	247.8	-3.2	286.7	-2.4	286.0
34.1	275.1	-2.1	311.3	-4.5	313.7
34.6	301.1	9.9	325.8	5.5	330.2
1.0	328.2	-2.4	331.6	-6.7	335.9
-23.2	355.3	4.1	327.9	3.5	328.5
-50.7	368.4	-5.0	322.7	-2.1	319.9
-72.9	382.1	-5.7	314.9	2.1	307.1

8. El voltaje de salida (y) de una batería se mide en un rango de temperaturas (x) desde 0°C hasta 50°C . La siguiente tabla representa un diagrama de dispersión del voltaje contra la temperatura, con tres curvas ajustadas sobrepuestas. Las curvas son el modelo lineal $y = \beta_0 + \beta_1x + \varepsilon$, el cuadrático $y = \beta_0 + \beta_1x + \beta_2x^2 + \varepsilon$, y el cúbico, $y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \varepsilon$. Con base en la gráfica, ¿cuál de los modelos se debe utilizar para describir los datos? Explique.

- El modelo lineal.
- El modelo cuadrático.
- El modelo cúbico.
- Todos parecen ser igualmente buenos.



9. En relación con el ejercicio 2 de la sección 8.2.

- Utilizando cada uno de los tres modelos, pronostique la emisión de NO_x cuando Velocidad = 1 500, Momento de torsión = 400 y HP = 150.
- Utilizando cada uno de los tres modelos, pronostique la emisión de NO_x cuando Velocidad = 1 600, Momento de torsión = 300 y HP = 100.
- Utilizando cada uno de los tres modelos, pronostique la emisión de NO_x cuando Velocidad = 1 400, Momento de torsión = 200, y HP = 75.
- ¿Cuál(es) modelo(s) parece(n) ser el(los) mejor(es)? Elija una de las respuestas y explique.
 - El modelo con Velocidad y Momento de torsión como variables independientes es el mejor.
 - El modelo con Velocidad y HP como variables independientes es el mejor.
 - El modelo con Velocidad, Momento de torsión y HP como variables independientes es el mejor.
 - El modelo con Velocidad y Momento de torsión y el modelo con Velocidad y HP son casi igualmente buenos; ambos son mejores que el modelo con Velocidad, Momento de torsión y HP.
 - El modelo con Velocidad y Momento de torsión y el modelo con Velocidad, Momento de torsión y HP son casi igualmente buenos; ambos son mejores que el modelo con Velocidad y HP.
 - El modelo con Velocidad y HP y el modelo con Velocidad, Momento de torsión y HP son casi igualmente buenos; ambos son mejores que el modelo con Velocidad y Momento de torsión.
 - Todos los modelos son casi igualmente buenos.

10. Este ejercicio ilustra una razón para las excepciones a la regla de la parsimonia (véase la p. 584).

- Un científico ajusta el modelo $Y = \beta_1 C + \varepsilon$, donde C representa la temperatura en $^{\circ}\text{C}$ y Y pueden representar cualquier resultado. Observe que el modelo no tiene intercepto. Ahora convierta $^{\circ}\text{C}$ a $^{\circ}\text{F}$ ($C = 0.556F - 17.78$). ¿Ahora el modelo tiene intercepto?
- Otro científico ajusta el modelo $Y = \beta_0 + \beta_2 C^2$, donde C y Y son como en el inciso a). Observe que el modelo tiene un término numérico cuadrático, pero ningún término lineal. Convierta $^{\circ}\text{C}$ a $^{\circ}\text{F}$ ($C = 0.556F - 17.78$). ¿Ahora el modelo tiene un término lineal?
- Suponga que x y z son dos unidades diferentes que se pueden utilizar para medir la misma cantidad, y que $z = a + bx$, donde $a \neq 0$. ($^{\circ}\text{C}$ y $^{\circ}\text{F}$ son un ejemplo.) Muestre que los modelos sin intercepto $y = \beta x$ y $y = \beta z$ no pueden ser correctos, por lo que la validez de que no tengan intercepto depende del punto cero de las unidades para la variable independiente.
- Sea x y z como en el inciso c). Muestre que los modelos $y = \beta_0 + \beta_2 x^2$ y $y = \beta_0 + \beta_2 z^2$ no pueden ser correctos, y, por tanto, la validez del modelo depende del punto cero para las unidades de la variable independiente.

11. Los datos que se presentan en la tabla siguiente dan la fuerza de tensión en psi (y) del papel como función de los porcentajes de contenido de madera dura (x).

Contenido de madera dura	Fuerza de tensión	Contenido de madera dura	Fuerza de tensión
1.0	26.8	7.0	52.1
1.5	29.5	8.0	56.1
2.0	36.6	9.0	63.1
3.0	37.8	10.0	62.0
4.0	38.2	11.0	62.5
4.5	41.5	12.0	58.0
5.0	44.8	13.0	52.9
5.5	44.7	14.0	38.2
6.0	48.5	15.0	32.9
6.5	50.1	16.0	21.9

- a) Ajuste modelos polinomiales de grados 1, 2, etcétera, para pronosticar la fuerza de tensión a partir del contenido de madera dura. Por cada uno, utilice la prueba F para comparar con el grado del modelo menos uno. Pare cuando el P -valor de la prueba F sea mayor de 0.05. ¿Cuál es el grado del modelo polinomial elegido por este método?
- b) Utilizando el modelo del inciso a), estime la concentración de madera dura que produce la fuerza de tensión más alta.
12. El artículo “Enthalpies and Entropies of Transfer of Electrolytes and Ions from Water to Mixed Aqueous Organic Solvents” (G. Hefter, Y. Marcus y W. Waghorne, en *Chemical Reviews*, 2002:2773-2836) presenta mediciones de entropía y cambios de entalpía para muchas sales bajo diferentes condiciones. La tabla siguiente presenta los resultados para las transferencias de entropía (en $\text{J/K} \cdot \text{mol}$) de agua a agua + metanol de NaCl (sal de mesa) en un rango de concentraciones de metanol:

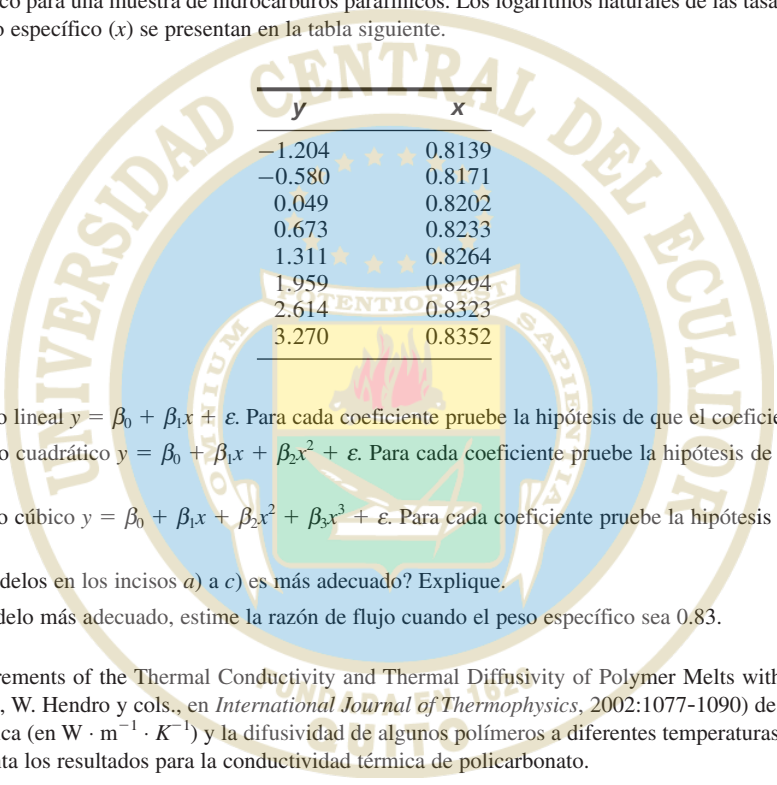
Concentración (%)	Entropía
5	1
10	-1
20	-7
30	-17
40	-28
50	-39
60	-52
70	-65
80	-80
90	-98
100	-121

- a) Ajuste modelos polinomiales de grados 1, 2 y 3 para pronosticar la entropía (y) a partir de la concentración (x).
- b) ¿Qué grado del polinomio es el más adecuado? Explique.
- c) Utilizando el modelo más adecuado, determine los intervalos de confianza de 99% para los coeficientes.
13. Una compañía de pintura recopila datos acerca de la durabilidad de su pintura y de sus competidores. Midieron los tiempos de vida de tres muestras de cada tipo de pintura en algunas ciudades estadounidenses. Los resultados se consignan en la tabla SEI3.

TABLA SEI3 Datos para el ejercicio 13

Ciudad	Temperatura promedio (°F)		Precipitación (pulg.)	Tiempo de vida (años)					
	Enero	Julio		Pintura del patrocinador			Pintura de la competencia		
Atlanta, GA	41.9	78.6	48.6	11.5	10.7	12.3	10.8	11.1	10.2
Boston, MA	29.6	73.5	43.8	11.7	10.1	12.5	10.7	11.6	11.0
Kansas City, KS	28.4	80.9	29.3	12.3	13.4	12.8	11.8	12.2	11.3
Minneapolis, MN	11.2	73.1	26.4	10.5	9.9	11.2	10.4	9.6	9.2
Dallas, TX	45.0	86.3	34.2	11.2	10.6	12.0	10.6	10.1	11.4
Denver, CO	29.5	73.3	15.3	15.2	14.2	13.8	13.4	14.4	13.2
Miami, FL	67.1	82.4	57.5	8.7	7.9	9.4	8.1	8.6	7.6
Phoenix, AZ	52.3	92.3	7.1	11.1	11.8	12.4	10.9	10.1	9.9
San Francisco, CA	48.5	62.2	19.7	16.7	17.2	15.9	15.8	15.4	14.9
Seattle, WA	40.6	65.3	38.9	14.2	14.1	13.6	12.6	13.6	14.1
Washington, DC	35.2	78.9	39.0	12.6	11.5	12.0	11.9	10.9	11.4

- a) La prueba previa sugiere que los factores más importantes que influyen en la vida de las capas de pintura son la temperatura mínima (estimado por el promedio de la temperatura de enero), la temperatura máxima (estimado por el promedio de la temperatura de julio) y la precipitación anual. Utilizando estas variables, y productos y potencias de estas variables, construya un buen modelo para predecir el tiempo de vida de la pintura del patrocinador y un buen modelo (quizás diferente) para pronosticar el tiempo de vida de la pintura del competidor.
- b) Utilizando los modelos desarrollados en el inciso a), calcule los tiempos de vida esperados para estas dos pinturas para alguien que vive en Cheyenne, Wyoming, donde la temperatura promedio de enero es 26.1°F; la de julio, 68.9°F, y la media de la precipitación anual, 13.3 pulg.
14. El artículo “Two Different Approaches for RDC Modelling When Simulating a Solvent Deasphalting Plant” (J. Aparicio, M. Heronimo y cols., *Computers and Chemical Engineering*, 2002:1369-1377) notifica la razón de flujo (en dm^3/h) y las mediciones de peso específico para una muestra de hidrocarburos parafínicos. Los logaritmos naturales de las tasas de flujo (y) y de las mediciones del peso específico (x) se presentan en la tabla siguiente.



y	x
-1.204	0.8139
-0.580	0.8171
0.049	0.8202
0.673	0.8233
1.311	0.8264
1.959	0.8294
2.614	0.8323
3.270	0.8352

- a) Ajuste al modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- b) Ajuste al modelo cuadrático $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- c) Ajuste el modelo cúbico $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- d) ¿Cuál de los modelos en los incisos a) a c) es más adecuado? Explique.
- e) Mediante el modelo más adecuado, estime la razón de flujo cuando el peso específico sea 0.83.
15. El artículo “Measurements of the Thermal Conductivity and Thermal Diffusivity of Polymer Melts with the Short-Hot-Wire Method” (X. Zhang, W. Hendro y cols., en *International Journal of Thermophysics*, 2002:1077-1090) describe mediciones de conductividad térmica (en $\text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$) y la difusividad de algunos polímeros a diferentes temperaturas (en 1 000°C). La siguiente tabla presenta los resultados para la conductividad térmica de policarbonato.

Conductividad	Temperatura	Conductividad	Temperatura	Conductividad	Temperatura	Conductividad	Temperatura
0.236	0.028	0.259	0.107	0.254	0.159	0.249	0.215
0.241	0.038	0.257	0.119	0.256	0.169	0.230	0.225
0.244	0.061	0.257	0.130	0.251	0.181	0.230	0.237
0.251	0.083	0.261	0.146	0.249	0.204	0.228	0.248

- a) Denotando la conductividad por y y la temperatura por x , ajuste el modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- b) Ajuste el modelo cuadrático $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.

- c) Ajuste el modelo cúbico $y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- d) Ajuste el modelo cuártico $y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \beta_4x^4 + \varepsilon$. Para cada coeficiente pruebe la hipótesis de que el coeficiente es igual a 0.
- e) ¿Cuál de los modelos de los incisos a) al d) es más adecuado? Explique.
- f) Utilizando el modelo más adecuado, calcule la conductividad a una temperatura de 120°C.

16. El artículo “Electrical Impedance Variation with Water Saturation in Rock” (Q. Su, Q. Feng y Z. Shang, en *Geophysics*, 2000:68-75) informa mediciones de permeabilidades (en $10^{-3} \mu\text{m}^2$), porosidades (en %), y área superficial por unidad de volumen del espacio del poro (en 10^4 cm^{-1}) para algunas muestras de roca. Los resultados se presentan en la tabla siguiente, denotando ln Permeabilidad por y , la porosidad por x_1 , y el área superficial por unidad de volumen por x_2 .

y	x_1	x_2	y	x_1	x_2
-0.27	19.83	9.55	0.58	10.52	20.03
2.58	17.93	10.97	-0.56	18.92	13.10
3.18	21.27	31.02	-0.49	18.55	12.78
1.70	18.67	28.12	-0.01	13.72	40.28
-1.17	7.98	52.35	-1.71	9.12	53.67
-0.27	10.16	32.82	-0.12	14.39	26.75
-0.53	17.86	57.66	-0.92	11.38	75.62
-0.29	13.48	21.10	2.18	16.59	9.95
4.94	17.49	9.15	4.46	16.77	7.88
1.94	14.18	11.72	2.11	18.55	88.10
3.74	23.88	5.43	-0.04	18.02	10.95

- a) Ajuste el modelo $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon$. Calcule la tabla de análisis de varianza.
- b) Ajuste el modelo $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \varepsilon$. Calcule la tabla de análisis de varianza.
- c) Ajuste el modelo $y = \beta_0 + \beta_1x_1 + \varepsilon$. Calcule la tabla de análisis de varianza.
- d) Calcule los estadísticos F para comparar los modelos en los incisos b) y c) con el modelo del inciso a). ¿Qué modelo prefiere usted? ¿Por qué?

17. El artículo “Groundwater Electromagnetic Imaging in Complex Geological and Topographical Regions: A Case Study of a Tectonic Boundary in the French Alps” (S. Houtot, P. Tarits y cols., *Geophysics*, 2002:1048-1060) presenta mediciones de concentraciones de algunos químicos (en mmol/L) y la conductividad eléctrica (en 10^{-2} S/m) para varias muestras de agua en diferentes ubicaciones cerca del Lago Gittaz, en los Alpes franceses. En la tabla siguiente se presentan los resultados para el magnesio y el calcio. Se han eliminado dos datos atípicos.

Conductividad	Magnesio	Calcio	Conductividad	Magnesio	Calcio
2.77	0.037	1.342	1.10	0.027	0.487
3.03	0.041	1.500	1.11	0.039	0.497
3.09	0.215	1.332	2.57	0.168	1.093
3.29	0.166	1.609	3.27	0.172	1.480
3.37	0.100	1.627	2.28	0.044	1.093
0.88	0.031	0.382	3.32	0.069	1.754
0.77	0.012	0.364	3.93	0.188	1.974
0.97	0.017	0.467	4.26	0.211	2.103

- a) Pronostique la conductividad (y) de las concentraciones de magnesio (x_1) y del calcio (x_2), ajuste al modelo cuadrático completo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + \varepsilon$. Calcule la tabla de análisis de varianza.
- b) Utilice la prueba F para investigar algunos submodelos del modelo cuadrático completo. Establezca qué modelo prefiere y por qué.
- c) Utilice la rutina de mejores subconjuntos para encontrar los submodelos con el máximo R^2 ajustado y el mínimo C_p de Mallows. ¿Son el mismo modelo? Comente sobre lo adecuado de este(estos) modelo(s).

18. El artículo “Low-Temperature Heat Capacity and Thermodynamic Properties of 1,1,1-trifluoro-2,2-dichloroethane” (R. Varushchenko y A. Druzhinina, en *Fluid Phase Equilibria*, 2002:109-119) presenta mediciones de la capacidad calorífica molar (y) de 1,1,1-trifluoro-2,2-dicloroetano (en $\text{J} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$) a diferentes temperaturas (x) en unidades de 10 K. Los resultados para cada décima medición se presentan en la tabla siguiente.

y	x	y	x
5.7037	1.044	60.732	6.765
16.707	1.687	65.042	7.798
29.717	2.531	71.283	9.241
41.005	3.604	75.822	10.214
48.822	4.669	80.029	11.266
55.334	5.722		

- a) Ajuste al modelo lineal simple $y = \beta_0 + \beta_1 x + \varepsilon$. Realice una gráfica de residuos y el comentario sobre lo adecuado del modelo.
- b) Ajuste al modelo lineal simple $y = \beta_0 + \beta_1 \ln x + \varepsilon$. Realice una gráfica de residuos y comente sobre lo adecuado del modelo.
- c) Calcule los coeficientes y sus desviaciones estándar para polinomios de grados 2, 3, 4 y 5. Realice gráficas de residuos para cada uno.
- d) El artículo citado al principio de este ejercicio recomienda el modelo cuártico $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \varepsilon$. ¿Le parece razonable? ¿Por qué sí o no?
19. El artículo “Lead Dissolution from Lead Smelter Slags Using Magnesium Chloride Solutions” (A. Xenidis, T. Lillis y I. Hallikia) analiza una investigación de las tasas de filtración de plomo en soluciones de cloruro de magnesio. Los datos en la tabla siguiente (leídos de una gráfica) presentan el porcentaje de plomo que se ha extraído en diferentes tiempos (en minutos).

Tiempo (t)	4	8	16	30	60	120
Porcentaje extraído (y)	1.2	1.6	2.3	2.8	3.6	4.4

- a) El artículo sugiere ajustar estos datos a un modelo cuadrático $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon$. Ajuste este modelo y calcule las desviaciones estándar de los coeficientes.
- b) La tasa reacción al tiempo t está dada por la derivada $dy/dt = \beta_1 + 2\beta_2 t$. Calcule el tiempo en que la tasa de reacción será igual a 0.05.
- c) La tasa de reacción en $t = 0$ es igual a β_1 . Determine un intervalo de confianza de 95% para la tasa de reacción en $t = 0$.
- d) ¿Puede concluir que la tasa de reacción está disminuyendo con el tiempo? Explique.
20. El artículo “The Ball-on-Three-Ball Test for Tensile Strength: Refined Methodology and Results for Three Hohokam Ceramic Types” (M. Beck, en *American Antiquity*, 2002:558-569) describe la resistencia de las cerámicas antiguas. La tabla siguiente presenta pesos medidos (en g), espesores (en mm), y las cargas (en kg) requeridas para agrietar un ejemplar de una colección de ejemplares fechados entre 1100 y 1300 que provienen del río Middle Gila, en Arizona.

Peso (x_1)	Espesor (x_2)	Carga (y)
12.7	5.69	20
12.9	5.05	16
17.8	6.53	20
18.5	6.51	36
13.4	5.92	27
15.2	5.88	35
13.2	4.09	15
18.3	6.14	18
16.2	5.73	24
14.7	5.47	21
18.2	7.32	30
14.8	4.91	20
17.7	6.72	24
16.0	5.85	23
17.2	6.18	21
14.1	5.13	13
16.1	5.71	21

a) Ajuste al modelo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$.

b) Elimine la variable cuyo coeficiente tiene el mayor P -valor, y repita el ajuste.

c) Realice la gráfica de residuos contra valores ajustados del modelo en el inciso b). ¿Hay señales de que el modelo es inadecuado?

21. El modelo lineal de partes en pasos: Sea $\tilde{x}$ una constante conocida, y suponga que una variable dependiente y está relacionada con una variable independiente x_1 , como se muestra a continuación:

$$y = \begin{cases} \beta_0 + \beta_1 x_1 + \varepsilon & \text{si } x_1 \leq \tilde{x} \\ \beta_0^* + \beta_1^* x_1 + \varepsilon & \text{si } x_1 > \tilde{x} \end{cases}$$

En otras palabras, y y x_1 están relacionados linealmente, pero son adecuadas diferentes rectas si $x_1 \leq \tilde{x}$ o $x_1 > \tilde{x}$. Defina una nueva variable independiente x_2 por

$$x_2 = \begin{cases} 0 & \text{si } x_1 \leq \tilde{x} \\ 1 & \text{si } x_1 > \tilde{x} \end{cases}$$

También defina $\beta_2 = \beta_0^* - \beta_0$ y $\beta_3 = \beta_1^* - \beta_1$. Determine un modelo de regresión múltiple que implique a y , x_1 , x_2 , β_0 , β_1 , β_2 y β_3 que exprese la relación descrita aquí.

22. El artículo “Seismic Hazard in Greece Based on Different Strong Ground Motion Parameters” (S. Koutrakis, G. Karakaisis y cols., en *Journal of Earthquake Engineering*, 2002:75-109) presenta un estudio de episodios sísmicos en Grecia durante 1978-1997. Es de interés la duración de los “fuertes movimientos de tierra”, que es el tiempo en que la aceleración de la tierra excede un valor específico. En cada episodio las mediciones de la duración de temblores fuertes de tierra se hicieron en una o más ubicaciones. La tabla SE22 de la página 618 presenta cada uno de 121 temblores medidos, los datos con el tiempo de duración y (en segundos) durante los cuales la aceleración de la tierra excedió el doble de la aceleración de la gravedad. La magnitud m del sismo, la distancia d (en km) de la medición desde el epicentro, y los dos indicadores del tipo de suelo s_1 y s_2 , definidos de la siguiente manera: $s_1 = 1$ si el suelo consta de depósitos aluviales blandos, $s_1 = 0$ de otra manera, y $s_2 = 1$ si el suelo consta de rocas terciarias o más viejas, $s_2 = 0$ de otra manera. Los casos donde tanto $s_1 = 0$ como $s_2 = 0$ corresponden a condiciones intermedias del suelo. El artículo presenta mediciones repetidas en algunas ubicaciones que no se incluyen aquí.

TABLA SE22 Datos para el ejercicio 22

y	m	d	s_1	s_2	y	m	d	s_1	s_2	y	m	d	s_1	s_2
8.82	6.4	30	1	0	4.31	5.3	6	0	0	5.74	5.6	15	0	0
4.08	5.2	7	0	0	28.27	6.6	31	1	0	5.13	6.9	128	1	0
15.90	6.9	105	1	0	17.94	6.9	33	0	0	3.20	5.1	13	0	0
6.04	5.8	15	0	0	3.60	5.4	6	0	0	7.29	5.2	19	1	0
0.15	4.9	16	1	0	7.98	5.3	12	1	0	0.02	6.2	68	1	0
5.06	6.2	75	1	0	16.23	6.2	13	0	0	7.03	5.4	10	0	0
0.01	6.6	119	0	1	3.67	6.6	85	1	0	2.17	5.1	45	0	1
4.13	5.1	10	1	0	6.44	5.2	21	0	0	4.27	5.2	18	1	0
0.02	5.3	22	0	1	10.45	5.3	11	0	1	2.25	4.8	14	0	1
2.14	4.5	12	0	1	8.32	5.5	22	1	0	3.10	5.5	15	0	0
4.41	5.2	17	0	0	5.43	5.2	49	0	1	6.18	5.2	13	0	0
17.19	5.9	9	0	0	4.78	5.5	1	0	0	4.56	5.5	1	0	0
5.14	5.5	10	1	0	2.82	5.5	20	0	1	0.94	5.0	6	0	1
0.05	4.9	14	1	0	3.51	5.7	22	0	0	2.85	4.6	21	1	0
20.00	5.8	16	1	0	13.92	5.8	34	1	0	4.21	4.7	20	1	0
12.04	6.1	31	0	0	3.96	6.1	44	0	0	1.93	5.7	39	1	0
0.87	5.0	65	1	0	6.91	5.4	16	0	0	1.56	5.0	44	1	0
0.62	4.8	11	1	0	5.63	5.3	6	1	0	5.03	5.1	2	1	0
8.10	5.4	12	1	0	0.10	5.2	21	1	0	0.51	4.9	14	1	0
1.30	5.8	34	1	0	5.10	4.8	16	1	0	13.14	5.6	5	1	0
11.92	5.6	5	0	0	16.52	5.5	15	1	0	8.16	5.5	12	1	0
3.93	5.7	65	1	0	19.84	5.7	50	1	0	10.04	5.1	28	1	0
2.00	5.4	27	0	1	1.65	5.4	27	1	0	0.79	5.4	35	0	0
0.43	5.4	31	0	1	1.75	5.4	30	0	1	0.02	5.4	32	1	0
14.22	6.5	20	0	1	6.37	6.5	90	1	0	0.10	6.5	61	0	1
0.06	6.5	72	0	1	2.78	4.9	8	0	0	5.43	5.2	9	0	0
1.48	5.2	27	0	0	2.14	5.2	22	0	0	0.81	4.6	9	0	0
3.27	5.1	12	0	0	0.92	5.2	29	0	0	0.73	5.2	22	0	0
6.36	5.2	14	0	0	3.18	4.8	15	0	0	11.18	5.0	8	0	0
0.18	5.0	19	0	0	1.20	5.0	19	0	0	2.54	4.5	6	0	0
0.31	4.5	12	0	0	4.37	4.7	5	0	0	1.55	4.7	13	0	1
1.90	4.7	12	0	0	1.02	5.0	14	0	0	0.01	4.5	17	0	0
0.29	4.7	5	1	0	0.71	4.8	4	1	0	0.21	4.8	5	0	1
6.26	6.3	9	1	0	4.27	6.3	9	0	1	0.04	4.5	3	1	0
3.44	5.4	4	1	0	3.25	5.4	4	0	1	0.01	4.5	1	1	0
2.32	5.4	5	1	0	0.90	4.7	4	1	0	1.19	4.7	3	1	0
1.49	5.0	4	1	0	0.37	5.0	4	0	1	2.66	5.4	1	1	0
2.85	5.4	1	0	1	21.07	6.4	78	0	1	7.47	6.4	104	0	0
0.01	6.4	86	0	1	0.04	6.4	105	0	1	30.45	6.6	51	1	0
9.34	6.6	116	0	1	15.30	6.6	82	0	1	12.78	6.6	65	1	0
10.47	6.6	117	0	0										

Con los datos de la tabla SE22 construya un modelo lineal para pronosticar la duración y a partir de alguna o de todas las variables m , d , s_1 y s_2 . Asegúrese de considerar las transformaciones de las variables, así como las potencias de y las interacciones entre las variables independientes. Describa sus pasos para construir su modelo. Realice una gráfica de residuos contra valores ajustados para comprobar que su modelo satisface los supuestos necesarios. Además, observe que los datos se presentan en orden cronológico, al leer hacia abajo en las columnas. Realice una gráfica para determinar si se debe incluir al tiempo como una variable independiente.

23. El artículo “Estimating Resource Requirements at Conceptual Design Stage Using Neural Networks” (A. Elazouni, I. Nosair y cols., en *Journal of Computing in Civil Engineering*, 1997:217-223) sugiere que ciertos requerimientos de recursos en la construcción de silos de concreto se puede pronosticar a partir de un modelo. Éstos incluyen la cantidad de concreto en m^3 (y), el número de personas por días de trabajo (z), o el número de horas de la mezcladora de concreto (w) necesarios para un trabajo dado. La tabla SE23A define 23 variables independientes potenciales que se pueden usar para pronosticar y , z o w . Los valores de las variables dependientes e independientes, reunidos en 28 trabajos de construcción, se presentan en las tablas SE23B de la página 620 y de la tabla SE23C en la página 621. A menos que se establezca otra situación, las longitudes están en m, las áreas en m^2 y los volúmenes en m^3 .

- Utilizando la regresión de los mejores subconjuntos, determine el mejor modelo para pronosticar y de acuerdo con el criterio de R^2 ajustado.
- Utilizando la regresión de los mejores subconjuntos, determine el mejor modelo para pronosticar y de acuerdo con el criterio del C_p mínimo de Mallows.
- Determine un modelo para pronosticar y usando la regresión paso a paso. Explique el criterio que está usando para determinar qué variables agregar o eliminar del modelo.
- Utilizando la regresión de los mejores subconjuntos, determine el mejor modelo para pronosticar z de acuerdo con el criterio del R^2 ajustado.
- Utilizando la regresión de los mejores subconjuntos, determine el mejor modelo para predecir z de acuerdo con el criterio del C_p mínimo de Mallows.
- Determine un modelo para pronosticar z usando la regresión paso a paso. Explique el criterio que está usando para determinar qué variables agregar o eliminar del modelo.
- Utilizando la regresión de los mejores subconjuntos, determine el mejor modelo para pronosticar w de acuerdo con el criterio del R^2 ajustado.
- Mediante la regresión de los mejores subconjuntos, determine el mejor modelo para pronosticar w de acuerdo con el criterio del C_p mínimo de Mallows.
- Determine un modelo para pronosticar w utilizando la regresión paso a paso. Explique el criterio que está usando para determinar qué variables agregar o eliminar del modelo.

TABLA SE23A Descripción de las variables para el ejercicio 23

x_1 Número de cubos	x_{13} Cociente amplitud-espesor
x_2 Máximo concreto requerido por hora	x_{14} Perímetro del complejo
x_3 Altura	x_{15} Capacidad de mezcla
x_4 Razón de deslizamiento de la cimbra (m/día)	x_{16} Densidad de material almacenado
x_5 Número de etapas de construcción	x_{17} Porcentaje de desechos para reforzar el acero
x_6 Perímetro de la cimbra	x_{18} Porcentaje de desechos de concreto
x_7 Volumen de complejo de silos	x_{19} Número de trabajadores en el equipo de concreto
x_8 Área superficial de las paredes del silo	x_{20} Espesores de la pared (cm)
x_9 Volumen de un cubo	x_{21} Número de trabajadores en los equipos que refuerzan el concreto
x_{10} Áreas de pared a piso	x_{22} Número de trabajadores que forman los equipos
x_{11} Número de gatos de elevación	x_{23} Cociente longitud-amplitud
x_{12} Cociente longitud-espesor	

TABLA SE23B Datos para el ejercicio 23

y	z	w	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1 850	9 520	476	33	4.5	19.8	4.0	4	223	11 072	14 751	335	26.1	72
932	4 272	268	24	3.5	22.3	4.0	2	206	2 615	8 875	109	27.9	64
556	3 296	206	18	2.7	20.3	5.0	2	130	2 500	5 321	139	28.4	48
217	1 088	68	9	3.2	11.0	4.5	1	152	1 270	1 675	141	11.6	40
199	2 587	199	2	1.0	23.8	5.0	1	79	1 370	7 260	685	17.1	21
56	1 560	120	2	0.5	16.6	5.0	1	43	275	1 980	137	22.0	15
64	1 534	118	2	0.5	18.4	5.0	1	43	330	825	165	23.6	12
397	2 660	133	14	3.0	16.0	4.0	1	240	5 200	18 525	371	12.8	74
1 926	11 020	551	42	3.5	16.0	4.0	4	280	15 500	3 821	369	12.8	88
724	3 090	103	15	7.8	15.0	3.5	1	374	4 500	5 600	300	12.2	114
711	2 860	143	25	5.0	16.0	3.5	1	315	2 100	6 851	87	24.8	60
1 818	9 900	396	28	4.8	22.0	4.0	3	230	13 500	13 860	482	17.6	44
619	2 626	202	12	3.0	18.0	5.0	1	163	1 400	2 935	115	26.4	36
375	2 060	103	12	5.8	15.0	3.5	1	316	4 200	4 743	350	11.8	93
214	1 600	80	12	3.5	15.0	4.5	1	193	1 300	2 988	105	20.6	40
300	1 820	140	6	2.1	14.0	5.0	1	118	800	1 657	133	17.0	24
771	3 328	256	30	3.0	14.0	5.0	3	165	2 800	2 318	92	19.9	43
189	1 456	91	12	4.0	17.0	4.5	1	214	2 400	3 644	200	13.6	53
494	4 160	320	27	3.3	20.0	4.5	3	178	6 750	3 568	250	14.0	44
389	1 520	95	6	4.1	19.0	4.0	1	158	2 506	3 011	401	11.8	38
441	1 760	110	6	4.0	22.0	5.0	1	154	2 568	3 396	428	14.1	35
768	3 040	152	12	5.0	24.0	4.0	1	275	5 376	6 619	448	14.5	65
797	3 180	159	9	5.0	25.0	4.0	1	216	4 514	5 400	501	14.8	52
261	1 131	87	3	3.0	17.5	4.0	1	116	1 568	2 030	522	10.5	24
524	1 904	119	6	4.4	18.8	4.0	1	190	3 291	3 572	548	9.8	42
1 262	5 070	169	15	7.0	24.6	3.5	1	385	8 970	9 490	598	12.9	92
839	7 080	354	9	5.2	25.5	4.0	1	249	5 845	6 364	649	13.9	60
1 003	3 500	175	9	5.7	27.7	4.0	1	246	6 095	6 248	677	15.1	60

TABLA SE23C Datos para el ejercicio 23

X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{21}	X_{22}	X_{23}
19.6	17.6	745	0.50	800	6.00	5.50	10	24	7	20	1.12
16.0	16.0	398	0.25	600	7.00	5.00	10	20	6	20	1.00
15.3	13.5	262	0.25	850	7.00	4.50	8	20	5	18	1.13
17.0	13.8	152	0.25	800	5.00	4.00	8	25	6	16	1.23
28.1	27.5	79	0.15	800	7.50	3.50	5	20	4	14	1.02
20.3	20.0	43	0.15	600	5.00	4.00	5	15	1	12	1.02
24.0	18.3	43	0.15	600	5.05	4.25	5	15	2	12	1.31
27.5	23.0	240	0.25	600	6.00	4.00	8	20	7	22	1.20
27.5	23.0	1121	0.25	800	8.00	4.00	10	20	9	24	1.20
21.2	18.4	374	0.75	800	5.00	3.50	10	25	12	24	1.15
10.6	10.0	315	0.50	800	6.00	4.00	10	25	11	20	1.06
20.0	20.0	630	0.50	800	7.00	5.00	10	25	9	18	1.00
13.7	13.9	163	0.25	600	6.00	4.50	8	18	11	18	1.20
20.4	20.4	316	0.50	800	6.50	3.50	10	25	6	14	1.00
13.6	10.2	193	0.50	800	5.00	3.50	10	25	4	14	1.33
13.6	12.8	118	0.25	800	5.00	3.75	8	25	6	14	1.06
13.6	9.6	424	0.25	800	5.00	3.75	8	25	6	14	1.42
18.5	16.0	214	0.50	600	6.00	4.00	8	20	4	14	1.15
19.5	16.0	472	0.25	600	6.50	4.50	10	20	3	14	1.20
21.0	12.8	158	0.50	800	5.50	3.50	6	25	8	14	1.30
20.8	16.0	154	0.50	800	7.00	4.00	8	36	8	14	1.35
23.4	17.3	275	0.50	600	7.50	5.50	8	22	11	16	1.40
16.8	15.4	216	0.50	800	8.00	5.50	8	28	12	16	1.10
26.8	17.8	116	0.25	850	6.50	3.00	6	25	5	14	1.50
23.6	16.1	190	0.50	850	6.50	4.50	5	28	9	16	1.45
23.6	16.6	385	0.75	800	8.00	6.50	15	25	16	20	1.43
25.6	16.0	249	0.50	600	8.00	5.50	12	25	13	16	1.60
22.3	14.3	246	0.50	800	8.50	6.00	8	28	16	16	1.55

24. El artículo que se cita en el ejercicio 23 presenta valores para las variables dependientes e independientes para diez trabajos de construcción adicionales. Estos valores se presentan en las tablas SE24A y SE24B de la página 622.

- Utilizando la ecuación construida en el inciso a) del ejercicio 23, pronostique la cantidad concreta (y) para cada uno de estos diez trabajos.
- Denotando los valores pronosticados por $\hat{y}_1, \dots, \hat{y}_{10}$ y los valores observados por $y_1, \dots, y_{10}$, calcule las cantidades $y_i - \hat{y}_i$. Estos son los *errores de predicción*.
- Calcule los valores ajustados $\hat{y}_1, \dots, \hat{y}_{28}$ a partir de los datos del ejercicio 23. Utilice los valores observados $y_1, \dots, y_{28}$ a partir de esos datos, calcule los residuos $y_i - \hat{y}_i$.
- En general, ¿cuáles son más grandes, los residuos o los errores de predicción? ¿Por qué será cierto esto en general?

TABLA SE24A Datos para el ejercicio 24

y	z	w	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1 713	3 400	170	6	4.2	27.0	4.0	1	179	4 200	4 980	700.0	15.1	42
344	1 616	101	3	3.4	20.0	5.0	1	133	2 255	2 672	751.5	16.7	30
474	2 240	140	3	3.4	28.0	5.0	1	116	2 396	3 259	798.8	17.0	24
1 336	5 700	190	15	7.0	26.0	3.5	1	344	12 284	9 864	818.9	16.0	86
1 916	9 125	365	18	5.6	26.5	3.5	2	307	15 435	8 140	852.5	12.4	68
1 280	11 980	599	9	2.1	28.3	4.0	1	283	8 064	8 156	896.0	14.0	68
1 683	6 390	213	12	7.9	29.0	3.5	1	361	11 364	10 486	947.0	13.4	87
901	2 656	166	6	5.4	29.5	4.5	1	193	5 592	5 696	932.0	14.8	39
460	2 943	150	3	3.0	30.0	5.0	1	118	2 943	3 540	981.0	17.2	26
826	3 340	167	6	4.9	29.8	4.5	1	211	6 000	6 293	1 000.0	15.1	50

TABLA SE24B Datos para el ejercicio 24

x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}
22.5	14.8	179	0.50	850	8.0	5.0	6	28	11	16	1.52
32.0	18.8	133	0.25	800	7.5	3.0	10	25	7	14	1.70
24.6	15.0	116	0.25	800	9.0	4.0	10	28	9	14	1.65
20.2	21.1	344	0.75	850	8.5	6.5	12	28	19	18	1.72
30.0	13.2	540	0.50	600	6.5	7.0	15	25	12	18	1.75
25.3	14.3	283	0.25	800	7.5	6.5	14	30	20	16	1.80
22.7	14.0	361	0.75	800	9.0	7.0	10	30	25	18	1.42
20.5	16.0	193	0.50	850	9.5	5.5	10	30	15	16	1.20
26.0	20.1	118	0.25	600	10.0	4.0	10	25	8	14	1.30
32.0	20.0	211	0.50	600	9.5	5.0	10	25	13	16	1.90

Capítulo 9

Experimentos factoriales

Introducción

Los experimentos son esenciales para desarrollar y mejorar los métodos científicos y de ingeniería. Sólo mediante la experimentación se pueden comparar las diferentes variantes de un método con el fin de comprobar cuál es el más efectivo. Para que sea útil, un experimento se debe diseñar adecuadamente, y los datos que se obtienen de éste se deben analizar en forma correcta. En este capítulo se analizan el diseño y el análisis de datos a partir de una clase de experimentos conocidos como **experimentos factoriales**.

9.1 Experimentos de un solo factor

Se inicia con un ejemplo. El artículo “An Investigation of the $\text{CaCO}_3\text{-CaF}_2\text{-K}_2\text{SiO}_3\text{-SiO}_2\text{-Fe}$ Flux System Using the Submerged Arc Welding Process on HSLA-100 and AISI-1081 Steels” (G. Fredrickson, tesis de maestría, Escuela de Minas de Colorado, 1992) describe un experimento en el que se prepararon flujos de soldadura con diferentes composiciones químicas. Se hicieron varias soldaduras utilizando cada flujo sobre metal con base de acero AISI-1018. La tabla 9.1 presenta los resultados de las mediciones de la dureza, en la escala de Brinell, de cinco soldaduras que usan cada uno de los cuatro flujos.

TABLA 9.1 Dureza de Brinell de soldaduras que utilizan cuatro flujos diferentes

Flujo	Valores de la muestra					Media muestral	Desviación estandar muestral
A	250	264	256	260	239	253.8	9.7570
B	263	254	267	265	267	263.2	5.4037
C	257	279	269	273	277	271.0	8.7178
D	253	258	262	264	273	262.0	7.4498

La figura 9.1 presenta diagramas de puntos para las durezas que utilizan los cuatro flujos. Cada media muestral está marcada con una “X”. Es obvio que las medias muestrales difieren. En particular, las soldaduras que se hicieron con un flujo C tienen la media muestral mayor y las que usan el flujo A tienen la más pequeña. Por supuesto, existe incertidumbre en la media muestral, y la pregunta es si ésta difiere de la otra en una cantidad mayor que la que se podría considerar para la incertidumbre de una sola cantidad. Otra manera de redactar la pregunta es: ¿Se puede concluir que hay diferencias en las medias poblacionales entre los cuatro tipos de flujos?

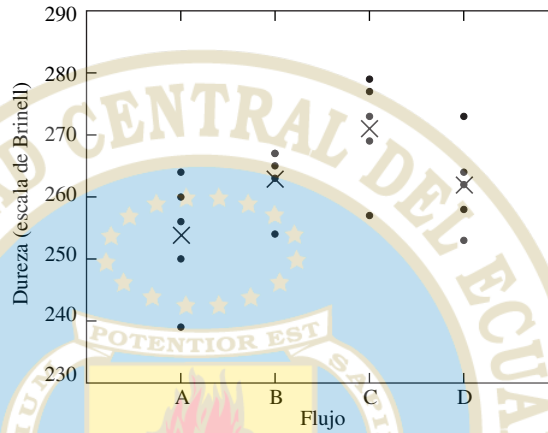


FIGURA 9.1 Diagramas de puntos para cada muestra en la tabla 9.1. Cada media muestral está marcada con una “X”. Las medias muestrales son un poco diferentes, pero los valores de muestra se traslapan considerablemente.

Éste es un ejemplo de un experimento factorial. En general, un experimento factorial implica varias variables. Una de éstas es la **variable respuesta**, que en ocasiones se llama **variable de salida** o **variable dependiente**. Las otras se denominan **factores**. La pregunta indicada por un experimento factorial es si la variación de los niveles de los factores produce una diferencia en la media de la variable respuesta. En el experimento descrito en la tabla 9.1, la dureza es la respuesta, y hay un factor: tipo de flujo. Debido a que sólo hay un factor, éste es un **experimento de un solo factor**. Hay cuatro valores diferentes para el factor tipo de flujo en este experimento. Estos valores diferentes se denominan **niveles** del factor y también **tratamientos**. Por último, los objetos con los que se realizaron dichas mediciones se llaman **unidades experimentales**. Las unidades asignadas a un tratamiento dado se llaman **réplicas**. En el experimento anterior, las soldaduras son las unidades experimentales, y hay cinco réplicas para cada tratamiento.

En este experimento de soldaduras, las cuatro composiciones particulares del flujo las eligió de manera intencional el experimentador, más que en forma aleatoria de una gran población de flujos. Se dice que este experimento sigue un **modelo de efectos fijos**. En algunos experimentos, los tratamientos se eligen aleatoriamente de una población de tratamientos posibles. En este caso se dice que el experimento sigue un **modelo de efectos aleatorios**. Los métodos de análisis para estos dos modelos son los mismos, aunque las conclusiones se pueden obtener a partir de sus diferencias. Se centrará la atención en los modelos de efectos fi-

jos. Posteriormente se analizarán algunas de las diferencias entre modelos de efectos fijos y aleatorios.

Experimentos completamente aleatorios

En este experimento de soldaduras se produjeron 20 soldaduras, cinco con cada uno de los cuatro flujos. Cada soldadura se produjo sobre una placa de base de acero diferente. Por tanto, para operar el experimentador tuvo que elegir, de un total de 20 placas de base, un grupo de cinco soldaduras con flujo A, otro grupo de cinco con soldaduras con flujo B, y así sucesivamente. El mejor camino para asignar las placas base para los flujos es en forma aleatoria. De este modo, el diseño experimental no favorecerá algún tratamiento sobre otro. Por ejemplo, el experimentador podría numerar las placas de la 1 a la 20, y luego generar un ordenamiento aleatorio de los enteros del 1 al 20. Las placas cuyos números corresponden a los primeros cinco números de la lista se les asigna el flujo A, y así sucesivamente. Éste es un ejemplo de un **experimento completamente aleatorio**.

Definición

Un experimento factorial en el cual las unidades experimentales son asignadas a los tratamientos en forma aleatoria, con todas las asignaciones posibles siendo igualmente probables, se llama **experimento completamente aleatorio**.

En muchas situaciones, en los resultados de un experimento puede incidir el orden en que se tomaron las observaciones. Por ejemplo, el desempeño de una máquina para hacer mediciones podría cambiar con el tiempo, debido, por ejemplo, a un movimiento de calibración, o a efectos de calentamiento. En casos como éstos, el procedimiento ideal es hacer observaciones en orden aleatorio. Esto último requiere cambiar de tratamiento a tratamiento conforme transcurren las observaciones, en lugar de hacer las que corresponden a un tratamiento específico en forma consecutiva. En algunos casos los cambios en los tratamientos implican tiempos o costos considerables, por lo que no es factible cambiar entre tratamientos. En situaciones de este tipo los tratamientos se deben realizar en un orden aleatorio, con todas las observaciones que corresponden al primer tratamiento elegido de forma aleatoria para que opere primero, y así sucesivamente.

En un experimento completamente aleatorio, es adecuado considerar que cada tratamiento representa una población, y a las respuestas observadas para las unidades asignadas a ese tratamiento como una muestra aleatoria simple de esa población. Por consiguiente, los datos del experimento constan de algunas muestras aleatorias, cada una proveniente de una población diferente. Las medias poblacionales se llaman **medias del tratamiento**. Las preguntas de interés relacionadas con las medias de tratamiento son si son todas iguales, y si no, cuáles son diferentes, qué tan grandes son las diferencias, etcétera.

Análisis de varianza de un sentido

La determinación formal respecto a si las medias del tratamiento son diferentes requiere una prueba de hipótesis. Se inicia introduciendo la notación. Se tienen I muestras, cada una proveniente de un tratamiento diferente. Las medias del tratamiento se denotan por

$$\mu_1, \dots, \mu_I$$

No es necesario que los tamaños de muestra sean iguales, aunque es deseable, como se analizará posteriormente en esta sección. Los tamaños de muestra se denotan por

$$J_1, \dots, J_I$$

El número total en todas las muestras combinadas se denota por N .

$$N = J_1 + J_2 + \dots + J_I$$

Las hipótesis que se desea probar son

$$H_0 : \mu_1 = \dots = \mu_I \text{ contra } H_1 : \text{dos o más de las } \mu_i \text{ son diferentes}$$

Si había solamente dos muestras, se podría utilizar la prueba t de dos muestras (sección 6.7) para probar la hipótesis nula. Debido a que hay más de dos muestras, se utiliza un método conocido como **análisis de varianza de un sentido** (ANOVA, por sus siglas en inglés). Para definir el estadístico de prueba de un ANOVA, primero se desarrolla la notación para las observaciones muestrales. Debido a que hay varias muestras, se usa un subíndice doble para denotar las observaciones. Específicamente, sea X_{ij} la j -ésima observación en la i -ésima muestra. La media de la i -ésima muestra se denota por $\bar{X}_{i.}$.

$$\bar{X}_{i.} = \frac{\sum_{j=1}^{J_i} X_{ij}}{J_i} \quad (9.1)$$

La **gran media muestral**, denotada por $\bar{X}_{..}$, es el promedio de todos los elementos que se muestrean tomados en conjunto:

$$\bar{X}_{..} = \frac{\sum_{i=1}^I \sum_{j=1}^{J_i} X_{ij}}{N} \quad (9.2)$$

Con un poco de álgebra se puede demostrar que la gran media muestral también es un promedio ponderado de las medias de las muestras:

$$\bar{X}_{..} = \frac{\sum_{i=1}^I J_i \bar{X}_{i.}}{N} \quad (9.3)$$

Ejemplo

9.1

Respecto a los datos en la tabla 9.1, determine $I, J_1, \dots, J_I, N, X_{23}, \bar{X}_3, \bar{X}_{..}$.

Solución

Hay cuatro muestras, por lo que $I = 4$. Cada muestra contiene cinco observaciones, así $J_1 = J_2 = J_3 = J_4 = 5$. El número total de observaciones es $N = 20$. La cantidad X_{23} es la tercera observación en la segunda muestra, que es 267. La cantidad $\bar{X}_3$ es la media muestral de la tercera muestra. Este valor es $\bar{X}_3 = 271.0$. Por último, con la ecuación (9.3) se calcula la gran media muestral $\bar{X}_{..}$.

$$\begin{aligned} \bar{X}_{..} &= \frac{(5)(253.8) + (5)(263.2) + (5)(271.0) + (5)(262.0)}{20} \\ &= 262.5 \end{aligned}$$

La figura 9.2 presenta la idea detrás de un ANOVA de un sentido. La figura ilustra varias muestras hipotéticas provenientes de tratamientos diferentes, junto con sus medias muestrales y la gran media muestral. Éstas están dispersas alrededor de la gran media muestral. El ANOVA de un sentido proporciona una manera de medir esta dispersión. Si aquéllas están muy dispersas, entonces es probable que las medias del tratamiento sean diferentes, y no se rechazará H_0 .

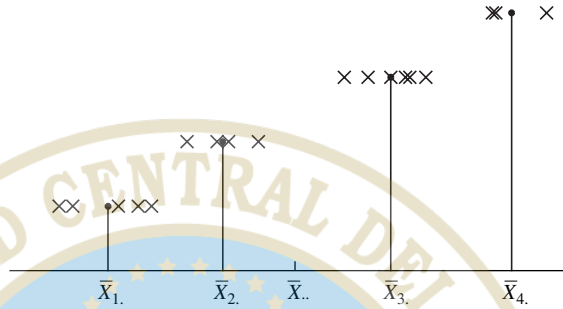


FIGURA 9.2 La variación de las medias muestrales alrededor de la gran muestra puede atribuirse tanto a la incertidumbre aleatoria como a las diferencias entre las medias del tratamiento. La variación dentro de una muestra particular alrededor de su propia media muestral sólo es atribuible a la incertidumbre aleatoria.

La variación de las medias muestrales alrededor de la gran media muestral se mide por una cantidad que se llama **suma total de cuadrados del tratamiento** (SSTr, por sus siglas en inglés), que está dada por

$$SSTr = \sum_{i=1}^I J_i (\bar{X}_{i.} - \bar{X}_{..})^2 \quad (9.4)$$

Cada término en la SSTr implica la distancia de las medias muestrales a la gran media. Observe que cada distancia al cuadrado está multiplicada por el tamaño de muestra que corresponde a su media muestral, por lo que las medias mayores para las muestras cuentan más. La SSTr indica cómo son diferentes las medias del tratamiento entre sí. Si la SSTr es grande, entonces las medias muestrales están muy dispersas, y es razonable concluir que las medias del tratamiento son diferentes y se rechaza H_0 . Si, por otro lado, la SSTr es pequeña, entonces todas las medias muestrales son cercanas a la gran muestra y, por tanto, entre sí, por lo que es creíble que las medias del tratamiento sean iguales.

Una fórmula equivalente para la SSTr, un poco más fácil de calcular a mano, es

$$SSTr = \sum_{i=1}^I J_i \bar{X}_{i.}^2 - N \bar{X}_{..}^2 \quad (9.5)$$

Para determinar si la SSTr es suficientemente grande para rechazar H_0 , se compara con otra suma de cuadrados, que se llama **suma de los cuadrados del error** (SSE, por sus siglas en inglés). La SSE mide la diferencia en los puntos muestrales individuales alrededor de sus respectivas medias muestrales. Esta diferencia se mide al sumar los cuadrados de las distancias de cada punto a su propia media muestral. La SSE está dada por

$$SSE = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2 \quad (9.6)$$

Las cantidades $X_{ij} - \bar{X}_{i.}$ se llaman **residuos**, por lo que la SSE representa la suma de los cuadrados de los residuos. La SSE, a diferencia de la SSTr, depende solamente de las distancias de los puntos muestrales a partir de sus propias medias y no la afecta la ubicación de las medias del tratamiento con respecto a otra. Por tanto, la SSE sólo mide la variación aleatoria subyacente en el proceso que se está estudiando. Es análoga a la suma de los cuadrados del error en la regresión.

Una fórmula equivalente para la SSE, que es un poco más fácil de calcular a mano, es

$$SSE = \sum_{i=1}^I \sum_{j=1}^{J_i} X_{ij}^2 - \sum_{i=1}^I J_i \bar{X}_{i.}^2 \quad (9.7)$$

Otra fórmula equivalente para la SSE que se basa en las variancias muestrales. Sea s_i^2 la variancia muestral de la i -ésima muestra. Entonces

$$s_i^2 = \frac{\sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2}{J_i - 1} \quad (9.8)$$

Por lo que de la ecuación (9.8) se tiene que $\sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2 = (J_i - 1)s_i^2$. Cuando se sustituye en la ecuación (9.6) se obtiene

$$SSE = \sum_{i=1}^I (J_i - 1)s_i^2 \quad (9.9)$$

Ejemplo

9.2

Para los datos en la tabla 9.1, calcule la SSTr y la SSE.

Solución

Las medias muestrales se presentan en la tabla 9.1. Éstas son

$$\bar{X}_{1.} = 253.8 \quad \bar{X}_{2.} = 263.2 \quad \bar{X}_{3.} = 271.0 \quad \bar{X}_{4.} = 262.0$$

En el ejemplo 9.1 se calculó la gran media muestral de $\bar{X}_{..} = 262.5$. Ahora se utiliza la ecuación (9.4) para calcular la SSTr:

$$\begin{aligned} SSTr &= 5(253.8 - 262.5)^2 + 5(263.2 - 262.5)^2 + 5(271.0 - 262.5)^2 + 5(262.0 - 262.5)^2 \\ &= 743.4 \end{aligned}$$

Para calcular a la SSE se utilizará la ecuación (9.9), ya que las desviaciones estándar muestrales s_i ya se han presentado en la tabla 9.1.

$$\begin{aligned} SSE &= (5 - 1)(9.7570)^2 + (5 - 1)(5.4037)^2 + (5 - 1)(8.7178)^2 + (5 - 1)(7.4498)^2 \\ &= 1\,023.6 \end{aligned}$$

Se pueden utilizar la SSTr y la SSE para construir un estadístico de prueba, suponiendo que se satisfacen los siguientes dos supuestos.

Supuestos para el ANOVA de un sentido

Las pruebas de hipótesis usuales del ANOVA de un sentido son válidas en las condiciones siguientes:

1. Las poblaciones en tratamiento deben ser normales.
2. Las poblaciones en tratamiento deben tener todas la misma varianza, que se denota por σ^2 .

Antes de presentar el estadístico de prueba se explicará cómo funciona. Si los dos supuestos para el ANOVA de un sentido se satisfacen aproximadamente, se pueden calcular las medias de la SSE y de la SSTR. La media de la SSTR depende de si H_0 es verdadera, porque la SSTR tiende a ser más pequeña cuando H_0 es verdadera y mayor cuando H_0 es falsa. La media de la SSTR satisface la condición

$$\mu_{SSTR} = (I - 1)\sigma^2 \quad \text{cuando } H_0 \text{ es verdadera} \quad (9.10)$$

$$\mu_{SSTR} > (I - 1)\sigma^2 \quad \text{cuando } H_0 \text{ es falsa} \quad (9.11)$$

El tamaño probable de la SSE, y por tanto, su media, no depende de que H_0 sea verdadera. La media de la SSE está dada por

$$\mu_{SSE} = (N - I)\sigma^2 \quad \text{si } H_0 \text{ es o no verdadera} \quad (9.12)$$

Las deducciones de las ecuaciones (9.10) y (9.12) se presentan al final de esta sección.

Las cantidades $I - 1$ y $N - I$ son los **grados de libertad** para la SSTR y la SSE, respectivamente. Cuando una suma de cuadrados se divide entre sus grados de libertad, la cantidad obtenida se llama **media cuadrática**. La **media cuadrática de tratamiento** se denota por MSTR, y la **media cuadrática del error** se denota por MSE. Se definen por

$$MSTR = \frac{SSTR}{I - 1} \quad MSE = \frac{SSE}{N - I} \quad (9.13)$$

De las ecuaciones (9.10) a la (9.13) se tiene que

$$\mu_{MSTR} = \sigma^2 \quad \text{cuando } H_0 \text{ es verdadera} \quad (9.14)$$

$$\mu_{MSTR} > \sigma^2 \quad \text{cuando } H_0 \text{ es falsa} \quad (9.15)$$

$$\mu_{MSE} = \sigma^2 \quad \text{si } H_0 \text{ es o no verdadera} \quad (9.16)$$

Las ecuaciones (9.14) y (9.16) muestran que cuando H_0 es verdadera, la MSTR y la MSE tienen la misma media. Por tanto, cuando H_0 es verdadera, se esperaría que su cociente esté cerca de 1. Este cociente es, en efecto, el estadístico de prueba. Éste para probar $H_0 : \mu_1 = \dots = \mu_I$ es

$$F = \frac{MSTR}{MSE} \quad (9.17)$$

Cuando H_0 es verdadera, el numerador y el denominador de F son, en promedio, del mismo tamaño, por lo que F tiende a estar cerca de 1. En efecto, cuando H_0 verdadera, su estadístico de prueba tiene una distribución F con $I - 1$ y $N - I$ grados de libertad, que se denota por $F_{I-1, N-I}$. Cuando H_0 es falsa, la MSTR tiende a ser mayor, pero la MSE no, por lo que F tiende a ser mayor que 1.

Resumen

La prueba F para el ANOVA de un sentido

Para probar $H_0: \mu_1 = \dots = \mu_I$ contra H_0 : dos o más de las μ_i son diferentes:

1. Calcule $SSTr = \sum_{i=1}^I J_i (\bar{X}_{i.} - \bar{X}_{..})^2 = \sum_{i=1}^I J_i \bar{X}_{i.}^2 - N \bar{X}_{..}^2$.
2. Calcule $SSE = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2 = \sum_{i=1}^I \sum_{j=1}^{J_i} X_{ij}^2 - \sum_{i=1}^I J_i \bar{X}_{i.}^2$
 $= \sum_{i=1}^I (J_i - 1) s_i^2$.
3. Calcule $MSTr = \frac{SSTr}{I - 1}$ y $MSE = \frac{SSE}{N - I}$.
4. Calcule el estadístico de prueba: $F = \frac{MSTr}{MSE}$.
5. Encuentre el P -valor consultando la tabla F (tabla A.7 en el Apéndice A) con $I - 1$ y $N - I$ grados de libertad.

Ahora se aplica el método de análisis de varianza al ejemplo con el que se introdujo esta sección.

Ejemplo

9.3

Para los datos en la tabla 9.1, calcule la $MSTr$, la MSE y F . Determine el P -valor para probar la hipótesis nula de que todas las medias son iguales. ¿Qué concluye?

Solución

Del ejemplo 9.2, $SSTr = 743.4$ y $SSE = 1\,023.6$. Se tienen $I = 4$ muestras y $N = 20$ observaciones en todas las muestras tomadas en conjunto. Utilizando la ecuación (9.13),

$$MSTr = \frac{743.4}{4 - 1} = 247.8 \quad MSE = \frac{1023.6}{20 - 4} = 63.975$$

El valor del estadístico de prueba F es, por tanto,

$$F = \frac{247.8}{63.975} = 3.8734$$

Para encontrar el P -valor se consulta la tabla F (tabla A.7). Los grados de libertad son $4 - 1 = 3$ para el numerador y $20 - 4 = 16$ para el denominador. Bajo H_0 , F tiene una distribución $F_{3, 16}$. Observando la tabla F debajo de 3 y 16 grados de libertad, se encuentra que el punto superior al 5% es 3.24, y que el punto superior al 1% es 5.29. Por tanto, el P -valor se encuentra entre 0.01 y 0.05 (véase la figura 9.3; un paquete de software da un valor exacto de 0.029 con dos dígitos significativos). Es razonable concluir que las medias poblacionales no son todas iguales, y, por consiguiente, la composición del flujo afecta la dureza.

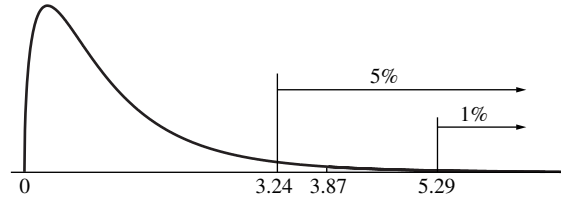


FIGURA 9.3 El valor observado del estadístico de prueba es 3.87. El punto superior al 5% de la distribución $F_{3,16}$ es 3.24. El punto superior al 1% de la distribución $F_{3,16}$ es 5.29. Por tanto, el P -valor se encuentra entre 0.01 y 0.05. Un paquete de software da un valor de 0.029.

Intervalos de confianza para medias de tratamiento

Las observaciones sobre el i -ésimo tratamiento se supone que son una muestra aleatoria simple de una población normal con media μ_i y varianza σ^2 . Para construir un intervalo de confianza para μ_i , el primer paso es estimar la varianza poblacional σ^2 . Una manera de hacer esto último sería utilizar la varianza muestral s_i^2 de las observaciones en el i -ésimo tratamiento. Sin embargo, debido a que se supone que todas las observaciones para todos los tratamientos tienen la misma varianza, es mejor combinar todas las variancias muestrales en una estimación “combinada”. Con este propósito, observe que la SSE es una suma ponderada de las variancias muestrales (ecuación 9.9) y la MSE es el promedio ponderado (ecuación 9.13). Por tanto, la cantidad MSE es la estimación combinada de la varianza σ^2 . Puesto que $\bar{X}_i$ es la media muestral de J_i observaciones, la varianza de $\bar{X}_i$ es σ^2/J_i , estimada con MSE/J_i . El número de grados de libertad para la MSE es $N - I$. La cantidad

$$\frac{\bar{X}_i - \mu_i}{\sqrt{MSE/J_i}}$$

tiene una distribución t de Student con $N - I$ grados de libertad. Por tanto, se puede construir un intervalo de confianza para μ_i con el método descrito en la sección 5.3.

Un intervalo de confianza de nivel $100(1 - \alpha)$ para μ_i está dado por

$$\bar{X}_i \pm t_{N-I, \alpha/2} \sqrt{\frac{MSE}{J_i}} \quad (9.18)$$

Ejemplo

9.4

Determine un intervalo de confianza del 95% para la media de la dureza de soldaduras producidas con el flujo A.

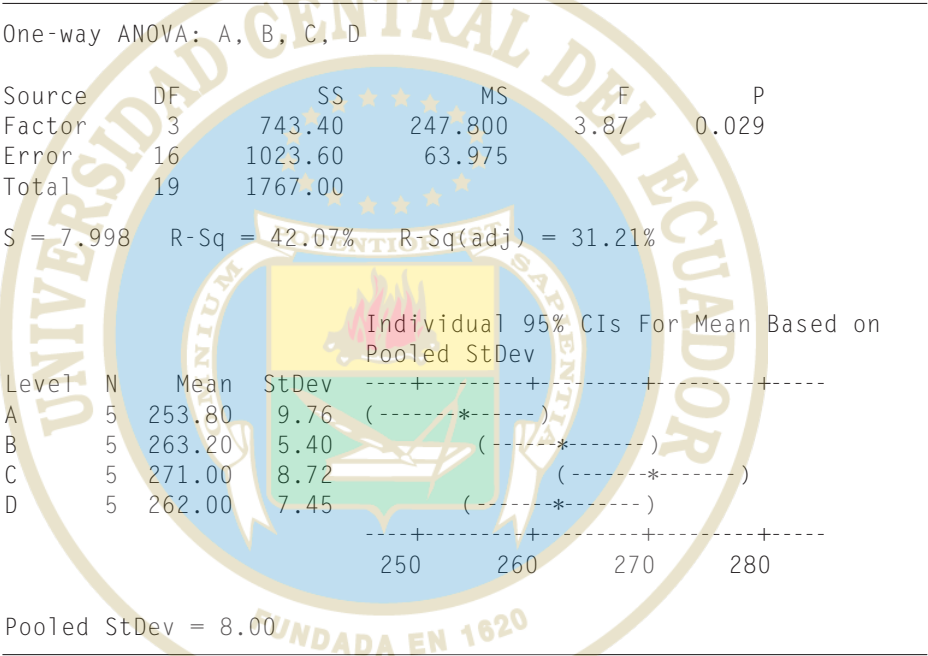
Solución

De la tabla 9.1, $\bar{X}_1 = 253.8$. El valor de la MSE se calculó en el ejemplo 9.3 de 63.975. Hay $I = 4$ tratamientos, $J_1 = 5$ observaciones para el flujo A, y $N = 20$ observaciones en conjunto. De la tabla t de Student se obtiene $t_{16, .025} = 2.120$. Por tanto, el intervalo de confianza del 95% es

$$253.8 \pm 2.120 \sqrt{\frac{63.975}{5}} = 253.8 \pm 7.6$$

La tabla ANOVA

Los resultados de un análisis de varianza generalmente se resumen en una tabla de análisis de varianza (ANOVA). Esta tabla es muy parecida a la que se obtiene en la regresión múltiple. El siguiente resultado de MINITAB muestra el análisis de varianza para los datos de soldaduras que se presentan en la tabla 9.1.



En la tabla ANOVA, la columna “DF” presenta el número de grados de libertad tanto para el tratamiento (“Factor”) como para la suma de los cuadrados del error (“Error”). La columna “SS” presenta la SSTr (en el renglón “Factor”) y la SSE (en el renglón “Error”). El renglón “Total” contiene la **suma total de los cuadrados**, que es la suma de SSTr y SSE. La columna “MS” presenta las medias cuadráticas MSTr y MSE. La columna “F” presenta el estadístico *F* para probar la hipótesis nula de que todas las medias poblacionales son iguales. Por último, la columna “P” representa el *P*-valor para la prueba *F*. Debajo de la tabla ANOVA, el valor “S” es la estimación combinada de la desviación estándar del error σ , que se calcula al tomar la raíz cuadrada de la MSE. La cantidad “R-sq” es R^2 , el coeficiente de determinación, que es igual al cociente SSTr/TSS. Esto es análogo al caso de la regresión múltiple (véase la ecuación 8.9 de la sección 8.1). El valor “R-Sq(adj)” es el R^2 ajustado, igual

a $R^2 - [(I - 1)/(N - I)](1 - R^2)$, nuevamente análogo a la regresión múltiple. Las cantidades R^2 y R^2 ajustadas no son muy utilizadas en el análisis de varianza como en la regresión múltiple. Por último, se presentan las medias muestrales y las desviaciones estándar para cada grupo de tratamiento, junto con una gráfica que muestra un intervalo de confianza del 95% para cada media de tratamiento.

Ejemplo

9.5

En el artículo “Review of Development and Application of CRSTER and MPTER Models” (R. Wilson, *Atmospheric Environment*, 1993:41-57) se presentan varias mediciones de las concentraciones máximas por hora (en $\mu\text{g}/\text{m}^3$) de SO_2 para cada una de las cuatro plantas de energía. Los resultados son los siguientes (se han eliminados dos datos atípicos):

Planta 1:	438	619	732	638		
Planta 2:	857	1 014	1 153	883	1 053	
Planta 3:	925	786	1 179	786		
Planta 4:	893	891	917	695	675	595

El siguiente resultado de MINITAB presenta resultados para un ANOVA de un sentido. ¿Puede concluir que las máximas concentraciones por hora difieren entre las plantas?

One-way ANOVA: Plant 1, Plant 2, Plant 3, Plant 4

Source	DF	SS	MS	F	P
Plant	3	378610	126203	6.21	0.006
Error	15	304838	20323		
Total	18	683449			

S = 142.6 R-Sq = 55.40% R-Sq(adj) = 46.48%

				Individual 95% CIs For Mean Based on Pooled StDev	
Level	N	Mean	StDev	-----+-----+-----+-----+-----+-----+-----	
1	4	606.8	122.9	(-----*-----)	
2	5	992.0	122.7		(-----*-----)
3	4	919.0	185.3		(-----*-----)
4	6	777.7	138.8	(-----*-----)	
				-----+-----+-----+-----+-----+-----+-----	
				600 800 1000 1200	

Pooled StDev = 142.6

Solución

En la tabla ANOVA el P -valor para la hipótesis nula es que todas las medias de tratamiento son iguales a 0.006. Por tanto, se concluye que no todas las medias de tratamiento son iguales.

Comprobación de supuestos

Como se mencionó antes, los métodos de análisis de varianza requieren los supuestos que las observaciones sobre cada tratamiento sean una muestra proveniente de una población normal y que todas las poblaciones normales tengan la misma varianza. Una buena manera de comprobar el supuesto de normalidad es con una gráfica de probabilidad normal. Si los tamaños muestrales son suficientemente grandes, se puede construir una gráfica distinta de probabilidad para cada muestra. Este caso rara vez se presenta en la práctica. Cuando los tamaños de muestra no son lo suficientemente grandes para que las gráficas individuales de probabilidad proporcionen información, los residuos $X_{ij} - \bar{X}_i$, pueden graficarse todos en una sola gráfica. Cuando se satisfacen los supuestos de normalidad y de varianza constante, estos residuos estarán distribuidos normalmente con media cero y se deberán graficar aproximadamente en una línea recta. La figura 9.4 presenta una gráfica de probabilidad normal de los residuos para los datos de soldaduras de la tabla 9.1. No hay evidencias de una violación grave del supuesto de normalidad.

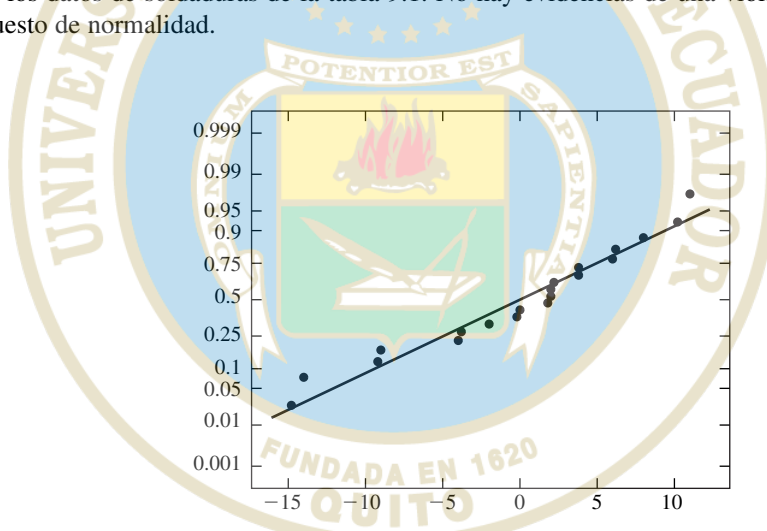


FIGURA 9.4 Gráfica de probabilidad para los residuos de los datos de soldadura. No hay evidencias de una violación grave del supuesto de normalidad.

El supuesto de variancias iguales puede ser difícil de comprobar, porque con sólo algunas pocas observaciones en cada muestra, las desviaciones estándar muestrales pueden ser muy diferentes (por un factor de 2 o más) incluso cuando vale el supuesto. Para los datos de las soldaduras, las desviaciones estándar muestrales varían de 5.4037 a 9.7570. Es razonable proseguir como si las variancias fueran iguales.

Las dispersiones de las observaciones dentro de varias muestras se pueden comprobar visualmente haciendo una gráfica de residuos. Esto último se hace al realizar la gráfica de residuos $X_{ij} - \bar{X}_i$ contra los valores ajustados, que son las medias muestrales $\bar{X}_i$. Si las disper-

siones difieren considerablemente entre las muestras, el supuesto de variancias iguales es sospechoso. Si una o más de las muestras contienen datos atípicos, el supuesto de normalidad también es dudoso. La figura 9.5 presenta una gráfica de residuos para los datos de soldadura. No hay datos atípicos graves, y las dispersiones no son muy diferentes entre las muestras.

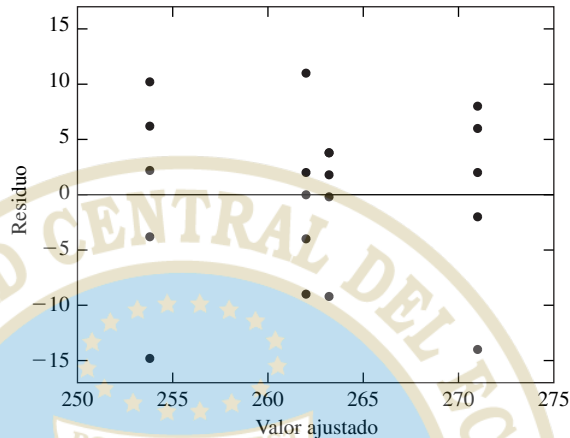


FIGURA 9.5 Gráfica de residuos de los valores $X_{ij} - \bar{X}_i$ contra $\bar{X}_i$ para los datos de soldaduras. Las dispersiones no son muy diferentes de muestra a muestra, y no hay ningún dato atípico grave.

Diseños balanceados contra no balanceados

Cuando se asignan números iguales de unidades a cada tratamiento, se dice que el diseño está **balanceado**. Aunque el **análisis de varianza** de un sentido se puede utilizar tanto con diseños balanceados como no balanceados, los primeros brindan una gran ventaja. Un diseño balanceado es mucho **menos sensible** a las violaciones del supuesto de la igualdad de la variancia que uno no balanceado. Debido a que las desviaciones moderadas de este supuesto pueden ser difíciles de detectar, es mejor utilizar un diseño balanceado siempre que sea posible, para que las violaciones no detectadas del supuesto no comprometan gravemente la validez de los resultados. Cuando es imposible obtener un diseño balanceado, es preferible un diseño ligeramente desbalanceado a uno gravemente desbalanceado.

Resumen

- Con un diseño balanceado el efecto de las variancias desiguales generalmente no es grande.
- Con un diseño desbalanceado el efecto de las variancias desiguales puede ser importante.
- Entre más desbalanceado sea el diseño, mayor será el efecto de las variancias desiguales.

La identidad del análisis de la varianza

Tanto en la regresión lineal como en el análisis de varianza, una cantidad que se llama la suma total de los cuadrados se obtiene al restar la gran media muestral de cada observación, elevando al cuadrado estas desviaciones, y después sumándolas. Una identidad del análisis de varianza es una ecuación que expresa la suma total de cuadrados como una suma de las otras sumas de los cuadrados. Se han presentado identidades del análisis de varianza para la regresión lineal simple (al final de la sección 7.2) y para la regresión múltiple (ecuación 8.7 en la sección 8.1).

La suma total de los cuadrados para el ANOVA en un sentido está dada por

$$SST = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{..})^2 \quad (9.19)$$

Una fórmula equivalente está dada por

$$SST = \sum_{i=1}^I \sum_{j=1}^{J_i} X_{ij}^2 - N \bar{X}_{..}^2 \quad (9.20)$$

Al examinar las ecuaciones (9.5), (9.7) y (9.20) se encuentra que la suma total de los cuadrados es igual a la suma de los cuadrados del tratamiento más la suma de los cuadrados del error. Esta es la identidad del análisis de varianza para el análisis de varianza de un sentido.

La identidad del análisis de varianza

$$SST = SSTr + SSE \quad (9.21)$$

Una parametrización alternativa

La presentación del análisis de varianza de un sentido, como un método para comparar varias medias de tratamiento utilizando muestras aleatorias extraídas de cada población de tratamiento, es una manera natural de ver el tema. Existe otra manera de expresar estas mismas ideas, con una notación algo diferente, que a veces es útil.

Para cada observación X_{ij} , se define $\varepsilon_{ij} = X_{ij} - \mu_i$, la diferencia entre la observación y su media. Por analogía con la regresión lineal, las cantidades ε_{ij} se llaman **errores**. Es cierto que

$$X_{ij} = \mu_i + \varepsilon_{ij} \quad (9.22)$$

Ahora, puesto que X_{ij} está distribuido normalmente con media μ_i y varianza σ^2 , se tiene que ε_{ij} está distribuido normalmente con media 0 y varianza σ^2 .

En un experimento de un solo factor se tiene interés en determinar si las medias de tratamiento son todas iguales. Dando las medias del tratamiento $\mu_1, \dots, \mu_I$, la cantidad

$$\mu = \frac{1}{I} \sum_{i=1}^I \mu_i \quad (9.23)$$

representa el promedio de todas las medias del tratamiento. La cantidad μ se llama **gran media de la población**. El i -ésimo **efecto del tratamiento**, que se denota por α_i , es la diferencia entre la media del i -ésimo tratamiento y la gran media de la población:

$$\alpha_i = \mu_i - \mu \quad (9.24)$$

que se obtiene a partir de la definición de α_i en la cual $\sum_{i=1}^I \alpha_i = 0$.

Ahora se pueden descomponer las medias del tratamiento de la siguiente manera:

$$\mu_i = \mu + \alpha_i \quad (9.25)$$

Combinando las ecuaciones (9.22) y (9.25) se obtiene el **modelo del análisis de la varianza de un sentido**:

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad (9.26)$$

La hipótesis nula $H_0 : \mu_1 = \dots = \mu_I$ es equivalente a $H_0 : \alpha_1 = \dots = \alpha_I = 0$.

En el ANOVA de un sentido es posible trabajar con las medias de tratamiento μ_i , como se ha hecho, en lugar de con los efectos de tratamiento α_i . Sin embargo, en los experimentos de múltiples factores las medias de tratamiento en sí mismas no son suficientes y se deben descomponer en una manera similar a la que se describió aquí. Lo anterior se analizará más profundamente en la sección 9.3.

Potencia

Cuando se diseña un experimento factorial es importante que la prueba F tenga buena potencia; es decir, gran probabilidad de rechazar la hipótesis nula de la igualdad si en efecto todas las medias de tratamiento no son iguales. Un experimento con potencia baja no es muy aconsejable, ya que es improbable detectar una diferencia en los tratamientos, aun si existiera una. En lo que sigue, se supondrá que el experimento está balanceado y que los supuestos de normalidad y de igualdad de varianza valen. Suponga que el número de niveles se fija en I .

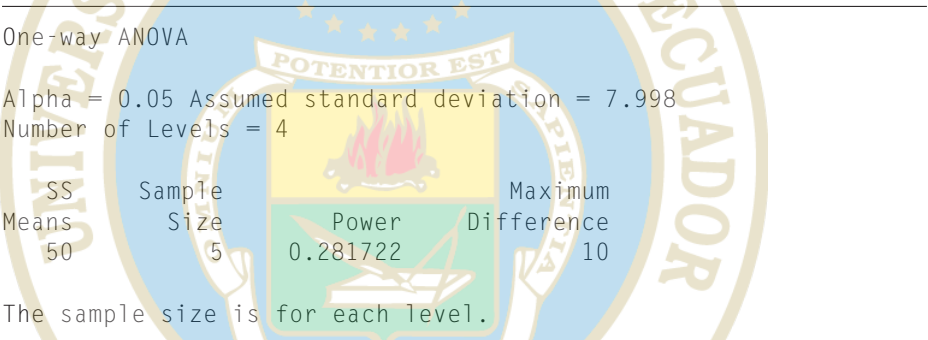
La potencia de cualquier prueba primero depende del criterio de rechazo: entre mayor sea el nivel en el que se quiere rechazar, mayor es la potencia. El nivel de 5% es el que con más frecuencia se utiliza en la práctica. En cuanto se establece el nivel de rechazo, la potencia de la prueba F depende de tres cantidades: 1) la dispersión de las medias verdaderas se mide mediante la cantidad $\sum_i \alpha_i^2$ donde α_i es el efecto del i -ésimo tratamiento, 2) la desviación estándar del error σ y 3) el tamaño muestral J . Observe que si la hipótesis nula es verdadera, entonces $\sum_i \alpha_i^2 = 0$. Entre más grande es $\sum_i \alpha_i^2$, más se aleja de la verdadera hipótesis nula, y mayor es la potencia, que es la probabilidad de que la hipótesis nula se rechace.

Un cálculo de potencia puede servir para cualquiera de dos propósitos: determinar el tamaño muestral para cada tratamiento necesario con el fin de obtener una potencia deseada, o determinar qué potencia tiene con un tamaño de muestra dado. En un cálculo de potencia tradicional, se especifica la cantidad $\sum_i \alpha_i^2$ con la que se desea detectar y el valor de σ que se espera encontrar. Entonces se puede calcular la potencia para un tamaño de muestra dado, o el tamaño de muestra necesario para obtener una potencia específica. En la práctica, rara vez se sabe cómo especificar un valor para $\sum_i \alpha_i^2$, pero se puede con frecuencia especificar el tamaño de una diferencia entre la mayor y la menor medias del tratamiento que se desea detectar. Por ejemplo, en el experimento de soldaduras, un metalúrgico es capaz de especificar que una diferencia de 10 o más entre la mayor y menor medias de tratamiento es científicamente importante, pero es improbable que haga lo mismo para $\sum_i \alpha_i^2$.

En MINITAB se puede especificar el tamaño de una diferencia científicamente importante entre las medias de tratamiento más grande y más pequeña y calcular el tamaño de muestra necesario para garantizar que la potencia para detectar la diferencia será al menos una cantidad específica. Se presenta un ejemplo.

Ejemplo
9.6

Un metalúrgico quiere repetir el experimento de soldaduras con cuatro flujos diferentes y quiere que el diseño sea bastante sensible para que sea probable detectar una diferencia de 10 o más en la dureza de Brinell con un nivel de 5%. Supone que la desviación estándar del error tiene aproximadamente el mismo valor de 7.998 calculado en el experimento que se ha estado estudiando. El siguiente resultado de MINITAB muestra un cálculo de potencia para un experimento con cinco observaciones por tratamiento. ¿Cuál es la potencia? ¿Qué le recomendaría al metalúrgico con respecto a la utilidad de este experimento propuesto?



Solución

La potencia es 0.281772. Esto último significa que la probabilidad de que el experimento propuesto detectará una diferencia de 10 entre el mayor y el menor tratamientos no podría ser mayor a 0.28. La recomendación adecuada es no operar este experimento; tiene muy poca oportunidad de éxito. En su lugar, se debe calcular el tamaño de muestra necesario para proporcionar la potencia adecuada, y, si es factible, el tamaño que debe tener un experimento operado.

Ejemplo
9.7

El metalúrgico del ejemplo 9.6 ha tomado su consejo y ha calculado el tamaño de muestra necesario para proporcionar una potencia de 0.90 y así detectar una diferencia de 10 con un nivel de 5%. En el siguiente resultado de MINITAB, ¿cuál es la potencia? ¿Cuántas observaciones serán necesarias en cada nivel? ¿Cuántas observaciones se necesitarán en total?

One-way ANOVA

Alpha = 0.05 Assumed standard deviation = 7.998

Number of Levels = 4

SS	Sample	Target		Maximum
Means	Size	Power	Actual Power	Difference
50	20	0.9	0.914048	10

The sample size is for each level.

Solución

El tamaño muestral necesario es 20 por nivel; con cuatro niveles habrá 80 observaciones en total. Observe que la potencia verdadera del experimento es aproximadamente 0.914, mayor que la “potencia objetivo” de 0.90 que fue pedida. La razón para esto último es que la potencia proporcionada por un tamaño de muestra de 19 por nivel podía ser algo menor que 0.90; un tamaño de muestra de 20 es el más pequeño que garantiza que proporciona una potencia de 0.90 o más.

Modelos de efectos aleatorios

En muchos experimentos factoriales los tratamientos los elige deliberadamente el experimentador. Se dice que estos experimentos siguen un **modelo de efectos fijos**. En algunos casos, los tratamientos se eligen en forma aleatoria de una población de tratamientos posibles. En estos casos se dice que los experimentos siguen un **modelo de efectos aleatorios**. En un modelo de efectos fijos interesan los tratamientos específicos elegidos por el experimento. En un modelo de efectos aleatorios, interesa toda la población de tratamientos posibles, y no sólo los que se eligen para el experimento.

El artículo que describe el experimento de soldaduras establece que los tratamientos fueron elegidos deliberadamente y no representa una muestra aleatoria proveniente de una población mayor de compuestos de flujo. Por tanto, este experimento sigue un modelo de efectos fijos. Las cuatro centrales hidroeléctricas en el ejemplo 9.5 son una muestra de conveniencia; son plantas en las que las mediciones estaban fácilmente disponibles. En algunos casos es adecuado tratar una muestra de conveniencia como si hubiera una muestra aleatoria simple (véase el análisis en la sección 1.1). Si estas condiciones valen, entonces se puede considerar que el experimento de la central hidroeléctrica sigue un modelo de efectos aleatorios; de otra forma, se debe tratar como un modelo de efectos fijo.

Hay una diferencia importante en la interpretación entre los resultados de un modelo de efectos fijos y la de un modelo de efectos aleatorios. En el primer modelo las únicas conclusiones que se pueden obtener son acerca de los tratamientos empleados en realidad en el experimento. Sin embargo, en un modelo de efectos aleatorios, debido a que los tratamientos son una muestra aleatoria simple de una población de tratamientos, se pueden obtener conclusiones con respecto a toda la población, incluyendo los tratamientos que no incluyen en realidad en el experimento.

Esta diferencia en las interpretaciones genera una diferencia en las hipótesis nulas que se prueban. En el modelo de efectos fijos la hipótesis nula de interés es $H_0 : \mu_1 = \cdots = \mu_L$. En el modelo de efectos aleatorios la hipótesis nula de interés es

H_0 : las medias del tratamiento son iguales para cada nivel en la población.

En el modelo de efectos aleatorios se hace el supuesto de que la población de las medias del tratamiento es normal.

Curiosamente, aunque la hipótesis nula para el modelo de efectos aleatorios difiere de la del modelo de efectos fijo, la prueba de hipótesis es exactamente la misma. La prueba F descrita antes se aplica para el modelo de efectos aleatorios y también para el modelo de efectos fijos.

Ejemplo

9.8

En el ejemplo 9.5, suponga que es razonable tratar las cuatro centrales hidroeléctricas como una muestra aleatoria proveniente de una gran población de centrales hidroeléctricas; además, suponga que las concentraciones de SO_2 en la población de plantas están distribuidas normalmente. ¿Se puede concluir que hay diferencias en las concentraciones de SO_2 entre las centrales hidroeléctricas en la población?

Solución

Éste es un modelo de efectos aleatorios, así que se puede aplicar la prueba F para probar la hipótesis nula de que todas las medias del tratamiento en la población son las mismas. Los resultados de la prueba F se muestran en el ejemplo 9.5. El P -valor es 0.006. Por tanto, se rechaza la hipótesis nula y se concluye que hay diferencias en las medias de las concentraciones de SO_2 entre las centrales hidroeléctricas en la población.

Deducciones de las ecuaciones (9.10) y (9.12)

En lo que sigue será más fácil usar la notación $E(\cdot)$ para denotar la media de una cantidad $V(\cdot)$ para denotar la varianza. Por lo que, por ejemplo, $E(\text{SSE}) = \mu_{\text{SSE}}$, $E(\text{SSTr}) = \mu_{\text{SSTr}}$, y $V(X_{ij})$ denota la varianza de X_{ij} .

Se mostrará que $E(\text{SSE}) = E[\sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_i)^2] = (N - I)\sigma^2$, si las medias de la población son o no iguales. Esta es la ecuación (9.12).

Se inicia por sumar y restar la media de tratamiento μ_i de cada término en $\sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_i)^2$ para obtener

$$\text{SSE} = \sum_{i=1}^I \sum_{j=1}^{J_i} [(X_{ij} - \mu_i) - (\bar{X}_i - \mu_i)]^2$$

Desarrollando se obtiene

$$\text{SSE} = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \mu_i)^2 - \sum_{i=1}^I \sum_{j=1}^{J_i} 2(X_{ij} - \mu_i)(\bar{X}_i - \mu_i) + \sum_{i=1}^I \sum_{j=1}^{J_i} (\bar{X}_i - \mu_i)^2 \quad (9.27)$$

Ahora $\sum_{j=1}^{J_i} (X_{ij} - \mu_i) = J_i(\bar{X}_i - \mu_i)$. Sustituyéndolo en el término de en medio del lado derecho de la ecuación (9.27) se obtiene

$$\text{SSE} = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \mu_i)^2 - 2 \sum_{i=1}^I J_i(\bar{X}_i - \mu_i)^2 + \sum_{i=1}^I \sum_{j=1}^{J_i} (\bar{X}_i - \mu_i)^2$$

Puesto que $\sum_{i=1}^I \sum_{j=1}^{J_i} (\bar{X}_i - \mu_i)^2 = \sum_{i=1}^I J_i(\bar{X}_i - \mu_i)^2$, esto simplifica a,

$$SSE = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \mu_i)^2 - \sum_{i=1}^I J_i (\bar{X}_{i.} - \mu_i)^2 \quad (9.28)$$

Tomando medias en ambos lados de la ecuación (9.28) se obtiene

$$E(SSE) = \sum_{i=1}^I \sum_{j=1}^{J_i} E(X_{ij} - \mu_i)^2 - \sum_{i=1}^I J_i E(\bar{X}_{i.} - \mu_i)^2 \quad (9.29)$$

Ahora $E(X_{ij}) = E(\bar{X}_{i.}) = \mu_i$. Todas las variancias de la población son iguales; se denota su valor común por σ^2 . De lo que se tiene que

$$\begin{aligned} E(X_{ij} - \mu_i)^2 &= V(X_{ij}) = \sigma^2 \\ E(\bar{X}_{i.} - \mu_i)^2 &= V(\bar{X}_{i.}) = \frac{\sigma^2}{J_i} \end{aligned}$$

Sustituyendo en la ecuación (9.29) se obtiene

$$E(SSE) = \sum_{i=1}^I \sum_{j=1}^{J_i} \sigma^2 - \sum_{i=1}^I \frac{J_i \sigma^2}{J_i} = N\sigma^2 - I\sigma^2 = (N - I)\sigma^2$$

Lo anterior termina la deducción de $E(SSE)$.

Ahora se muestra que $E(SSTr) = E[\sum_{i=1}^I J_i (\bar{X}_{i.} - \bar{X}_{..})^2] = (I - 1)\sigma^2$ bajo el supuesto de que las medias del tratamiento son todas iguales a un valor común denotado por μ . Esto es la ecuación (9.10).

Se inicia por sumar y restar la media común del tratamiento μ de cada término en $\sum_{i=1}^I J_i (\bar{X}_{i.} - \bar{X}_{..})^2$ para obtener

$$SSTr = \sum_{i=1}^I J_i [(\bar{X}_{i.} - \mu) - (\bar{X}_{..} - \mu)]^2$$

Desarrollando, se obtiene

$$SSTr = \sum_{i=1}^I J_i (\bar{X}_{i.} - \mu)^2 - 2 \sum_{i=1}^I J_i (\bar{X}_{i.} - \mu)(\bar{X}_{..} - \mu) + \sum_{i=1}^I J_i (\bar{X}_{..} - \mu)^2 \quad (9.30)$$

Ahora

$$\bar{X}_{..} = \sum_{i=1}^I \frac{J_i \bar{X}_{i.}}{N}$$

por lo que

$$\bar{X}_{..} - \mu = \sum_{i=1}^I \frac{J_i (\bar{X}_{i.} - \mu)}{N}$$

y

$$\sum_{i=1}^I J_i (\bar{X}_{i.} - \mu) = N(\bar{X}_{..} - \mu)$$

Sustituyendo en el término de en medio del lado derecho de (9.30), se obtiene

$$SSTr = \sum_{i=1}^I J_i (\bar{X}_{i.} - \mu)^2 - 2N(\bar{X}_{..} - \mu)^2 + \sum_{i=1}^I J_i (\bar{X}_{..} - \mu)^2$$

Puesto que $\sum_{i=1}^I J_i = N$, se obtiene

$$SSTr = \sum_{i=1}^I J_i (\bar{X}_{i.} - \mu)^2 - N(\bar{X}_{..} - \mu)^2$$

Tomando medias en ambos lados se obtiene

$$E(SSTr) = \sum_{i=1}^I J_i E(\bar{X}_{i.} - \mu)^2 - NE(\bar{X}_{..} - \mu)^2 \quad (9.31)$$

Ahora $E(\bar{X}_{i.}) = E(\bar{X}_{..}) = \mu$, por lo que

$$E(\bar{X}_{i.} - \mu)^2 = V(\bar{X}_{i.}) = \frac{\sigma^2}{J_i}$$

$$E(\bar{X}_{..} - \mu)^2 = V(\bar{X}_{..}) = \frac{\sigma^2}{N}$$

Sustituyendo en (9.31) se obtiene

$$E(SSTr) = \sum_{i=1}^I \frac{J_i \sigma^2}{J_i} - \frac{N \sigma^2}{N} = (I - 1) \sigma^2$$

Ejercicios para la sección 9.1

- Uno de los factores que determina el grado del riesgo que un pesticida plantea a la salud humana es la tasa con la que lo absorbe la piel después del contacto. Una pregunta importante es si la cantidad en la piel aumenta con la duración del contacto, o si sólo aumenta durante un tiempo corto antes de estabilizarse. Para investigar esto último se aplicaron cantidades medidas de cierto pesticida a 20 muestras de piel de rata. Se analizaron cuatro pieles en cada uno de los intervalos de 1, 2, 4, 10 y 24 horas. Las cantidades de compuestos químicos (en μg) que estaban en la piel están dadas en la tabla siguiente.

Duración	Cantidades absorbidas			
1	1.7	1.5	1.2	1.5
2	1.8	1.6	1.8	1.9
4	1.9	1.7	2.1	2.0
10	2.3	1.9	1.7	1.5
24	2.1	2.2	2.5	2.3

- Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- ¿Puede concluir que la cantidad en la piel varía con el tiempo?

2. La fuerza producida de soldaduras de titanio CP se midió para soldaduras enfriadas en tasas de 10°C/s, 15°C/s y 28°C/s. Los resultados se presentan en la tabla siguiente. (Basada en el artículo “Advances in Oxygen Equivalence Equations for Predicting the Properties of Titanium Welds” D. Harwig, W. Ittiwattana, y H. Castner, en *The Welding Journal*, 2001:126s-136s.)

Razón de enfriamiento	Producción de la fuerza									
10	71.00	75.00	79.67	81.00	75.50	72.50	73.50	78.50	78.50	
15	63.00	68.00	73.00	76.00	79.67	81.00				
28	68.65	73.70	78.40	84.40	91.20	87.15	77.20	80.70	84.85	88.40

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que la producción de la fuerza de soldaduras de titanio CP varía con la tasa de enfriamiento?
3. La eliminación de nitrógeno de amoníaco es un aspecto importante del tratamiento de filtraciones en basureros. Durante varios días se registró la tasa de eliminación (en % por día) para cada uno de los diferentes métodos de tratamiento. Los resultados se presentan en la tabla siguiente. (Basado en el artículo “Removal of Ammoniacal Nitrogen from Landfill Leachate by Irrigation onto Vegetated Treatment Planes”, S. Tyrrel, P. Leeds-Harrison y K. Harrison, en *Water Research*, 2002:291-299.)

Tratamiento	Tasa de eliminación			
A	5.21	4.65		
B	5.59	2.69	7.57	5.16
C	6.24	5.94	6.41	
D	6.85	9.18	4.94	
E	4.04	3.29	4.52	3.75

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que los métodos de tratamiento difieren en sus tasas de eliminación?
4. En el artículo “Calibration of an FTIR Spectrometer” (P. Pankratz, *Statistical Case Studies for Industrial and Process Improvement*, SIAM-ASA, 1997:19-38) se usa un espectrómetro para hacer cinco mediciones del contenido de carbono (en ppm) de cierta placa de silicio en cuatro días consecutivos. Los resultados son:

Día 1:	358	390	380	372	366
Día 2:	373	376	413	367	368
Día 3:	336	360	370	368	352
Día 4:	368	359	351	349	343

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que la calibración del espectrómetro es diferente entre los cuatro días?
5. El artículo “Quality of the Fire Clay Coal Bed, Southeastern Kentucky” (J. Hower, W. Andrews y cols., en *Journal of Coal Quality*, 1994:13-26) contiene mediciones de muestras de carbón de varios lugares en Kentucky. Los datos en porcentaje de ceniza de TiO_2 son los siguientes (se ha eliminado un dato atípico):

Buckeye Hollow:	0.96	0.86	0.94	0.91	0.70	1.28	1.19	1.04	1.42	0.82	0.89	1.45	1.66	1.68
	2.10	2.19												
Bear Branch:	0.91	1.42	2.54	2.23	2.20	1.44	1.70	1.53	1.84					
Defeated Creek:	1.30	1.39	2.58	1.49	1.49	2.07	1.87	1.39	1.02	0.91	0.82	0.67	1.34	1.51
Turkey Creek:	1.20	1.60	1.32	1.24	1.08	2.33	1.81	1.76	1.25	0.81	0.95	1.92		

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que hay diferencias en contenido de TiO_2 entre estos lugares?

6. Los arqueólogos pueden determinar las dietas de las antiguas civilizaciones midiendo el cociente de carbono 13 entre carbono 12 en huesos encontrados en sitios de entierro. Las grandes cantidades de carbono 13 sugieren una dieta rica en hierbas como maíz, mientras que las cantidades pequeñas sugieren una dieta basada de plantas herbáceas. El artículo “Climate and Diet in Fremont Prehistory: Economic Variability and Abandonment of Maize Agriculture in the Great Salt Lake Basin” (J. Coltrain y S. Leavitt, en *American Antiquity*, 2002:453-485) informa cocientes, como una diferencia de un patrón en unidades de partes por mil, para huesos de personas en varios grupos de edades. Los datos se presentan en la tabla siguiente.

Grupo de edades (años)	Cociente											
0-11	17.2	18.4	17.9	16.6	19.0	18.3	13.6	13.5	18.5	19.1	19.1	13.4
12-24	14.8	17.6	18.3	17.2	10.0	11.3	10.2	17.0	18.9	19.2		
25-45	18.4	13.0	14.8	18.4	12.8	17.6	18.8	17.9	18.5	17.5	18.3	15.2
	19.2	15.4	13.2								10.8	19.8
46+	15.5	18.2	12.7	15.1	18.2	18.0	14.4	10.2	16.7			

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que los coeficientes de concentración son diferentes entre los grupos de edades?
7. Un experimento se realizó para medir la producción proporcionada por cada uno de tres catalizadores en cierta reacción. El experimento se repitió tres veces para cada catalizador. Las producciones del reactor, en gramos, son:

Catalizador 1: 84.33 90.25 85.62
Catalizador 2: 88.44 89.81 86.53
Catalizador 3: 94.71 91.19 92.81

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que hay diferencias en las medias de las producciones entre los catalizadores?
8. Se llevó a cabo un experimento para comparar las vidas de cuatro marcas diferentes de bujía de encendido. Se usaron cinco bujías de cada marca y se registró el número de millas hasta que fallaba cada una. Lo siguiente es una parte del resultado MINITAB para un ANOVA de un sentido.

One-way Analysis of Variance

Analysis of Variance

Source	DF	SS	MS	F	P
Brand	3	176.482	(a)	(e)	(f)
Error	(b)	(c)	(d)		
Total	19	235.958			

Complete los números faltantes del a) al f) en la tabla. Puede dar un rango para el P -valor.

9. En relación con el ejercicio 8, ¿es creíble que las marcas de bujía de encendido tengan la misma media de tiempos de vida?
10. Se compararon tres métodos de separación en cierto proceso químico para estudiar sus efectos sobre la producción. Se hicieron tres operaciones con cada método, y las producciones, en porcentaje de un máximo teórico, son los siguientes:

Método A: 84.6 83.3 85.1
Método B: 87.3 85.9 88.2
Método C: 87.2 86.0 86.3

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que hay diferencias entre las medias de las producciones?

11. Se realizó un experimento para determinar si la temperatura de recocido de hierro dúctil afecta su fuerza de tensión. Se recocieron cinco elementos para cada una de las cuatro temperaturas. Se midió la fuerza de tensión (en ksi) para cada una. Los resultados se presentan en la tabla siguiente.

Temperatura (°C)	Valores muestrales				
750	19.72	20.88	19.63	18.68	17.89
800	16.01	20.04	18.10	20.28	20.53
850	16.66	17.38	14.49	18.21	15.58
900	16.93	14.49	16.15	15.53	13.25

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que hay diferencias entre las medias de las fuerzas de tensión?
12. Considere el ejercicio 10.
- a) Calcule la cantidad $s = \sqrt{\text{MSE}}$, la estimación de la desviación estándar del error σ .
- b) Suponiendo que s sea la desviación estándar del error, encuentre el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.90 y detectar una diferencia máxima de 2 en las medias de tratamiento a un nivel del 5%.
- c) Utilizando una estimación más conservadora de $1.5s$ como la desviación estándar de error, encuentre el tamaño muestral necesario en cada tratamiento para proporcionar una potencia de 0.90 para detectar una diferencia máxima de 2 en las medias de tratamiento con un nivel de 5%.
13. Considérese el ejercicio 11.
- a) Calcule la cantidad $s = \sqrt{\text{MSE}}$, la estimación de la desviación estándar de error σ .
- b) Suponiendo que s sea la desviación estándar de error, determine el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.90 y detectar una diferencia máxima de 2 en las medias del tratamiento con un nivel de 5%.
- c) Utilizando una estimación más conservadora de $1.5s$ como la desviación estándar del error, determine el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.90 y detectar una diferencia máxima de 2 en las medias de tratamiento con un nivel de 5%.
14. El artículo “The Lubrication of Metal-on-Metal Total Hip Joints: A Slide Down the Stribeck Curve k” (S. Smith, D. Dowson, y A. Goldsmith, en *Proceedings of the Institution of Mechanical Engineers*, 2001:483-493) presenta los resultados de pruebas de desgaste hechas en uniones artificiales de cadera de metal. Se probaron uniones con diferentes diámetros. Los datos que se presentan en la tabla siguiente para la rugosidad de la cabeza son consistentes con las medias y las desviaciones estándar descritas en el artículo.

Diámetro (mm)	Rugosidad de la cabeza				
	(nm)				
16	0.83	2.25	0.20	2.78	3.93
28	2.72	2.48	3.80		
36	5.99	5.32	4.59		

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que la media de la rugosidad varía con el diámetro? Explique.
15. El artículo “Mechanical Grading of Oak Timbers” (D. Kretschmann y D. Green, en *Journal of Materials in Civil Engineering*, 1999:91-97) presenta mediciones de los módulos de ruptura, en MPa, para maderas de 7 por 9 de roble mezclados verdes de Virginia occidental y Pensilvania. Las medias, desviaciones estándar y tamaños muestrales para cuatro diferentes grados de madera se presentan en la tabla siguiente.

Grado	Media	Desviación estándar	Tamaño muestral
Selecta	45.1	8.52	32
Núm. 1	42.0	5.50	11
Núm. 2	33.2	6.71	15
Grado bajo	38.1	8.04	42

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que la media de los módulos de ruptura difieren para los diferentes grados de madera?
16. El artículo “Withdrawal Strength of Threaded Nails” (D. Rammer, S. Winistorfer y D. Sender, en *Journal of Structural Engineering* 2001:442-449) describe un experimento para comparar las fuerzas de retirada para algunos tipos de clavos. Los datos contenidos en la tabla siguiente son consistentes con las medias y las desviaciones estándar que se describen en el artículo para tres tipos de clavos: con estrías anulares, con estrías en espiral y comunes. Todos los clavos tenían diámetros a 0.1 mm del otro, y todos fueron hechos con el mismo tipo de madera.

Tipo de clavo	Fuerza de retirada (N/mm)									
Con estrías anulares	36.57	29.67	43.38	26.94	12.03	21.66	41.79	31.50	35.84	40.81
Con estrías en espiral	14.66	24.22	23.83	21.80	27.22	38.25	28.15	36.35	23.89	28.44
Común	12.61	25.71	17.69	24.69	26.48	19.35	28.60	42.17	25.11	19.98

- a) Construya una tabla ANOVA. Puede dar un rango para el P -valor.
- b) ¿Puede concluir que la media de la fuerza de retirada es diferente para los diferentes tipos de clavo?
17. El siguiente resultado de MINITAB presenta un cálculo de potencia.

Alpha = 0.05	Assumed standard deviation = 142.6			Number of Levels = 4
SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
20000	14	0.85	0.864138	200
The sample size is for each level.				

- a) ¿Cuál es la potencia que requiere el experimentador?
- b) Para garantizar una potencia de 0.864138, ¿cuántas observaciones se deben tomar para todos los tratamientos combinados?
- c) ¿Cuál es la diferencia entre las medias del tratamiento que puedan detectarse con una potencia de al menos 0.864138?
- d) ¿La potencia para detectar es una diferencia máxima de 250 mayor que 0.864138 o menor que 0.864138? Explique.

9.2 Comparaciones apareadas en los experimentos de un solo factor

En un ANOVA de un sentido se utiliza una prueba F para probar la hipótesis nula de que todas las medias de tratamiento son iguales. Si ésta es rechazada se puede concluir que las medias de tratamiento no son todas iguales. Pero esta prueba no indica cuáles son diferentes del resto. A veces un experimentador considera dos tratamientos específicos, i y j , y quiere estudiar la diferencia $\mu_i - \mu_j$. En este caso el método de la diferencia significativa mínima de Fisher (DSM)

es adecuado y puede usarse para construir intervalos de confianza para $\mu_i - \mu_j$ o para probar la hipótesis nula que $\mu_i - \mu_j = 0$. Otras veces, un experimentador tal vez desee determinar todos los pares de medias que se pueda concluir que difieren de otro. En este caso se debe utilizar otro tipo de procedimiento que se llama **método de comparaciones múltiples**. Se analizarán dos métodos de comparaciones múltiples, el de Bonferroni y el de Tukey-Kramer.

El método de la diferencia significativa mínima de Fisher

Se inicia mediante la descripción del método DSM de Fisher para construir intervalos de confianza de Fisher. El intervalo de confianza para la diferencia $\mu_i - \mu_j$ se centra en la diferencia de las medias muestrales $\bar{X}_i - \bar{X}_j$. Determinar qué tan ancho hacer el intervalo de confianza requiere calcular la desviación estándar de $\bar{X}_i - \bar{X}_j$. Sean J_i y J_j los tamaños de muestra en los niveles i y j , respectivamente. Debido al supuesto de que todas las observaciones están distribuidas normalmente con varianza σ^2 , por lo que $\bar{X}_i - \bar{X}_j$ está distribuido normalmente con media $\mu_i - \mu_j$ y varianza $\sigma^2(1/J_i + 1/J_j)$. La varianza σ^2 se estima con MSE, para las razones que se explicaron previamente en el análisis de los intervalos de confianza para las medias de tratamiento (sección 9.1). Ahora la cantidad

$$\frac{(\bar{X}_i - \bar{X}_j) - (\mu_i - \mu_j)}{\sqrt{\text{MSE}(1/J_i + 1/J_j)}}$$

tiene una distribución t de Student con $N - I$ grados de libertad. (El valor $N - I$ es el número de grados de libertad usado para calcular MSE; véase la ecuación 9.13.) La cantidad $t_{N-I, \alpha/2} \sqrt{\text{MSE}(1/J_i + 1/J_j)}$ se llama diferencia significativa mínima. Esta última constituye la base para los intervalos de confianza y las pruebas de hipótesis.

Método de Fisher de la diferencia significativa mínima para intervalos de confianza y pruebas de hipótesis

El intervalo de confianza de la diferencia significativa mínima de Fisher, al nivel $100(1 - \alpha)\%$, para la diferencia $\mu_i - \mu_j$ es

$$\bar{X}_i - \bar{X}_j \pm t_{N-I, \alpha/2} \sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)} \quad (9.32)$$

Para probar la hipótesis nula $H_0 : \mu_i - \mu_j = 0$, el estadístico de prueba es

$$\frac{\bar{X}_i - \bar{X}_j}{\sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)}} \quad (9.33)$$

Si H_0 es verdadera, este estadístico tiene una distribución t Student con $N - I$ grados de libertad. Específicamente, si

$$|\bar{X}_i - \bar{X}_j| > t_{N-I, \alpha/2} \sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)} \quad (9.34)$$

Entonces H_0 es rechazada al nivel α .

La razón de que la cantidad $t_{N-I, \alpha/2} \sqrt{\text{MSE}(1/J_i + 1/J_j)}$ se llame la diferencia significativa mínima es que la hipótesis nula de las medias iguales se rechazó con un nivel de α siempre que la diferencia en las medias muestrales $|\bar{X}_i - \bar{X}_j|$ supere este valor. Cuando el diseño es balanceado, con todos los tamaños muestral iguales a J , la diferencia significativa mínima es igual a $t_{N-I, \alpha/2} \sqrt{2\text{MSE}/J}$ para todos los pares de medias.

Ejemplo

9.9

En el experimento de las soldaduras analizado en la sección 9.1 se midió la dureza para cinco soldaduras de cada uno de cuatro flujos A, B, C, y D. La media muestral de los valores de la dureza fueron $\bar{X}_A = 253.8$, $\bar{X}_B = 263.2$, $\bar{X}_C = 271.0$ y $\bar{X}_D = 262.0$. La tabla ANOVA se muestra en el siguiente cuadro.

One-way ANOVA: A, B, C, D

Source	DF	SS	MS	F	P
Factor	3	743.40	247.800	3.87	0.029
Error	16	1023.60	63.975		
Total	19	1767.00			

S = 7.998 R-Sq = 42.07% R-Sq(adj) = 31.21%

Antes de realizar el experimento se midieron los contenidos de carbono de los flujos. El flujo B tenía el contenido de carbono más bajo (2.67% por peso), y el flujo C tenía el más alto (5.05% por peso). Por tanto, el experimentador está particularmente interesado en comparar las durezas obtenidas con estos dos flujos. Determine un intervalo de confianza del 95% para la diferencia de las medias de durezas entre soldaduras producidas con los flujos B y C. ¿Se puede concluir que las dos medias son diferentes?

Solución

Se utiliza la expresión (9.32). Las medias muestrales son 271.0 y 263.2 para los flujos C y B, respectivamente. El producto anterior de MINITAB da la cantidad MSE de 63.975. (Este valor también se calculó en el ejemplo 9.3 en la sección 9.1.) Ambos tamaños muestrales son iguales a 5. Hay $I = 4$ niveles y $N = 20$ observaciones en total. Para un intervalo de confianza de 95% se consulta la tabla t para determinar el valor $t_{16, .025} = 2.120$. Por tanto, el intervalo de confianza de 95% es $271.0 - 263.2 \pm 2.120 \sqrt{63.975(1/5 + 1/5)}$ o $(-2.92, 18.52)$.

Para realizar una prueba de hipótesis nula con el fin de que las dos medias de tratamiento sean iguales, se calcula el valor del estadístico de prueba (expresión 9.33) y se obtiene

$$\frac{271.0 - 263.2}{\sqrt{63.975(1/5 + 1/5)}} = 1.54$$

Al consultar la tabla t con $N - I = 16$ grados de libertad, se encuentra que P está entre $2(0.05) = 0.10$ y $2(0.10) = 0.20$ (observe que ésta es una prueba de dos colas). No se puede concluir que las medias del tratamiento difieren.

Si se desea realizar una prueba de nivel fijo a un nivel $\alpha = 0.05$ como una alternativa para calcular el P -valor, el valor crítico t es $t_{16, .025} = 2.120$. El lado izquierdo de la desigualdad (9.34) es $|271.0 - 263.2| = 7.8$. El lado derecho es $2.120\sqrt{63.975(1/5+1/5)} = 10.72$. Dado que 7.8 no es mayor que 10.72, no se rechaza H_0 al nivel 5%.

El siguiente resultado de MINITAB presenta los intervalos de confianza al 95% de Fisher de DSM para cada diferencia entre las medias de tratamiento en el experimento de soldaduras.

Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons

Simultaneous confidence level = 81.11%

A subtracted from:

	Lower	Center	Upper	
B	-1.324	9.400	20.124	(-----*-----)
C	6.476	17.200	27.924	(-----*-----)
D	-2.524	8.200	18.924	(-----*-----)

B subtracted from:

	Lower	Center	Upper	
C	-2.924	7.800	18.524	(-----*-----)
D	-11.924	-1.200	9.524	(-----*-----)

C subtracted from:

	Lower	Center	Upper	
D	-19.724	-9.000	1.724	(-----*-----)

-12 0 12 24

Los valores “Center” son las diferencias entre pares de medias de tratamiento. Las cantidades “Lower” y “Upper” son los límites superior e inferior, respectivamente, del intervalo de confianza. De particular interés es el nivel de confianza simultáneo de 81.11%. Esto último indica que aunque se tiene una confianza de 95% de que cualquier intervalo de confianza dado contiene su diferencia verdadera en las medias, sólo se tiene una confianza de 81.11% de que *todos* los intervalos de confianza contienen sus diferencias verdaderas.

En el ejemplo 9.9 se realizó una sola prueba sobre la diferencia entre dos medias específicas. ¿Qué sucede si se quería probar cada par de medias para averiguar acerca de cuáles se puede concluir que son diferentes? Podría parecer razonable realizar la prueba DSM sobre cada par. Sin embargo, esto último no es adecuado porque cuando se realizan varias pruebas aumenta la probabilidad de rechazar una hipótesis nula verdadera. Éste es problema de pruebas múltiples, que se analizaron con cierto detalle en la sección 6.14. Este problema se mostró en el resultado anterior de MINITAB, que muestra que sólo se tiene una confianza de

81.11% de que todos los intervalos de confianza de 95% contengan sus valores verdaderos. Cuando se consideran simultáneamente, algunos intervalos de confianza o pruebas de hipótesis, aquéllos deben ser más amplios, y el criterio para rechazar las hipótesis nulas más estricto, que en situaciones donde sólo está implicado un intervalo o una prueba. En estas situaciones se utilizan métodos de comparaciones múltiples para producir **intervalos de confianza simultáneos** y **pruebas de hipótesis simultáneas**. Si se construyen intervalos de confianza simultáneos de nivel $100(1 - \alpha)\%$ para las diferencias entre cada par de medias, entonces se tiene una confianza de nivel $100(1 - \alpha)\%$ de que *cada* intervalo de confianza contenga la diferencia verdadera. Si se realizan pruebas de hipótesis simultáneas para todas las hipótesis nulas de la forma $H_0 : \mu_i - \mu_j = 0$, entonces se puede rechazar, al nivel α , cada hipótesis nula cuyo P -valor sea menor que α .

El método de Bonferroni de comparaciones múltiples

El método de Bonferroni, analizado en la sección 6.14, es un método general, válido siempre que algunos intervalos de confianza o pruebas sean considerados simultáneamente. La aplicación del método es simple. Sea C el número de pares de las diferencias que van a compararse. Por ejemplo, si hay I tratamientos, y todos los pares de diferencias que van a compararse, entonces $C = I(I - 1)/2$. El método de Bonferroni es el mismo que el método de DSM, excepto que α se sustituye con α/C .

El método de Bonferroni para intervalos de confianza simultáneos y pruebas de hipótesis

Se supone que se están considerando C diferencias de la forma $\mu_i - \mu_j$. Los intervalos de confianza simultáneos de Bonferroni, al nivel $100(1 - \alpha)\%$, para las C diferencias $\mu_i - \mu_j$ son

$$\bar{X}_{i.} - \bar{X}_{j.} \pm t_{N-I, \alpha/(2C)} \sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)} \quad (9.35)$$

Se tiene una confianza del $100(1 - \alpha)\%$ de que los intervalos de confianza de Bonferroni contengan el valor verdadero de la diferencia $\mu_i - \mu_j$ para todos los pares C que se están considerando.

Con el fin de probar las C hipótesis nulas de la forma $H_0 : \mu_i - \mu_j = 0$, los estadísticos de prueba son

$$\frac{\bar{X}_{i.} - \bar{X}_{j.}}{\sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)}}$$

Determine el P -valor para cada prueba consultando la tabla t de Student con $N - 1$ grados de libertad, y multiplique el P -valor encontrado por C .

Específicamente, si

$$|\bar{X}_{i.} - \bar{X}_{j.}| > t_{N-I, \alpha/(2C)} \sqrt{\text{MSE} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)}$$

entonces H_0 es rechazada con un nivel de α .

Ejemplo

9.10

Con los datos de soldaduras analizados en el ejemplo 9.9 utilice el método de Bonferroni para determinar cuáles pares de flujos, si los hay, se puede concluir que difieren en su efecto de dureza, a un nivel del 5%.

Solución

Hay $I = 4$ niveles, con $J = 5$ observaciones en cada nivel, de un total de $N = 20$ observaciones. Con cuatro niveles, hay un total de $C = (4)(3)/2 = 6$ pares de medias que comparar.

Para probar al nivel $\alpha = 5\%$, se calcula $\alpha/(2C) = 0.004167$. El valor crítico de t es $t_{16, .004167}$. Este valor no está en la tabla; se encuentra entre $t_{16, .005} = 2.921$ y $t_{16, .001} = 3.686$. Mediante software se calcula $t_{16, .004167} = 3.0083$. Sin el respectivo software se podría aproximar este valor en forma aproximada por interpolación. Ahora $MSE = 63.975$ (véase el ejemplo 9.9), por lo que $t_{N-I, \alpha/(2C)} \sqrt{MSE(1/J_i + 1/J_j)} = 3.0083 \sqrt{63.975(1/5 + 1/5)} = 15.22$. Las cuatro medias muestrales son:

Flujo	A	B	C	D
Medida de la dureza	253.8	263.2	271.0	262.0

Hay solamente un par de las medias muestrales, 271.0 y 253.8, cuya diferencia es mayor a 15.22. Por tanto, se concluye que las soldaduras producidas con el flujo A tienen una media de dureza diferente que las generadas con el flujo C. Ninguna de las otras diferencias son importantes con un nivel de 5%.

Aunque es fácil de utilizar, el método de Bonferroni tiene la desventaja que cuando C se hace grande, los intervalos de confianza resultan muy anchos, y las pruebas de hipótesis tienen potencia baja. La razón de esto último consiste en que el método de Bonferroni es un método general, no específicamente diseñado para el análisis de varianza o para poblaciones normales. En muchos casos C es bastante grande, en particular con frecuencia se desean comparar todos los pares de medias. En estos casos, el **método de Tukey-Kramer** es superior, porque está diseñado para comparaciones múltiples de las medias de poblaciones normales. A continuación se le describe.

El método de Tukey-Kramer de comparaciones múltiples

El método de Tukey-Kramer está basado en la **distribución de rango studentizado**, en lugar de la distribución t de Student. Dicha distribución tiene dos grados de libertad como valores, que para el método de Tukey-Kramer son I y $N - I$. (En comparación, la prueba F utiliza $I - 1$ y $N - I$ grados de libertad.) El método de Tukey-Kramer utiliza el cuantil $1 - \alpha$ de la distribución de rango studentizado con I y $N - I$ grados de libertad; esta cantidad se denota por $q_{I, N-I, \alpha}$. La tabla A.8 (en el Apéndice A) presenta valores de $q_{I, N-I, \alpha}$ para diferentes valores de I , N , y α . Los mecanismos del método de Tukey-Kramer son los mismos que los del método de DSM, excepto que $t_{N-I, \alpha/2} \sqrt{MSE(1/J_i + 1/J_j)}$ se sustituye con $q_{I, N-I, \alpha} \sqrt{(MSE/2)(1/J_i + 1/J_j)}$. Algunas veces la cantidad $q_{I, N-I, \alpha} \sqrt{(MSE/2)(1/J_i + 1/J_j)}$ se denomina **diferencia honestamente significativa** (DHS), en contraparte con la diferencia significativa mínima de Fisher.

El método de Tukey-Kramer para los intervalos de confianza simultáneos y pruebas de hipótesis

Los intervalos de confianza simultáneos de Tukey-Kramer de nivel $100(1 - \alpha)\%$ para todas las diferencias $\mu_i - \mu_j$ son

$$\bar{X}_i. - \bar{X}_j. \pm q_{I, N-I, \alpha} \sqrt{\frac{\text{MSE}}{2} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)} \quad (9.36)$$

Ahora se tiene una confianza del $100(1 - \alpha)\%$ para los intervalos de confianza de Tukey-Kramer de que contengan el valor verdadero de la diferencia $\mu_i - \mu_j$ para cada i y j .

Con el fin de probar todas las hipótesis nulas simultáneas $H_0 : \mu_i - \mu_j = 0$, los estadísticos de prueba son

$$\frac{\bar{X}_i. - \bar{X}_j.}{\sqrt{\frac{\text{MSE}}{2} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)}}$$

El P -valor para cada prueba se encuentra al consultar la tabla de la distribución de rango studentizado (tabla A.8) con I y $N - I$ grados de libertad.

Para cada par de niveles i y j para los cuales

$$|\bar{X}_i. - \bar{X}_j.| > q_{I, N-I, \alpha} \sqrt{\frac{\text{MSE}}{2} \left(\frac{1}{J_i} + \frac{1}{J_j} \right)}$$

la hipótesis nula $H_0 : \mu_i - \mu_j = 0$ se rechaza con un nivel de α .

Una nota de terminología: cuando el diseño es balanceado, con todos los tamaños de muestra iguales a J , la cantidad $\sqrt{(\text{MSE}/2)(1/J_i + 1/J_j)}$ es igual a $\sqrt{\text{MSE}/J}$ para todos los pares de niveles. En este caso, con frecuencia el método sólo se llama método de Tukey.

Ejemplo

9.11

Para los datos de las soldaduras en la tabla 9.1 (en la sección 9.1), ¿cuáles pares de flujos, si hay, se puede concluir, con un nivel de 5%, que difieren en su efecto en la dureza?

Solución

Hay $I = 4$ niveles, con $J = 5$ observaciones en cada nivel, para un total de $N = 20$ observaciones en total. Con el propósito de probar con un nivel de $\alpha = 0.05$, se consulta la tabla de rango studentizado (tabla A.8) encontrando $q_{4, 16, .05} = 4.05$.

El valor de MSE es 63.975 (véase el ejemplo 9.9). Por tanto, $q_{I, N-I, \alpha} \sqrt{\text{MSE}/J} = 4.05 \sqrt{63.975/5} = 14.49$. Las cuatro medias muestrales son:

Flujo	A	B	C	D
Medida de la dureza	253.8	263.2	271.0	262.0

Hay solamente un par de medias muestrales, 271.0 y 253.8, cuya diferencia es mayor a 14.49. Por tanto, se concluye que las soldaduras producidas con el flujo A tienen una media de dureza diferente que las soldaduras producidas con el flujo C. Ninguna de las otras diferencias son importantes con un nivel de 5%.

Al comparar los resultados del ejemplo 9.11 con los del 9.10 se encuentra, en este caso, que el método de Tukey-Kramer es ligeramente más poderoso que el método de Bonferroni, ya que su valor crítico es solamente de 14.49, mientras que el de Bonferroni fue de 15.22. Cuando es posible comparar todos los pares posibles, como en este ejemplo, el método de Tukey-Kramer es siempre más poderoso que el de Bonferroni. Cuando sólo se comparan algunos de los pares posibles, el método de Bonferroni es a veces más poderoso.

En ocasiones sólo se realiza una sola prueba, pero la diferencia que se prueba se elige al examinar las medias muestrales y al elegir dos cuya diferencia es grande. En estos casos se debe utilizar un método de comparaciones múltiples, aunque sólo se esté realizando una prueba. El ejemplo 9.12 ilustra la idea.

Ejemplo

9.12

Un ingeniero examina los datos de las soldaduras en la tabla 9.1 y advierte que los dos tratamientos con la diferencia mayor en las medias muestrales son los flujos A y C. Decide probar la hipótesis nula de que la media de la dureza para las soldaduras producidas con el flujo A difiere de las que se producen con el flujo C. Debido a que sólo se realizará una prueba, utiliza el método de Fisher de DSM en lugar del de Bonferroni o de Tukey-Kramer. Explique por qué está equivocado.

Solución

El ingeniero ha revisado cada par de medias y ha elegido dos cuya diferencia es mayor. Aunque formalmente está realizando sólo una prueba, ha elegido ésta al comparar cada par de medias muestrales. Por esta razón debe utilizar un procedimiento de comparaciones múltiples, como el método de Bonferroni o de Tukey-Kramer.

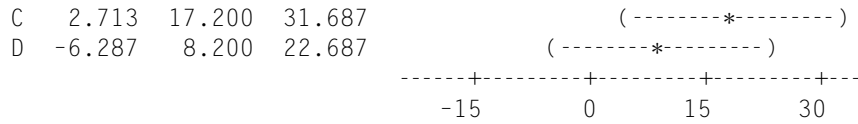
El siguiente resultado de MINITAB presenta intervalos de confianza simultáneos de 95% de Tukey-Kramer para los datos de las soldaduras.

Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons

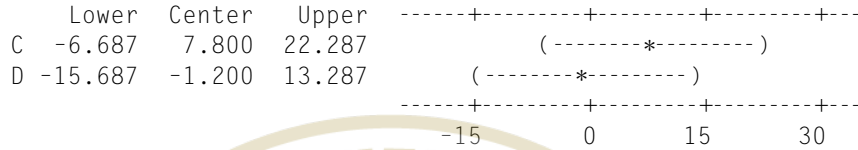
Individual confidence level = 98.87%

A subtracted from:

	Lower	Center	Upper	-----+-----+-----+-----+-----
B	-5.087	9.4000	23.887	(-----*-----)



B subtracted from:



C subtracted from:



Los valores “Center” son las diferencias entre pares de medias del tratamiento. Las cantidades “Lower” y “Upper” son los límites inferior y superior, respectivamente, del intervalo de confianza. Se tiene una confianza de 95% de que cada uno de estos intervalos contenga la diferencia verdadera de las medias de tratamiento. Observe que el “Individual confidence level” es 98.87%. Esto significa que se tiene una confianza con ese porcentaje de que cualquier intervalo de confianza específico contenga su valor verdadero. Por último, se indica que el nivel de confianza para los intervalos de Tukey-Kramer es mayor que para los intervalos de Fisher DSM, los intervalos de Tukey-Kramer son más amplios.

Ejemplo

9.13

En el ejemplo 9.5 (en la sección 9.1) se presentaron varias mediciones de las máximas concentraciones por hora (en $\mu\text{g}/\text{m}^3$) de SO_2 para cada una de cuatro centrales hidroeléctricas, y se concluyó que la media de las concentraciones en las cuatro plantas no eran todas iguales. El siguiente resultado de MINITAB presenta los intervalos de confianza simultáneos del 95% de Tukey-Kramer para la media de las concentraciones de las cuatro plantas. ¿Cuáles pares de plantas, si hay alguna, puede concluir con una confianza del 95% que tienen medias diferentes?

Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons

Individual confidence level = 98.87%

1 subtracted from:

	Lower	Center	Upper	
2	109.4	385.3	661.1	(-----*-----)
3	21.4	312.3	603.1	(-----*-----)
4	-94.6	170.9	436.4	(-----*-----)
				-----+-----+-----+-----+-----
				-300 0 300 600

2 subtracted from:

	Lower	Center	Upper	
3	-348.9	-73.0	202.9	(-----*-----)
4	-463.4	-214.3	34.7	(-----*-----)
				-----+-----+-----+-----+-----
				-300 0 300 600

3 subtracted from:

	Lower	Center	Upper	
4	-406.8	-141.3	124.1	(-----*-----)
				-----+-----+-----+-----+-----
				-300 0 300 600

Solución

Entre los intervalos de confianza simultáneos hay dos que no contienen 0. Éstos son los intervalos para $\mu_1 - \mu_2$ y para $\mu_1 - \mu_3$. Por tanto, se concluye que la media de las concentraciones difiere entre las plantas 1 y 2 y entre las 1 y 3.

Ejercicios para la sección 9.2

1. El artículo "Organic Recycling for Soil Quality Conservation in a Sub-Tropical Plateau Region" (K. Chakrabarti, B. Sarkar y cols., en *J. Agronomy and Crop Science*, 2000:137-142) describe un experimento en el que se trataron muestras de suelo con seis tratamientos diferentes, con dos réplicas por tratamiento, y se registró la actividad de fosfato ácido (en μmol de *p*-nitrofenol liberado por gramo de tierra seca por hora). Una tabla ANOVA para un ANOVA de un sentido se presenta en el cuadro siguiente.

One-way ANOVA: Treatments A, B, C, D, E, F

Source	DF	SS	MS	F	P
Treatment	5	1.18547	0.23709	46.64	0.000
Error	6	0.03050	0.00508		
Total	11	1.21597			

Las medias del tratamiento fueron

Tratamiento	A	B	C	D	E	F
Media	0.99	1.99	1.405	1.63	1.395	1.22

- ¿Puede concluir que hay diferencias en la actividad de fosfato ácido entre los tratamientos?
 - Utilice el método de Tukey-Kramer para determinar cuáles pares de medias de tratamiento, si los hay, son diferentes a un nivel de 5%.
 - Utilice el método de Bonferroni para determinar cuáles pares de medias de tratamiento, si los hay, son diferentes con un nivel de 5%.
 - ¿Cuál método es más poderoso en este caso: el método de Tukey-Kramer o el de Bonferroni?
 - El experimentador observa que el tratamiento A tenía la media muestral más pequeña, mientras que la media del tratamiento B era mayor. De los métodos de Fisher de DSM, de Bonferroni y del de Tukey-Kramer, ¿cuál, si hay alguno, se puede utilizar para probar que la hipótesis de estas dos medias de tratamiento son iguales?
2. El artículo “Optimum Design of an A-pillar Trim with Rib Structures for Occupant Head Protection” (H. Kim y S. Kang, en *Proceedings of the Institution of Mechanical Engineers*, 2001:1161-1169) analiza un estudio en que varios tipos de columnas A se compararon para determinar cuál proporciona la mayor protección para los ocupantes de los automóviles durante una colisión. Se sigue una tabla ANOVA de un sentido, donde los tratamientos son tres niveles de espacio longitudinal del saliente (el artículo también analiza dos factores insignificantes, que aquí se omitieron). Había nueve réplicas en cada nivel. La respuesta es el criterio de lesión de cabeza (HIC), una cantidad sin unidades que mide la absorción de energía de impacto de la columna.

One-way ANOVA: Spacing

Source	DF	SS	MS	F	P
Spacing	2	50946.6	25473.3	5.071	0.015
Error	24	120550.9	5023.0		
Total	26	171497.4			

Las medias de tratamiento fueron

Tratamiento	A	B	C
Media	930.87	873.14	979.41

- ¿Puede concluir que el espaciado longitudinal afecta la absorción de energía de impacto?
 - Utilice el método de Tukey-Kramer para determinar cuáles pares de medias de tratamiento, si las hay, son diferentes con un nivel de 5%.
 - Utilice el método de Bonferroni para determinar cuáles pares de medias de tratamiento, si las hay, son diferentes con un nivel de 5%.
 - ¿Cuál método es más poderoso en este caso, el método de Tukey-Kramer o el de Bonferroni?
3. Las resinas de acrílico utilizadas en la fabricación de dentaduras postizas no deben absorber mucha agua, ya que se reduce la fuerza. El artículo “Reinforcement of Acrylic Resin for Provisional Fixed Restorations. Part III: Effects of Addition of Titania and Zirconia Mixtures on Some Mechanical and Physical Properties” (W. Panyayong, Y. Oshida y cols., *Bio-Medical Materials and Engineering*, 2002:353-366) describe un estudio del efecto sobre la absorción de agua al agregar dióxido de titanio (TiO_2) y dióxido de circonio (ZrO_2) a una resina acrílica usual. Se midió la absorción de agua (en $\mu\text{g}/\text{mm}^2$) de cada uno de doce elementos de cada una de diferentes formulaciones, con diferentes cantidades de TiO_2 y ZrO_2 , que fueron inmersas en agua durante una semana. Los resultados se presentan en la tabla siguiente.

Formulación	Volumen %		Media	Desviación estándar
	TiO ₂	ZrO ₂		
A (control)	0	0	24.03	2.50
B	1	1	14.88	1.55
C	1	2	12.81	1.08
D	1	0,5	11.21	2.98
E	2	2	16.05	1.66
F	2	4	12.87	0.96
G	2	1	15.23	0.97
H	3	3	15.37	0.64

- Utilice el método de Bonferroni para determinar cuál de las formulaciones no controladas (de la B a la H) difiere, con un nivel de 5%, en su media de absorción de agua a partir de la formulación de control A.
 - Repita el inciso a) utilizando el método de Tukey-Kramer.
 - ¿Qué método es más poderoso para estas comparaciones? ¿Por qué?
4. Considere el ejercicio 1 de la sección 9.1.
- Con el método de Bonferroni determine qué pares de medias, si hay, son diferentes con un nivel de 5%.
 - Con el método de Tukey-Kramer determine qué pares de medias, si hay, son diferentes con un nivel de 5%.
 - ¿Cuál es el método más poderoso para encontrar todos los pares de tratamientos cuyas medias son diferentes, el método de Bonferroni o el de Tukey-Kramer?
5. Considere el ejercicio 11 de la sección 9.1.
- Mediante el método de Bonferroni determine cuáles pares de medias, si hay, son diferentes con un nivel de 5%.
 - Con el método de Tukey-Kramer determine cuáles pares de medias, si hay, son diferentes con un nivel de 5%.
 - ¿Cuál es el método más poderoso para encontrar todos los pares de tratamientos cuyas medias son diferentes, el de Bonferroni o el de Tukey-Kramer?
6. Considérese el ejercicio 1 en la sección 9.1. Un científico quiere determinar si la media de la cantidad absorbida durante 24 horas difiere de la media de las cantidades absorbidas durante una, dos, cuatro y diez horas.
- Con el método de Bonferroni determine cuáles de las medias, si hay, para una, dos, cuatro y diez horas difieren de la media de 24 horas. Utilice un nivel de 5%.
 - Mediante el método de Tukey-Kramer determine cuáles de las medias, si hay, para una, dos, cuatro y diez horas difieren de la media durante 24 horas. Utilice un nivel de 5%.
 - ¿Cuál es el método más poderoso para encontrar todos los tratamientos cuyas medias difieren de la de duración de 24 horas, el de Bonferroni o el de Tukey-Kramer?
7. Considérese el ejercicio 11 de la sección 9.1. Un minero quiere determinar si la media fuerza de tensión para elementos endurecidos a 900°C difiere de la media para fuerzas endurecidas a 750, 800 y 850°C.
- Con el método de Bonferroni determine cuáles de las medias, si hay, para 750, 800 y 850°C difiere de la media para 900°C.
 - Con el método de Tukey-Kramer determine cuáles de las medias, si hay, para 750, 800, y 850°C difiere de la media para 900°C.
 - ¿Cuál es el método más poderoso para encontrar todos los pares de los tratamientos cuyas medias difieren de la de 900°C, el de Bonferroni o el de Tukey-Kramer?

8. Considere el ejercicio 3 de la sección 9.1.
- Con el método de Fisher de DSM encuentre un intervalo de confianza de 95% para la diferencia entre las medias de los tratamientos B y D.
 - Con el método de Tukey-Kramer determine qué pares de tratamientos, si hay, son diferentes con un nivel de 5%.
9. Considérese el ejercicio 5 de la sección 9.1.
- Con el método de Fisher de DSM encuentre un intervalo de confianza de 95% para la diferencia entre las medias para Buckeye Hollow y Bear Branch.
 - Con el método de Tukey-Kramer determine cuáles pares de tratamientos, si hay, difieren con un nivel de 5%.
10. Considere el ejercicio 7 de la sección en 9.1.
- Con el método de Fisher de DSM encuentre un intervalo de confianza de 95% para la diferencia entre las medias para los catalizadores 1 y 3.
 - Con el método de Tukey-Kramer determine qué pares de catalizadores, si hay, difieren con un nivel de 5%.
11. Con referencia al ejercicio 14 en la sección 9.1.
- Con el método de Fisher de DSM encuentre un intervalo de confianza de 95% para la diferencia entre las medias para un diámetro de 16 y otro de 36.
 - Con el método de Tukey-Kramer determine qué pares de diámetros, si hay, difieren con un nivel de 5%.
12. Considere el ejercicio 16 de la sección 9.1.
- Con el método de Fisher de DSM encuentre un intervalo de confianza de 95% para la diferencia entre las medias para clavos de estrías anulares y comunes.
 - Con el método de Tukey-Kramer determine qué pares de tipos de clavo, si hay, difieren con un nivel de 5%.
13. En un experimento para determinar el efecto del catalizador en la producción de cierta reacción, la media de las producciones para las reacciones operadas con cada uno de los cuatro catalizadores fue $\bar{X}_1 = 89.88$, $\bar{X}_2 = 89.51$, $\bar{X}_3 = 86.98$ y $\bar{X}_4 = 85.79$. Suponga que se hicieron cinco operaciones con cada catalizador.
- Si $MSE = 3.85$, calcule el valor del estadístico F para probar la hipótesis nula de que los cuatro catalizadores tienen la misma media de la producción. ¿Puede concluir que difiere con un nivel de 5%?
 - Con el método de Tukey-Kramer determine qué pares de catalizadores, si hay, se puede concluir que difieren con un nivel de 5%.
14. En un experimento para determinar el efecto del tiempo de endurecimiento sobre la fuerza compresiva de cierto tipo de concreto, la media de las fuerzas, en MPa, para elementos endurecidos para cada uno de los cuatro tiempos de endurecimiento fueron $\bar{X}_1 = 1\,316$, $\bar{X}_2 = 1\,326$, $\bar{X}_3 = 1\,375$ y $\bar{X}_4 = 1\,389$. Suponga que se endurecieron cuatro elementos para cada tiempo de endurecimiento.
- Si $MSE = 875.2$, calcule el valor del estadístico F para probar la hipótesis nula de que los cuatro tiempos de endurecimiento tienen la misma media de la fuerza. ¿Esta hipótesis nula puede ser rechazada con un nivel de 5%?
 - Con el método de Tukey-Kramer determine qué pares de tiempos de endurecimiento, si hay, se puede concluir que difieren con un nivel de 5%.
15. Para algunos conjuntos de datos, el estadístico F rechazará la hipótesis nula de ninguna diferencia en la media de la producción, pero el método de Tukey-Kramer no encontrará algún par de medias que pueden concluir que difieren. Para las cuatro medias muestrales dadas en el ejercicio 13, suponiendo un tamaño de muestra de 5 para cada tratamiento, determine un valor de MSE con el propósito de que el estadístico F rechace la hipótesis nula de que no hay diferencia con un nivel de 5%, mientras que el método de Tukey-Kramer no encuentra algún par de medias que difieran con un nivel de 5%.
16. Para algunos conjuntos de datos, el estadístico F rechazará la hipótesis nula de la no diferencia en la media de las producciones, pero el método de Tukey-Kramer no encontrará algún par de medias que se pueda concluir que difiere. Para las cuatro me-

días muestrales dadas en el ejercicio 14, suponiendo un tamaño de muestra de 4 para cada tratamiento, encuentre un valor de MSE con el propósito de que el estadístico F rechace la hipótesis nula de ninguna diferencia con un nivel de 5%, mientras que el método de Tukey-Kramer no encuentra algún par de medias que difieran con un nivel de 5%.

9.3 Experimentos de dos factores

En los experimentos de un factor, analizados en las secciones 9.1 y 9.2, el propósito es determinar si al cambiar el nivel de un solo factor sólo se afecta la respuesta. Muchos experimentos implican variar algunos factores, cada uno de ellos puede afectar la respuesta. En esta sección se analizará el caso en el que existen dos factores. Los experimentos, naturalmente suficientes, son llamados **experimentos de dos factores**. Se muestra con un ejemplo.

Un ingeniero químico está estudiando los efectos de varios reactivos y catalizadores en la producción de cierto proceso. Esta última se expresa como un porcentaje de un máximo teórico. Se hicieron cuatro operaciones del proceso para cada combinación de tres reactivos y cuatro catalizadores. Los resultados se presentan en la tabla 9.2. En este experimento hay dos factores, el catalizador y el reactivo. El primero se llama **factor renglón**, ya que su valor varía de renglón a renglón en la tabla; el segundo se denomina **factor columna**. Estas designaciones son arbitrarias, en la tabla se podía haber presentado tan fácilmente como que los renglones representen los reactivos y las columnas, los catalizadores.

TABLA 9.2 Producciones para varias operaciones de un proceso químico con varias combinaciones de reactivos y catalizadores

Catalizador	Reactivo											
	1				2				3			
A	86.8	82.4	86.7	83.5	93.4	85.2	94.8	83.1	77.9	89.6	89.9	83.7
B	71.9	72.1	80.0	77.4	74.5	87.1	71.9	84.1	87.5	82.7	78.3	90.1
C	65.5	72.4	76.6	66.7	66.7	77.1	76.7	86.1	72.7	77.8	83.5	78.8
D	63.9	70.4	77.2	81.2	73.7	81.6	84.2	84.9	79.8	75.7	80.5	72.9

En general hay I niveles del factor renglón y J niveles del factor columna. (En la tabla 9.2, $I = 4$ y $J = 3$). Por tanto, hay IJ combinaciones diferentes de los dos factores. La terminología para estas combinaciones de factores no está estandarizada. Se llamará a cada combinación de factores **tratamiento**, pero algunos escritores usan el término **combinación de tratamiento**. Se debe recordar que las unidades asignadas a un tratamiento específico se denominan réplicas. Cuando el número de éstas es el mismo para cada tratamiento, se denotará con K . Por consiguiente, en la tabla 9.2, $K = 4$.

Cuando las observaciones se toman en cada tratamiento posible, el diseño se llama **diseño completo** o **diseño factorial completo**. Los diseños incompletos, en que no hay datos para uno o más tratamientos, pueden ser difíciles de interpretar, excepto para algunos casos especiales. Cuando sea posible se deben utilizar los diseños completos. Cuando el número de réplicas es el mismo para cada tratamiento, se dice que el diseño es **balanceado**. Para los experimentos de un solo factor no se necesitaba suponer que el diseño era balanceado. Con los experimentos de dos factores los diseños no balanceados son mucho más difíciles de analizar que los balanceados. Se restringirá el análisis a diseños balanceados. Como con los experi-

mentos de un solo factor, los factores pueden ser fijos o aleatorios. Los dos métodos que se describirán son aplicables a modelos de efectos fijos. Después se describirán brevemente modelos donde uno o ambos factores son aleatorios.

En un diseño totalmente aleatorio, cada tratamiento representa una población, y las observaciones sobre éste representan una muestra aleatoria simple de esa población. Se denotarán los valores muestrales para el tratamiento que corresponderá al i -ésimo nivel del factor de renglón y al j -ésimo nivel del factor columna por $X_{ij1}, \dots, X_{ijK}$. Se denotará el resultado de la media de la población para este tratamiento con μ_{ij} . Los valores μ_{ij} con frecuencia se llaman **medias de tratamiento**. En general, el propósito de un experimento de dos factores consiste en determinar si las medias de tratamiento están afectados por la variación del factor renglón, del factor columna, o de ambos. El método de análisis adecuado para los experimentos de dos factores se denomina **análisis de varianza de dos sentidos**.

Parametrización para el análisis de varianza de dos sentidos

En un análisis de varianza de dos sentidos se desea determinar si al cambiar el nivel de los factores de renglón o de columna cambia también el valor μ_{ij} . Para hacer esto último se debe expresar μ_{ij} en función de los parámetros que describen los factores renglón y columna por separado. Se iniciará esta tarea describiendo alguna notación para los promedios de las medias de tratamiento respecto de los diferentes niveles de los factores renglón y columna.

Para cualquier nivel i del factor de renglón, el promedio de todas las medias de tratamiento μ_{ij} en el i -ésimo renglón se denota mediante $\bar{\mu}_i$. Se expresa $\bar{\mu}_i$ en relación con las medias de tratamiento de la siguiente manera:

$$\bar{\mu}_i = \frac{1}{J} \sum_{j=1}^J \mu_{ij} \quad (9.37)$$

De manera similar, para cualquier nivel j del factor de columna, el promedio de todas las medias de tratamiento μ_{ij} en la j -ésima columna se denota por $\bar{\mu}_j$. Se expresa $\bar{\mu}_j$ en función de las medias de tratamiento, como se muestra a continuación:

$$\bar{\mu}_j = \frac{1}{I} \sum_{i=1}^I \mu_{ij} \quad (9.38)$$

Por último, se define la **gran media de la población**, se denota por μ , que representa el promedio de todas las medias de tratamiento μ_{ij} . La gran media de la población también se puede expresar como el promedio de las $\bar{\mu}_i$ o de las cantidades $\bar{\mu}_j$:

$$\mu = \frac{1}{I} \sum_{i=1}^I \bar{\mu}_i = \frac{1}{J} \sum_{j=1}^J \bar{\mu}_j = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J \mu_{ij} \quad (9.39)$$

La tabla 9.3 ilustra las relaciones entre μ_{ij} , $\bar{\mu}_i$, $\bar{\mu}_j$ y μ .

Utilizando las cantidades $\bar{\mu}_i$, $\bar{\mu}_j$ y μ , se puede descomponer la media de tratamiento μ_{ij} como sigue:

$$\mu_{ij} = \mu + (\bar{\mu}_i - \mu) + (\bar{\mu}_j - \mu) + (\mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \mu) \quad (9.40)$$

La ecuación (9.40) expresa la media de tratamiento μ_{ij} como una suma de cuatro términos. En la práctica, se usa la notación más simple para los tres términos que están más a la derecha en la ecuación (9.40):

$$\alpha_i = \bar{\mu}_{i.} - \mu \quad (9.41)$$

$$\beta_j = \bar{\mu}_{.j} - \mu \quad (9.42)$$

$$\gamma_{ij} = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \mu \quad (9.43)$$

TABLA 9.3 Medias de tratamiento y sus promedios a través de los renglones y hacia debajo de las columnas

Nivel renglón	Nivel columna				Media del renglón
	1	2	...	J	
1	μ_{11}	μ_{12}	...	μ_{1J}	$\bar{\mu}_{1.}$
2	μ_{21}	μ_{22}	...	μ_{2J}	$\bar{\mu}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	μ_{I1}	μ_{I2}	...	μ_{IJ}	$\bar{\mu}_{I.}$
Media de la columna	$\bar{\mu}_{.1}$	$\bar{\mu}_{.2}$	...	$\bar{\mu}_{.J}$	μ

Cada una de las cantidades μ , α_i , β_j y γ_{ij} tiene una interpretación importante:

- La cantidad μ es la gran media de la población, que es el promedio de todas las medias de tratamiento.
- La cantidad $\alpha_i = \bar{\mu}_{i.} - \mu$ se llama el i -ésimo **efecto de renglón**. Representa la diferencia entre el promedio del tratamiento para el nivel del i -ésimo del factor renglón y la gran media de la población. El valor de α_i indica el grado con el cual el i -ésimo nivel factor de renglón tiende a producir resultados que son mayores o menores que la gran media de la población.
- La cantidad $\beta_j = \bar{\mu}_{.j} - \mu$ se llama el j -ésimo **efecto de columna**. Es la diferencia entre el promedio de las medias del tratamiento para el j -ésimo nivel del factor columna y la gran media de la población. El valor de β_j indica el grado con el cual el j -ésimo nivel del factor columna tiende a producir resultados mayores o menores que la gran media de la población.
- La cantidad $\gamma_{ij} = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \mu$ se llama **interacción ij** . El efecto de un nivel de factor renglón (o columna) puede depender de qué nivel del factor columna (o renglón) está apareado con éste. Los términos de interacción miden el grado con el que esto último ocurre. Por ejemplo, suponga que el nivel 1 del factor renglón tiende a producir un resultado grande cuando se aparea con la columna de nivel 1, pero un resultado pequeño cuando se aparea con una columna de nivel 2. En este caso $\gamma_{1,1}$ sería positiva, y $\gamma_{1,2}$ sería negativo.

Tanto los efectos de renglón como los de columna se llaman **efectos principales** para distinguirlos de las interacciones. Observe que hay I efectos renglón, uno por cada nivel del factor renglón, los efectos columna J , uno por cada nivel del factor columna, y las IJ interacciones, una por cada tratamiento. Además, se tiene de las definiciones de las cantidades $\bar{\mu}_{i.}$, $\bar{\mu}_{.j}$ y μ de las ecuaciones (9.37) a la (9.39) que los efectos de renglón, efectos de columna, y las interacciones satisfacen las restricciones siguientes:

$$\sum_{i=1}^I \alpha_i = 0 \quad \sum_{j=1}^J \beta_j = 0 \quad \sum_{i=1}^I \gamma_{ij} = \sum_{j=1}^J \gamma_{ij} = 0 \quad (9.44)$$

Ahora se puede expresar a las medias de tratamiento μ_{ij} en función de α_i , β_j y de γ_{ij} . De la ecuación (9.40) se tiene que

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad (9.45)$$

Para cada observación X_{ijk} se define $\varepsilon_{ijk} = X_{ijk} - \mu_{ij}$, la diferencia entre la media de la observación y su media de tratamiento. Las cantidades ε_{ijk} se llaman **errores**. De aquí que

$$X_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad (9.46)$$

Al combinar las ecuaciones (9.46) y (9.45) se obtiene el modelo ANOVA de dos sentidos:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad (9.47)$$

Cuando todas las interacciones γ_{ij} son iguales a 0, se dice que el **modelo aditivo** es aplicable. Bajo el modelo aditivo, la ecuación (9.45) será

$$\mu_{ij} = \mu + \alpha_i + \beta_j \quad (9.48)$$

y la ecuación (9.47) será

$$X_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} \quad (9.49)$$

Bajo el modelo aditivo, la media de tratamiento μ_{ij} es igual a la gran media de la población μ , más una cantidad α_i que es resultado de utilizar el renglón i más otra β_j que es resultado de utilizar la columna j . En otras palabras, el efecto combinado de utilizar el renglón de nivel i junto con la columna de nivel j se encuentra al sumar los efectos principales individuales de los dos niveles. Cuando alguna o todas las interacciones no son iguales a 0, el modelo aditivo no vale, y el efecto combinado de un nivel de renglón y un nivel de columna no puede determinarse a partir de sus principales efectos individuales.

Ahora se mostrará cómo estimar los parámetros para el modelo completo de dos sentidos (9.47). El procedimiento para el modelo aditivo es exactamente el mismo, excepto que no se estiman las interacciones γ_{ij} . El procedimiento es directo. Primero se define alguna notación para varios promedios de los datos X_{ijk} , utilizando como ejemplo los datos de la tabla 9.2. La tabla 9.4 presenta el promedio de la producción de las cuatro operaciones para cada reactivo y catalizador de la tabla 9.2.

TABLA 9.4 Promedio de las producciones $\bar{X}_{ij}$ para operaciones de un proceso químico utilizando diferentes combinaciones de reactivos y catalizadores

Catalizador	Reactivo			Media del renglón $\bar{X}_{i.}$
	1	2	3	
A	84.85	89.13	85.28	86.42
B	75.35	79.40	84.65	79.80
C	70.30	76.65	78.20	75.05
D	73.18	81.10	77.23	77.17
Media de la columna $\bar{X}_{.j}$	75.92	81.57	81.34	Gran media muestral $\bar{X}_{..} = 79.61$

Cada número en el cuerpo de la tabla 9.4 es el promedio de los cuatro números en la celda correspondiente de la tabla 9.2. Éstos se llaman las **medias de las celdas**. Se denotan por $\bar{X}_{ij}$, y se definen como

$$\bar{X}_{ij} = \frac{1}{K} \sum_{k=1}^K X_{ijk} \quad (9.50)$$

Al promediar las medias de las celdas a través de los renglones se obtienen las **medias de renglón** $\bar{X}_{i..}$:

$$\bar{X}_{i..} = \frac{1}{J} \sum_{j=1}^J \bar{X}_{ij.} = \frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K X_{ijk} \quad (9.51)$$

Asimismo, cuando se promedian las medias de las celdas hacia abajo de las columnas se obtienen las **medias de columna** $\bar{X}_{.j.}$:

$$\bar{X}_{.j.} = \frac{1}{I} \sum_{i=1}^I \bar{X}_{ij.} = \frac{1}{IK} \sum_{i=1}^I \sum_{k=1}^K X_{ijk} \quad (9.52)$$

La gran media muestral $\bar{X}_{...}$ se puede encontrar al calcular el promedio de las medias de renglón, el promedio de las medias de la columna, el promedio de las medias de la celda, o el promedio de todas las observaciones:

$$\bar{X}_{...} = \frac{1}{I} \sum_{i=1}^I \bar{X}_{i..} = \frac{1}{J} \sum_{j=1}^J \bar{X}_{.j.} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J \bar{X}_{ij.} = \frac{1}{IJK} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk} \quad (9.53)$$

Ahora se describe como estimar los parámetros en el modelo ANOVA de dos sentidos. La idea fundamental es que la mejor estimación de la media del tratamiento μ_{ij} es la media de la celda $\bar{X}_{ij.}$, que representa el promedio de las observaciones muestrales que tiene ese tratamiento. Se tiene que la mejor estimación de la cantidad $\bar{\mu}_{i.}$ es la media del renglón $\bar{X}_{i..}$, la mejor estimación de la cantidad $\bar{\mu}_{.j.}$ es la media de la columna $\bar{X}_{.j.}$, y la mejor estimación de la gran media de la población μ es la gran media muestral $\bar{X}_{...}$. Se estiman los efectos de renglón que α_i , los efectos columna β_j , y las interacciones γ_{ij} al sustituir estas estimaciones en las ecuaciones de la (9.41) a la (9.43).

$$\hat{\alpha}_i = \bar{X}_{i..} - \bar{X}_{...} \quad (9.54)$$

$$\hat{\beta}_j = \bar{X}_{.j.} - \bar{X}_{...} \quad (9.55)$$

$$\hat{\gamma}_{ij} = \bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...} \quad (9.56)$$

Los efectos renglón, los efectos columna y las interacciones satisfacen las restricciones dadas en la ecuación (9.44). Realizando un poco de álgebra, se puede demostrar que sus estimaciones satisfacen las mismas restricciones:

$$\sum_{i=1}^I \hat{\alpha}_i = 0 \quad \sum_{j=1}^J \hat{\beta}_j = 0 \quad \sum_{i=1}^I \hat{\gamma}_{ij} = \sum_{j=1}^J \hat{\gamma}_{ij} = 0 \quad (9.57)$$

Ejemplo

9.14

Calcule los efectos estimados de renglón, de columna y las interacciones para los datos de la tabla 9.2.

Solución

Mediante las cantidades de la tabla 9.4 y las ecuaciones (9.54) a (9.56) se calcula

$$\begin{aligned} \hat{\alpha}_1 &= 86.42 - 79.61 = 6.81 & \hat{\alpha}_2 &= 79.80 - 79.61 = 0.19 \\ \hat{\alpha}_3 &= 75.05 - 79.61 = -4.56 & \hat{\alpha}_4 &= 77.17 - 79.61 = -2.44 \end{aligned}$$

$$\begin{aligned}\hat{\beta}_1 &= 75.92 - 79.61 = -3.69 & \hat{\beta}_2 &= 81.57 - 79.61 = 1.96 \\ \hat{\beta}_3 &= 81.34 - 79.61 = 1.73\end{aligned}$$

$$\begin{array}{lll}\hat{\gamma}_{11} = 2.12 & \hat{\gamma}_{12} = 0.75 & \hat{\gamma}_{13} = -2.87 \\ \hat{\gamma}_{21} = -0.76 & \hat{\gamma}_{22} = -2.36 & \hat{\gamma}_{23} = 3.12 \\ \hat{\gamma}_{31} = -1.06 & \hat{\gamma}_{32} = -0.36 & \hat{\gamma}_{33} = 1.42 \\ \hat{\gamma}_{41} = -0.30 & \hat{\gamma}_{42} = 1.97 & \hat{\gamma}_{43} = -1.67\end{array}$$

Utilizando un ANOVA de dos sentidos para probar hipótesis

Un análisis de varianza de dos sentidos está diseñado para responder tres preguntas principales:

1. ¿El modelo aditivo vale?
2. ¿Si es así, la media del resultado es la misma para todos los niveles del factor renglón?
3. ¿Si es así, la media del resultado es la misma para todos los niveles del factor columna?

En general, se hacen las preguntas 2 y 3 solamente cuando se cree que el modelo aditivo puede valer. Esto último se analizará con profundidad más adelante. Las tres preguntas son indicadas para la realización de pruebas de hipótesis. Las hipótesis nulas para estas pruebas son:

1. Para probar si el modelo aditivo vale se prueba la hipótesis nula de que todas las interacciones son iguales a 0:

$$H_0 : \gamma_{11} = \gamma_{12} = \cdots = \gamma_{IJ} = 0$$

Si esta hipótesis nula es verdadera, el modelo aditivo vale.

2. Para probar si la media del resultado es igual para todos los niveles del factor renglón, se prueba la hipótesis nula de que todos los efectos renglón son iguales a 0:

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

Si esta hipótesis nula es verdadera, entonces la media del resultado es igual para todos los niveles del factor renglón.

3. Para probar si la media del resultado es igual para todos los niveles del factor columna, se prueba la hipótesis nula de que todos los efectos columna son iguales a 0:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_J = 0$$

Si esta hipótesis nula es verdadera, entonces la media del resultado es igual para todos los niveles del factor columna.

Ahora se describen las pruebas usuales para estas hipótesis nulas. Para que las pruebas sean válidas, las siguientes condiciones deben valer:

Supuestos para un ANOVA de dos sentidos

Las pruebas usuales de hipótesis ANOVA de dos sentidos son válidas bajo las siguientes condiciones:

1. El diseño debe estar completo.
2. El diseño debe ser balanceado.
3. El número de réplicas por tratamiento, K , deber ser al menos 2.
4. Dentro de cualquier tratamiento, las observaciones $X_{ij1}, \dots, X_{ijK}$ constituyen una muestra aleatoria simple de una población normal.
5. La varianza poblacional es igual para todos los tratamientos. Esta varianza se denota mediante σ^2 .

Igual que con un ANOVA de un sentido, las pruebas usuales para estas hipótesis nulas están basadas en las sumas de los cuadrados. Específicamente, son la suma de los cuadrados de renglón (SSA), de los cuadrados de columna (SSB), de los cuadrados de interacción (SSAB), y de los cuadrados del error (SSE). También es de interés la suma total de los cuadrados (SST), que es igual a la suma de las otras. Las fórmulas para estas sumas de cuadrados son:

$$SSA = JK \sum_{i=1}^I \hat{\alpha}_i^2 = JK \sum_{i=1}^I (\bar{X}_{i..} - \bar{X}_{...})^2 = JK \sum_{i=1}^I \bar{X}_{i..}^2 - IJK \bar{X}_{...}^2 \quad (9.58)$$

$$SSB = IK \sum_{j=1}^J \hat{\beta}_j^2 = IK \sum_{j=1}^J (\bar{X}_{.j.} - \bar{X}_{...})^2 = IK \sum_{j=1}^J \bar{X}_{.j.}^2 - IJK \bar{X}_{...}^2 \quad (9.59)$$

$$\begin{aligned} SSAB &= K \sum_{i=1}^I \sum_{j=1}^J \hat{\gamma}_{ij}^2 = K \sum_{i=1}^I \sum_{j=1}^J (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...})^2 \\ &= K \sum_{i=1}^I \sum_{j=1}^J \bar{X}_{ij.}^2 - JK \sum_{i=1}^I \bar{X}_{i..}^2 - IK \sum_{j=1}^J \bar{X}_{.j.}^2 + IJK \bar{X}_{...}^2 \end{aligned} \quad (9.60)$$

$$SSE = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (X_{ijk} - \bar{X}_{ij.})^2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}^2 - K \sum_{i=1}^I \sum_{j=1}^J \bar{X}_{ij.}^2 \quad (9.61)$$

$$SST = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (X_{ijk} - \bar{X}_{...})^2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}^2 - IJK \bar{X}_{...}^2 \quad (9.62)$$

Se puede ver, de las expresiones que están más a la derecha en las ecuaciones (9.58) a (9.62), que la suma total de los cuadrados, SST, es igual a la suma de las otras. Ésta es la identidad del análisis de varianza para un ANOVA de dos sentidos.

La identidad del análisis de varianza

$$SST = SSA + SSB + SSAB + SSE \quad (9.63)$$

Junto con cada suma de los cuadrados aparece una cantidad conocida como sus grados de libertad. Las sumas de los cuadrados y sus grados de libertad, en general, se presentan en una tabla ANOVA. La tabla 9.5 presenta los grados de libertad para cada suma de los cuadrados, junto con la fórmula más conveniente para el cálculo. Se indica que los grados de libertad para SST es la suma de los grados de libertad para las otras sumas de los cuadrados.

TABLA 9.5 Tabla ANOVA para un ANOVA de dos sentidos

Fuente	Grados de libertad	Suma de los cuadrados
Renglones (SSA)	$I - 1$	$JK \sum_{i=1}^I \hat{\alpha}_i^2 = JK \sum_{i=1}^I \bar{X}_{i..}^2 - IJK \bar{X}_{...}^2$
Columnas (SSB)	$J - 1$	$IK \sum_{j=1}^J \hat{\beta}_j^2 = IK \sum_{j=1}^J \bar{X}_{.j.}^2 - IJK \bar{X}_{...}^2$
Interacciones(SSAB)	$(I - 1)(J - 1)$	$K \sum_{i=1}^I \sum_{j=1}^J \hat{\gamma}_{ij}^2 = K \sum_{i=1}^I \sum_{j=1}^J \bar{X}_{ij.}^2 - JK \sum_{i=1}^I \bar{X}_{i..}^2 - IK \sum_{j=1}^J \bar{X}_{.j.}^2 + IJK \bar{X}_{...}^2$
Error (SSE)	$IJ(K - 1)$	$\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (X_{ijk} - \bar{X}_{ij.})^2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}^2 - K \sum_{i=1}^I \sum_{j=1}^J \bar{X}_{ij.}^2$
Total (SST)	$IJK - 1$	$\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (X_{ijk} - \bar{X}_{...})^2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}^2 - IJK \bar{X}_{...}^2$

Observe que la magnitud de SSA depende de la magnitud de los efectos de renglón *estimados* $\hat{\alpha}_i$. Por tanto, cuando los efectos de renglón *verdaderos* α_i son iguales a 0, SSA tenderá a ser más pequeña, y cuando algunos de los efectos de renglón verdaderos no sean iguales a 0, SSA será mayor. Por tanto, se rechazará a $H_0 : \alpha_1 = \dots = \alpha_I = 0$ cuando SSA sea suficientemente grande. De manera semejante, SSB tenderá a ser más pequeña cuando los efectos de columna verdaderos β_j sean iguales a 0 y mayor cuando algunos efectos de columna no sean 0, y SSAB tenderá a ser más pequeña cuando las interacciones verdaderas γ_{ij} sean todas iguales a 0 y mayor cuando algunas interacciones no sean 0. Por tanto, se rechazará $H_0 : \beta_1 = \dots = \beta_J = 0$ cuando SSB sea suficientemente grande, y se rechazará $H_0 : \gamma_{11} = \dots = \gamma_{IJ} = 0$ cuando SSAB sea muy grande.

Se puede determinar si SSA, SSB y SSAB son bastante grandes comparándolas con la suma de los cuadrados del error, SSE. Como en el ANOVA de un sentido (sección 9.1), SSE

depende solamente de las distancias entre las observaciones y sus propias medias de celda. Por tanto, SSE mide sólo la variación aleatoria inherente al proceso y no está afectado por los valores de los efectos de renglón, los de columna o las interacciones.

Para comparar SSA, SSB y SSAB con SSE, primero se divide cada suma de los cuadrados entre sus grados de libertad, produciendo cantidades conocidas como **medias cuadráticas**. Estas últimas se denotan mediante MSA, MSB, MSAB y MSE, están definidas así:

$$\begin{aligned} \text{MSA} &= \frac{\text{SSA}}{I-1} & \text{MSB} &= \frac{\text{SSB}}{J-1} & \text{MSAB} &= \frac{\text{SSAB}}{(I-1)(J-1)} \\ \text{MSE} &= \frac{\text{SSE}}{IJ(K-1)} \end{aligned} \quad (9.64)$$

Los estadísticos de prueba para las tres hipótesis nulas son los cocientes de MSA, MSB y MSAB con MSE. Las distribuciones nulas de estos estadísticos de prueba son las distribuciones F . Específicamente,

- Bajo $H_0: \alpha_1 = \cdots = \alpha_I = 0$, el estadístico $\frac{\text{MSA}}{\text{MSE}}$ tiene una distribución $F_{I-1, IJ(K-1)}$.
- Bajo $H_0: \beta_1 = \cdots = \beta_J = 0$, el estadístico $\frac{\text{MSB}}{\text{MSE}}$ tiene una distribución $F_{J-1, IJ(K-1)}$.
- Bajo $H_0: \gamma_{11} = \cdots = \gamma_{IJ} = 0$, el estadístico $\frac{\text{MSAB}}{\text{MSE}}$ tiene una distribución $F_{(I-1)(J-1), IJ(K-1)}$.

En la práctica, las sumas de cuadrados, las medias cuadráticas y los estadísticos de prueba generalmente se calculan utilizando una computadora. El siguiente resultado de MINITAB presenta la tabla ANOVA para los datos de la tabla 9.2.

Two-way ANOVA: Yield versus Catalyst, Reagent

Source	DF	SS	MS	F	P
Catalyst	3	877.56	292.521	9.36	0.000
Reagent	2	327.14	163.570	5.23	0.010
Interaction	6	156.98	26.164	0.84	0.550
Error	36	1125.33	31.259		
Total	47	2487.02			

S = 5.591

R-sq = 54.75%

R-Sq(adj) = 40.93%

Las etiquetas DF, SS, F y P se refieren a los grados de libertad, suma de cuadrados, media cuadrática, los estadísticos F , y P -valor, respectivamente. Como en un ANOVA de un sentido, la media cuadrática para el error (MSE) es una estimación de la varianza del error σ^2 , la cantidad “S” representa la raíz cuadrada de MSE y es una estimación de la desviación estándar del error σ . Las cantidades “R-sq” y “R-sq(adj)” se calculan con fórmulas similares a las del ANOVA de un sentido.

Ejemplo

9.15

Utilice la tabla ANOVA anterior para determinar si el modelo aditivo es creíble para los datos de producción. Si lo fuera, ¿se puede concluir que el catalizador o el reactivo afecta la producción?

Solución

Primero se verifica si el modelo aditivo es creíble. El P -valor para las interacciones es 0.55, que no es pequeño. Por tanto, no se rechaza la hipótesis nula de que todas las interacciones son iguales a 0, y se concluye que el modelo aditivo es creíble. En consecuencia, ahora la siguiente pregunta es si los factores renglón o columna afectan el resultado. Se ve de la tabla que el P -valor para los efectos de renglón (catalizador) es 0 aproximadamente, por lo que se concluye que el catalizador afecta la producción. De manera similar, el P -valor para los efectos de columna (reactivo) es pequeño (0.010), por lo que se concluye que el reactivo también afecta la producción.

Ejemplo

9.16

El artículo “Uncertainty in Measurements of Dermal Absorption of Pesticides” (W. Navidi y A. Bunge, en *Risk Analysis*, 2002:1175-1182) describe un experimento en el que se aplicó un pesticida a piel en diferentes concentraciones y distintos tiempos. El resultado es la cantidad del pesticida que absorbió la piel. El siguiente resultado de MINITAB presenta la tabla ANOVA. ¿El modelo aditivo es creíble? Si es así, ¿afecta la concentración o la duración la cantidad absorbida?

Two-way ANOVA: Absorbed versus Concentration, Duration

Source	DF	SS	MS	F	P
Concent	2	49.991	24.996	107.99	0.000
Duration	2	19.157	9.579	41.38	0.000
Interaction	4	0.337	0.084	0.36	0.832
Error	27	6.250	0.231		
Total	35	75.735			

Solución

El P -valor para la interacción es 0.832, por lo que se concluye que el modelo aditivo es creíble. Los P -valores tanto para la concentración como para la dosis son muy pequeños. Por tanto, se deduce que tanto la concentración como la duración afectan la cantidad absorbida.

Comprobación de los supuestos

Se puede utilizar una gráfica de residuos para comprobar el supuesto de las variancias iguales, y una gráfica de probabilidad normal de los residuos para verificar la normalidad. La gráfica de residuos traza los residuos $X_{ijk} - \bar{X}_{ij}$, contra los valores ajustados, que son las medias muestrales $\bar{X}_{ij}$. Las figuras 9.6 y 9.7 presentan tanto una gráfica de probabilidad normal co-

mo una de residuos para los datos de la producción que se encuentran en la tabla 9.2. Parece que los supuestos se satisfacen bien.

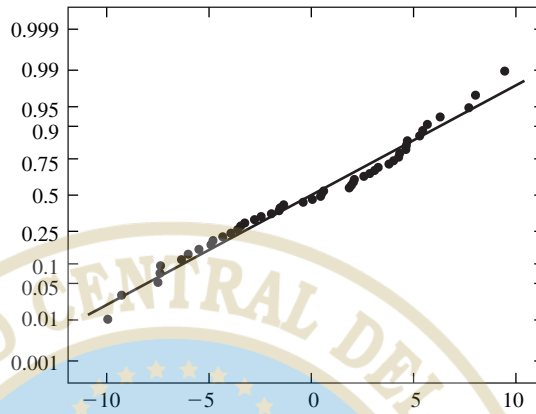


FIGURA 9.6 Gráfica de probabilidad normal para los residuos a partir de los datos de producción. No hay evidencia de fuerte desviación de la normalidad.

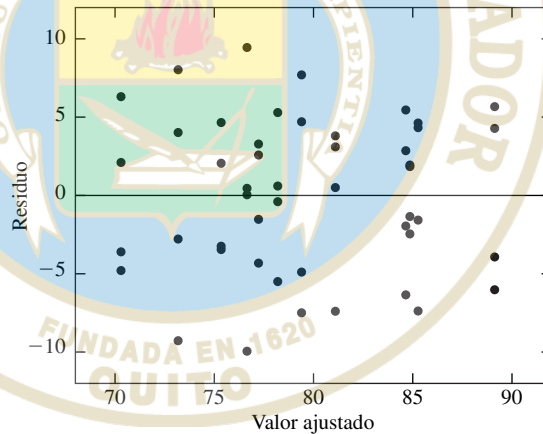


FIGURA 9.7 Gráfica de residuos para los datos de producción. No hay evidencias contra el supuesto de las varianzas iguales.

No interprete los efectos principales cuando el modelo aditivo no vale

Cuando las interacciones son suficientemente pequeñas para que el modelo aditivo sea creíble, la interpretación de los efectos principales es muy directa, como se muestra en los ejemplos 9.15 y 9.16. Sin embargo, cuando el modelo aditivo no vale, no siempre es fácil interpretar los efectos principales. He aquí un ejemplo hipotético para ilustrar el punto. Su-

ponga que un proceso está operando bajo condiciones obtenidas al variar dos factores diferentes en dos niveles cada uno. Se hacen dos operaciones en cada una de las cuatro combinaciones de niveles de renglón y de columna. Cada vez se mide la producción y se obtienen los resultados que se presentan en la tabla siguiente.

Nivel de renglón	Nivel de columna	
	1	2
1	51, 49	43, 41
2	43, 41	51, 49

Obviamente, si se desea maximizar la producción, los factores de renglón y de columna son importantes, se quiere aparear ya sea el nivel de renglón 1 con el nivel de columna 1 o el nivel de renglón 2 con el nivel de columna 2.

Ahora mire la tabla siguiente de ANOVA.

Source	DF	SS	MS	F	P
Row	1	0.0000	0.0000	0.00	1.000
Column	1	0.0000	0.0000	0.00	1.000
Interaction	1	128.00	128.00	64.00	0.001
Error	4	8.0000	2.0000		
Total	7	136.00			

Los efectos principales de la suma de los cuadrados tanto para el renglón como para la columna son iguales a 0, y sus *P*-valores son iguales a 1, que es el *P*-valor más grande que puede haber. Si se continúa con el procedimiento usado en los ejemplos 9.15 y 9.16, se concluirá que ni el factor de renglón ni el factor de columna afectan la producción. Pero es claro de los datos que los factores de renglón y de columna afectan la producción. Lo que está ocurriendo es que ambos no importan *en promedio*. El nivel 1 del factor renglón es mejor si se usa el nivel 1 del factor columna, y el nivel 2 del factor renglón es mejor si se usa nivel 2 del factor columna. Cuando se determina el promedio sobre dos niveles del factor columna, los niveles del factor renglón tienen la misma media de producción. Asimismo, los niveles columna tienen la misma media de producción cuando se promedian sobre los niveles del factor renglón. Cuando los efectos de los niveles de renglón dependen de cuáles niveles columna se han apareado, y viceversa, los efectos principales pueden ser engañosos.

El *P*-valor para las interacciones es el que indica que no se trata de interpretar los efectos principales. Este *P*-valor es muy pequeño, por lo que se rechaza el modelo aditivo. Entonces se sabe que algunas de las interacciones son diferentes de cero, por lo que los efectos de los niveles de renglón dependen de los niveles de columna, y viceversa. En consecuencia, cuando se rechaza el modelo aditivo no se debe tratar de interpretar los efectos principales. Se necesita observar en las mismas medias de celda para determinar cómo afectan el resultado las diferentes combinaciones de los niveles de renglón y de columna.

Resumen

En un análisis de varianza de dos sentidos:

- Si el modelo aditivo *no* es rechazado, entonces las pruebas de hipótesis para los efectos principales se pueden utilizar para determinar si los factores renglón o columna afectan el resultado.
- Si el modelo aditivo *es* rechazado, entonces las pruebas de hipótesis para los efectos principales no se deben utilizar. En lugar de eso, se deben examinar las medias de celda para determinar cómo afectan el resultado las diferentes combinaciones de los niveles de renglón y de columna.

Ejemplo

9.17

El espesor de la capa de dióxido de silicio en una placa semiconductora es crucial para su desempeño. En el artículo “Virgin Versus Recycled Wafers for Furnace Qualification: Is the Expense Justified?” (V. Czitrom y J. Reece, en *Statistical Case Studies for Process Improvement*, SIAM-ASA, 1997:87-103) se midió el espesor de la capa de óxido para tres tipos de placas: placas vírgenes, placas recicladas en la empresa, y placas recicladas por un proveedor externo. Además se utilizaron diferentes posiciones en el horno para crecer la capa de óxido. Se realizó un ANOVA de dos sentidos por tres operaciones en una empresa de fabricación de placas para los tres tipos de placas en tres ubicaciones del horno. Los datos se presentan en la tabla siguiente, seguida por el resultado de MINITAB.

Ubicación en el horno	Tipo de placa	Espesor de la capa de óxido (A)		
1	Virgen	90.1	90.7	89.4
1	En la empresa	90.4	88.8	90.6
1	Externa	92.6	90.0	93.3
2	Virgen	91.9	88.6	89.7
2	En la empresa	90.3	91.9	91.5
2	Externa	88.3	88.2	89.4
3	Virgen	88.1	90.2	86.6
3	En la empresa	91.0	90.4	90.2
3	Externa	91.5	89.8	89.8

Two-way ANOVA for Thickness versus Wafer, Location

Source	DF	SS	MS	F	P
Wafer	2	5.8756	2.9378	2.07	0.155
Location	2	4.1089	2.0544	1.45	0.262
Interaction	4	21.349	5.3372	3.76	0.022
Error	18	25.573	1.4207		
Total	26	56.907			

Debido a que las placas recicladas son más baratas, la compañía espera que no haya ninguna diferencia en el espesor de la capa de óxido entre los tres tipos de chips. Si es posible, determine si los datos son consistentes con la hipótesis de ninguna diferencia. Si no posible, explique por qué.

Solución

El *P*-valor para las interacciones es 0.022, que es pequeño. Por tanto, el modelo aditivo no es creíble por lo que no se pueden interpretar los efectos principales. Un aspecto positivo es hacer una tabla de las medias de la celda. La tabla 9.6 presenta la media muestral para cada tratamiento.

TABLA 9.6 Medias muestrales para cada tratamiento

Ubicación en el horno	Tipo de placa			Media del renglón
	Virgen	En la empresa	Externa	
1	90.067	89.933	91.967	90.656
2	90.067	91.233	88.633	89.978
3	88.300	90.533	90.367	89.733
Media de la columna	89.478	90.566	90.322	

De la tabla 9.6 se puede ver que los espesores varían entre los tipos de placa, pero ninguno de éstos produce la capa de óxido más gruesa o más angosta. Por ejemplo, en la ubicación 1 del horno las placas recicladas externamente producen la capa más gruesa, mientras las placas recicladas en la empresa producen la más angosta. En la ubicación 2 del horno se invierte el orden: las placas en la empresa producen la capa más gruesa mientras las externas la más angosta. Esto último es atribuible a la interacción de la ubicación del horno y del tipo de placa.

Un ANOVA de dos sentidos no es lo mismo que dos ANOVA de un sentido

En el ejemplo 9.17 se presentó un ANOVA de dos sentidos con tres niveles de renglón y tres de columna, para un total de nueve tratamientos. Si se corrieran los ANOVA de un sentido separados, habría solamente seis tratamientos. Lo anterior significa que en la práctica correr por separado los ANOVA de un sentido para cada factor podría ser menos costoso que correr un ANOVA de dos sentidos. Desgraciadamente, a veces este diseño de “uno-a-la-vez” se usa en la práctica por esta razón. Es importante darse cuenta que correr análisis de un sentido de factores individuales puede dar resultados engañosos cuando las interacciones están presentes. Para ver esto, mire la tabla 9.6. Suponga que un ingeniero está tratando de encontrar la combinación de horno y ubicación que producirá la capa de óxido más angosta. Primero corra el proceso una vez en cada ubicación del horno, utilizando placas recicladas de la empresa, porque esas placas son las que actualmente se usan en la producción. La ubicación 1 del horno produce la capa más angosta para las placas de la empresa. Ahora el ingeniero corre el proceso una vez para cada tipo de placa, todas en la ubicación 1, que fue la mejor para las placas de la empresa. De los tres tipos de placa las placas en la empresa producen la capa más angosta en la ubicación 1. Así que la conclusión que se obtiene del análisis uno-a-la-vez es que las capas más angostas se producen mediante la combinación de placas de la empresa en la ubica-

ción 1 del horno. Una mirada a la tabla 9.6 muestra que la conclusión es falsa. Hay dos combinaciones de ubicación de horno y tipo de placa que producen capas más angostas que ésta.

El método de uno-a-la-vez supone que la placa que produce las capas más angostas en una ubicación producirá las más angostas en todas las ubicaciones, y que la ubicación que produce las capas más angostas para un tipo de placa producirá las capas más angostas para todos los tipos. Esto último equivale a suponer que no hay ninguna interacción entre los factores, que en el caso de las placas y las ubicaciones es incorrecto. En resumen, el método uno-a-la-vez falla porque no puede detectar las interacciones entre los factores.

Resumen

- Cuando hay dos factores se debe usar un diseño de dos factores.
- El examen de un factor a la vez no puede revelar las interacciones entre los factores.

Gráficas de interacción

Las gráficas de interacción pueden ayudar a visualizar las interacciones. La figura 9.8 presenta una gráfica de interacción para los datos de placa. Se describe el método con el que se construyó esta gráfica. El eje vertical representa la respuesta: el espesor de capa. Se elige un factor para que se represente en el eje horizontal. Se eligió la ubicación de horno; habría sido igualmente aceptable haber elegido el tipo de placa. Ahora se continúa a través de los niveles del factor tipo de placa. Se inicia con las placas externas. Las tres medias de la celda para placas externas, como se muestran en la tabla 9.6, son 91.967, 88.633, y 90.367, que corresponden a las ubicaciones del horno 1, 2 y 3, respectivamente. Estos valores se trazan arriba de sus ubicaciones respectivas en el horno y se conectan con segmentos de recta. Este procedimiento se repite en los otros dos tipos de placa para terminar la gráfica.

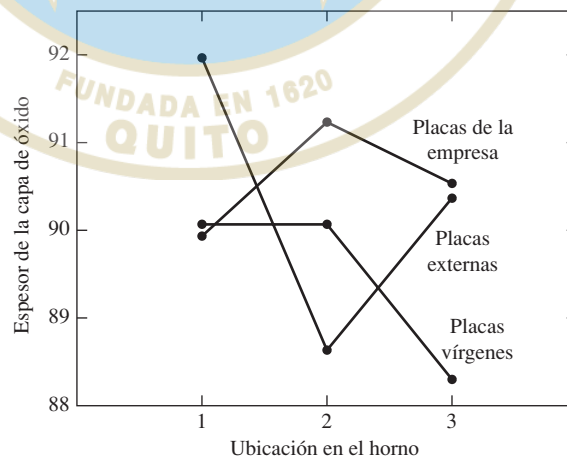


FIGURA 9.8 Gráfica de interacción para los datos de placa. Las rectas están lejos de ser paralelas, ello indica una importante interacción entre los factores.

En cuanto a los datos de la placa, las medias para placas externas siguen un patrón considerablemente diferente que las de los otros dos tipos de placa. Ésta es la fuente de la interacción importante y es la razón por la cual los efectos principales de placa y tipo de horno no se pueden interpretar con facilidad. En comparación, para los datos perfectamente aditivos, para los cuales las interacciones estimadas $\hat{\gamma}_{ij}$ son iguales a 0, los segmentos de recta en la gráfica de interacción son paralelos. La figura 9.9 ilustra este caso hipotético.

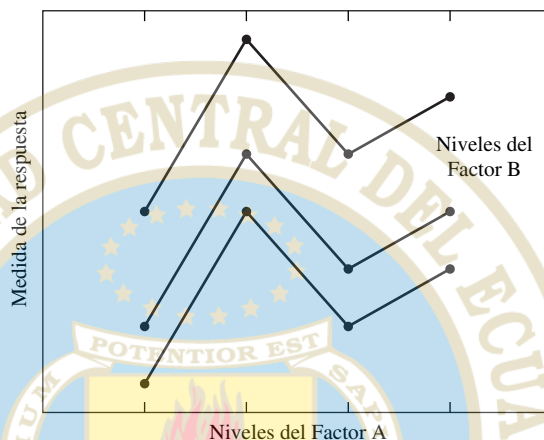


FIGURA 9.9 Gráfica de interacción para datos hipotéticos con estimaciones de la interacción $\hat{\gamma}_{ij}$ iguales a 0. Los segmentos de recta son paralelos.

La figura 9.10 presenta una gráfica de interacción para los datos de producción. Las medias de celda se presentaron en la tabla 9.4. Las líneas no son paralelas, pero sus pendientes se acoplan mejor que para los datos de placa. Esto último indica que las estimaciones de la interacción son diferentes de cero, pero son más pequeñas que los datos de la placa. En efecto, el P -valor para la prueba de la hipótesis nula de ninguna interacción fue 0.550 (véase la p. 667). La desviación del paralelismo que se presenta en la figura 9.10 es, por tanto, bastante pequeña para ser consistente con la hipótesis de ninguna interacción.

Comparaciones múltiples en un ANOVA de dos sentidos

Se utiliza una prueba F para probar la hipótesis nula de que todos los efectos renglón (o todos los efectos columna) sean iguales a 0. Si la hipótesis nula es rechazada, se puede concluir que algunos de dichos efectos difieren entre sí. Pero la hipótesis de la prueba no dice cuáles son diferentes del resto. Si el modelo aditivo es creíble, entonces se puede aplicar el método de comparaciones múltiples de Tukey (relacionado con el método de Tukey-Kramer descrito en la sección 9.2) para determinar cuáles pares de los efectos renglón o columna difieren entre sí. El método se describe en el cuadro siguiente.

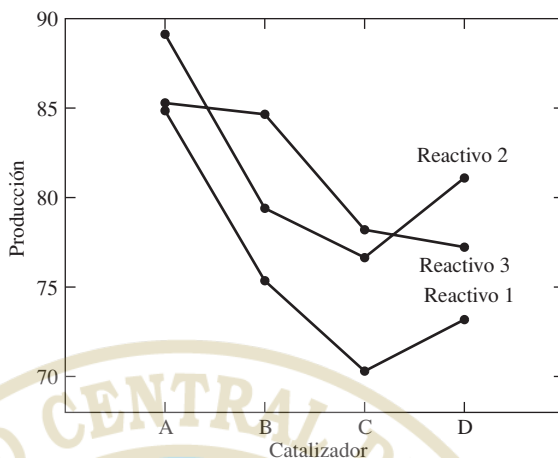


FIGURA 9.10 Gráfica de interacción para los datos de producción.

Método de Tukey para intervalos de confianza y pruebas de hipótesis para un ANOVA de dos sentidos

Sea I el número de niveles de un factor renglón, J el número de niveles del factor columna, y K el tamaño muestral para cada tratamiento. Entonces, si el modelo aditivo es creíble, los intervalos de confianza simultáneos de nivel $100(1 - \alpha)\%$ para todas las diferencias $\alpha_i - \alpha_j$ (o para todas las diferencias $\beta_i - \beta_j$) son

$$\hat{\alpha}_i - \hat{\alpha}_j \pm q_{I, IJ(K-1), \alpha} \sqrt{\frac{MSE}{JK}} \quad \hat{\beta}_i - \hat{\beta}_j \pm q_{J, IJ(K-1), \alpha} \sqrt{\frac{MSE}{IK}}$$

Se tiene una confianza de $100(1 - \alpha)\%$ de que los intervalos de confianza de Tukey contengan el verdadero valor de la diferencia $\alpha_i - \alpha_j$ (o $\beta_i - \beta_j$) para cada i y j .

Para cada par de niveles i y j para los cuales $|\hat{\alpha}_i - \hat{\alpha}_j| > q_{I, IJ(K-1), \alpha} \sqrt{\frac{MSE}{JK}}$, la hipótesis nula $H_0 : \alpha_i - \alpha_j = 0$ se rechaza con un nivel de α .

Para cada par de niveles i y j para los cuales $|\hat{\beta}_i - \hat{\beta}_j| > q_{J, IJ(K-1), \alpha} \sqrt{\frac{MSE}{IK}}$, la hipótesis nula $H_0 : \beta_i - \beta_j = 0$ se rechaza al nivel α .

Ejemplo

9.18

En el ejemplo 9.14 los efectos principales y las interacciones fueron calculados para los datos de producción de la tabla 9.2. Una tabla ANOVA para estos datos se presentó en la p. 667. Si es adecuado, utilice el método de Tukey para determinar qué pares de catalizadores y de reactivos difieren con un nivel de 5%, en su efecto sobre la producción.

Solución

De la tabla ANOVA, el P -valor para las interacciones es 0.550. Por tanto, el modelo aditivo es creíble, por lo que es adecuado utilizar el método de Tukey. El catalizador es el factor renglón y el reactivo lo es de columna, por lo que $I = 4$, $J = 3$ y $K = 4$. De la tabla ANOVA, $MSE = 31.259$.

Primero se encuentran todos los pares para los cuales los efectos de renglón son diferentes con un nivel de 5%. Para los efectos de renglón se debe utilizar el valor $q_{4, 36, .05}$. Este valor no se encuentra en la tabla de rango studentizado (tabla A.8, Apéndice A). Por tanto, se utiliza el valor $q_{4, 36, .05} = 3.85$, que está cercano (sólo ligeramente mayor que) $q_{4, 36, .05}$. Se calcula $q_{4, 30, .05} \sqrt{MSE/JK} = 3.85 \sqrt{31.259/12} = 6.21$.

En el ejemplo 9.14 los efectos de renglón estimados fueron calculados de

$$\hat{\alpha}_1 = 6.81 \quad \hat{\alpha}_2 = 0.19 \quad \hat{\alpha}_3 = -4.56 \quad \hat{\alpha}_4 = -2.44$$

Los pares de efectos de renglón cuyas diferencias son mayores a 6.21 son $\hat{\alpha}_1$ y $\hat{\alpha}_2$, $\hat{\alpha}_1$ y $\hat{\alpha}_3$, y $\hat{\alpha}_1$ y $\hat{\alpha}_4$. Se concluye que la media de la producción del catalizador A difiere de la media de las producciones de los catalizadores B, C y D, pero no se puede concluir que la media de las producciones de los catalizadores B, C, y D difieren entre sí.

Ahora se encuentran todos los pares para los cuales los efectos de columna difieren con un nivel de 5%. Para los efectos de columna se debe utilizar el valor, $q_{3, 36, .05}$, pero debido a que este valor no se encuentra en la tabla de rango studentizado, se utilizará el valor, $q_{3, 30, .05} = 3.49$. Se calcula $q_{3, 30, .05} \sqrt{MSE/IK} = 3.49 \sqrt{31.259/16} = 4.88$.

En el ejemplo 9.14 los efectos de columna estimados fueron calculados de

$$\hat{\beta}_1 = -3.69 \quad \hat{\beta}_2 = 1.96 \quad \hat{\beta}_3 = 1.73$$

Los pares de los efectos de columna cuyas diferencias son mayores a 4.88 son $\hat{\beta}_1$ y $\hat{\beta}_2$ y $\hat{\beta}_1$ y $\hat{\beta}_3$. Se concluye que la media de la producción del reactivo 1 es diferente de la media de las producciones de los reactivos 2 y 3, pero no se puede concluir que la media de las producciones de los reactivos 2 y 3 difieran entre sí.

ANOVA de dos sentidos cuando $K = 1$

Las pruebas F que se han presentado requieren del supuesto que el tamaño muestral K para cada tratamiento sea al menos de 2. La razón para esto último consiste en que cuando $K = 1$, la suma de los cuadrados del error (SSE) es igual a 0, puesto que $X_{ijk} = \bar{X}_{ij}$ para cada i y j . Además, los grados de libertad para SSE, que están dados por $IJ(K - 1)$, es igual a 0 cuando $K = 1$.

Cuando $K = 1$, no se puede realizar un ANOVA de dos sentidos, a menos que se tenga la seguridad de que el modelo aditivo vale. En este caso, puesto que se ha supuesto que las interacciones sean iguales a cero, se pueden utilizar la media cuadrática para la interacción (MSAB; véase la ecuación 9.64) y sus grados de libertad; en este sentido, en lugar de MSE, pruébense los efectos principales de renglón y de columna.

Factores aleatorios

El análisis de los experimentos de dos factores se ha concentrado en el caso donde ambos factores son fijos. Se dice que un experimento sigue un **modelo de efectos fijos**. También se pueden diseñar los experimentos para que uno o los dos factores sean aleatorios. Si ambos

factores lo son, se dice que el experimento sigue un **modelo de efectos aleatorios**. Si un factor es fijo y el otro es aleatorio, se dice que el experimento sigue un **modelo mezclado**.

Para el caso de un factor, el análisis es el mismo tanto para modelos de efectos fijos como aleatorios, mientras que la hipótesis nula de la prueba es diferente. En este contexto, para el caso de dos factores, tanto los métodos de análisis como las hipótesis nulas difieren entre los modelos de efectos fijos, los aleatorios y los mezclados. Los métodos para modelos en los cuales uno o más efectos son aleatorios se pueden encontrar en textos más avanzados, como Hocking (2003).

Diseños desbalanceados

Se ha supuesto que el diseño sea balanceado; es decir, que el número de réplicas sea igual en cada tratamiento. Los métodos que aquí se han descrito no son aplicables a diseños desbalanceados. Sin embargo, los diseños desbalanceados que están completos se pueden analizar con los métodos de regresión múltiple. Se puede consultar un texto avanzado, como Draper y Smith (1998), para los detalles.

Ejercicios para la sección 9.3

- Para tasar el efecto del tipo de anillo del pistón y del tipo de aceite en el desgaste del anillo del pistón, se estudiaron tres tipos del primero y cuatro tipos del segundo. Se midió el número de miligramos de material perdido por el anillo en cuatro horas de uso en tres réplicas de un experimento, se realizaron para cada una de las 12 combinaciones del tipo de aceite y del tipo de anillo de pistón. Con el tipo de aceite como el efecto renglón y tipo de anillo del pistón como el efecto columna, se observaron las siguientes sumas de los cuadrados: $SSA = 1.0926$, $SSB = 0.9340$, $SSAB = 0.2485$, $SSE = 1.7034$.
 - ¿Cuántos grados de libertad hay para el efecto del tipo de aceite?
 - ¿Cuántos grados de libertad hay para el efecto del tipo de anillo del pistón?
 - ¿Cuántos grados de libertad hay para las interacciones?
 - ¿Cuántos grados de libertad hay para el error?
 - Construya una tabla ANOVA. Puede dar rangos para los P -valores.
 - ¿El modelo aditivo es creíble? Proporcione el valor del estadístico de prueba y el P -valor.
 - ¿Es creíble que los efectos principales del tipo de aceite son todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
 - ¿Es creíble que los efectos principales del tipo de anillo del pistón sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
- Una tienda de maquinaria utiliza tres máquinas para pulir con precisión rodillos para levas. Se emplearon tres operarios para pulir los rodillos en las máquinas. En un experimento para determinar si hay diferencias en el producto entre las máquinas o entre sus operarios, cada uno de éstos trabajó en cada máquina durante cuatro días diferentes. El resultado medido fue la producción diaria de partes que satisfacían las especificaciones. Con el operario como el efecto renglón y la máquina como el efecto columna, se observaron las siguientes sumas de los cuadrados: $SSA = 3\,147.0$, $SSB = 136.5$, $SSAB = 411.7$, $SSE = 1\,522.0$.
 - ¿Cuántos grados de libertad hay para el efecto operario?
 - ¿Cuántos grados de libertad hay para el efecto máquina?
 - ¿Cuántos grados de libertad hay para las interacciones?
 - ¿Cuántos grados de libertad hay para el error?
 - Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
 - ¿El modelo aditivo es creíble? Proporcione el valor del estadístico de prueba y el P -valor.

- g) ¿Es creíble que los efectos principales de operador sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
- h) ¿Es creíble que los efectos principales de máquina sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
3. Un experimento para determinar el efecto de la temperatura del molde sobre la fuerza de tensión implicaba tres aleaciones diferentes y cinco temperaturas de molde distintas. Se fundieron cuatro elementos de cada aleación para cada temperatura de molde. Con la temperatura de molde como el factor renglón y la aleación como el factor columna, las sumas de los cuadrados fueron: $SSA = 69\,738$, $SSB = 8\,958$, $SSAB = 7\,275$ y $SST = 201\,816$.
- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Es creíble el modelo aditivo? Explique.
- c) ¿Es creíble que los efectos principales de la temperatura de molde sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Es creíble que los efectos principales de aleación sean todos iguales a 0? Proporcione el estadístico de prueba del valor y el P -valor.
4. El efecto de la presión de endurecimiento sobre la fuerza de enlace se probó en cuatro adhesivos diferentes. Había tres niveles de presión de endurecimiento. Se realizaron cinco réplicas para cada combinación de presión de endurecimiento y adhesivo. Con adhesivo como el factor renglón y la presión de endurecimiento como el factor columna, las sumas de los cuadrados fueron: $SSA = 155.7$, $SSB = 287.9$, $SSAB = 156.7$ y $SST = 997.3$.
- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Es creíble el modelo aditivo? Explique.
- c) ¿Es creíble que los efectos principales de la presión de endurecimiento sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Es creíble que los efectos principales de adhesivo sean todos iguales a 0? Proporcione el valor del estadístico de prueba y el P -valor.
5. El artículo “Change in Creep Behavior of Plexiform Bone with Phosphate Ion Treatment” (R. Regimbal, C. DePaula y N. Guzelsu, en *Bio-Medical Materials and Engineering*, 203:11-25) describe un experimento para estudiar los efectos con soluciones salinas y con soluciones del ion fosfato sobre las propiedades mecánicas del hueso plexiforme. La tabla siguiente presenta las mediciones de tensión producidas para seis elementos tratados con la solución salina (NaCl) o con la solución de ion fosfato (Na_2HPO_4), a 25 o 37°C. (El artículo presenta solamente las medias y las desviaciones estándar; los valores en la tabla son consistentes con éstos.)

Solución	Temperatura	Tensión producida (MPa)						
NaCl	25°C	138.40	130.89	94.646	96.653	116.90	88.215	
NaCl	37°C	92.312	147.28	116.48	88.802	114.37	90.737	
Na_2HPO_4	25°C	120.18	129.43	139.76	132.75	137.23	121.73	
Na_2HPO_4	37°C	123.50	128.94	102.86	99.941	161.68	136.44	

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto de solución (NaCl contra Na_2HPO_4) sobre la tensión producida interpretando los efectos principales de la solución? Si es así, interprete los efectos principales, incluyendo el estadístico de prueba adecuado y el P -valor. Si no, explique por qué.
- e) ¿Puede describirse el efecto de la temperatura sobre la tensión producida interpretando los efectos principales de la temperatura? Si es así, interprete los efectos principales, incluyendo el estadístico de prueba adecuado y el P -valor. Si no, explique por qué.

6. Se realizó un estudio para determinar los efectos de dos factores sobre lo cremoso del jabón. Los dos factores eran tipo de agua (de la llave contra desionizada) y glicerina (presente o ausente). El resultado medido fue la cantidad de espuma producida (en ml). Se repitió el experimento tres veces para cada combinación de los factores. Los resultados se presentan en la tabla siguiente.

Tipo de agua	Glicerina	Espuma (mL)		
Desionizada	Ausente	168	178	168
Desionizada	Presente	160	197	200
De la llave	Ausente	152	142	142
De la llave	Presente	139	160	160

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto del tipo de agua en la cantidad de espuma al interpretar los efectos principales del tipo de agua? Si es así, intérpretelos. Si no, explique por qué.
- e) ¿Puede describirse el efecto de glicerina en la cantidad de espuma interpretando los efectos principales de glicerina? Si es así, intérpretelos. Si no, explique.
7. Se realizó un experimento para medir los efectos de dos factores sobre la capacidad de las soluciones de limpieza para eliminar aceite de una pieza de tela. Los factores fueron la concentración del jabón (en % por peso) y la fracción de ácido láurico en la solución. Se repitió dos veces el experimento para la combinación de cada uno de los factores. El resultado medido fue el porcentaje de aceite eliminado. Los resultados se presentan en la tabla siguiente.

Peso % de jabón	Fracción de ácido láurico	% de aceite eliminado	
15	10	52.8	54.0
15	30	57.8	53.3
25	10	56.4	58.4
25	30	42.7	45.1

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto de concentración de jabón sobre la cantidad de aceite eliminado al interpretar los efectos principales de concentración de jabón? Si es así, interprete los efectos principales. Si no, explique.
- e) ¿Puede describirse el efecto de fracción de ácido láurico sobre la cantidad de aceite eliminado con la interpretación de los efectos principales de la fracción de ácido láurico? Si es así, intérpretelos. Si no, explique.
8. El artículo "A 4-Year Sediment Trap Record of Alkenones from the Filamentous Upwelling Region Off Cape Blanc, NW Africa and a Comparison with Distributions in Underlying Sediments" (P. Müller y G. Fiseher, en *Deep Sea Research*, 2001:1877-1903) estudió registros de sedimentos atrapados para valorar la transferencia de señales de agua de superficie en el registro geológico. Los datos en la tabla siguiente son las mediciones del flujo de masa total (en mg/m^2 por día) para trampas en dos ubicaciones y a profundidades diferentes.

Ubicación	Profundidad	Flujo								
A	Superior	109.8	86.5	150.5	69.8	63.2	107.8	72.4	74.4	
A	Inferior	163.7	139.4	176.9	170.6	123.5	142.9	130.3	111.6	
B	Superior	93.2	123.6	143.9	163.2	82.6	63.0	196.7	160.1	
B	Inferior	137.0	88.3	53.4	104.0	78.0	39.3	117.9	143.0	

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) Calcule todas las medias de celda. Utilícelas para describir la manera en que afectan al flujo la profundidad y la ubicación.
9. Las articulaciones artificiales constan de una pelota de cerámica montada sobre un cono. El artículo “Friction in Orthopaedic Zirconia Taper Assemblies” (W. Macdonald, A. Aspenberg y cols., en *Proceedings of the Institution of Mechanical Engineers*, 2000:685-692) presenta datos del coeficiente de fricción para un empujón hacia delante de 2 kN para el ensamble cónico hecho de dos aleaciones de circonio y empleando tres diferentes largos. Se hicieron cinco mediciones para cada combinación de material y de largo. Los resultados presentados en la tabla siguiente son consistentes con las medias de celda y desviaciones estándar presentadas en el artículo.

Material del cono	Longitud	Coeficiente de fricción				
CPTi-ZrO ₂	Corto	0.254	0.195	0.281	0.289	0.220
CPTi-ZrO ₂	Medio	0.196	0.220	0.185	0.259	0.197
CPTi-ZrO ₂	Largo	0.329	0.481	0.320	0.296	0.178
TiAlloy-ZrO ₂	Corto	0.150	0.118	0.158	0.175	0.131
TiAlloy-ZrO ₂	Medio	0.180	0.184	0.154	0.156	0.177
TiAlloy-ZrO ₂	Largo	0.178	0.198	0.201	0.199	0.210

- a) Calcule los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba, su distribución nula, y el P -valor.
- d) ¿Puede describirse el efecto del material sobre el coeficiente de fricción interpretando los efectos principales del material? Si es así, intérpretelos. Si no, explique.
- e) ¿Puede describirse el efecto del largo sobre el coeficiente de fricción al interpretar los efectos principales del largo? Si es así, intérpretelos. Si no, explique.
10. El artículo “Anodic Fenton Treatment of Treflan MTF” (D. Saltmiras y A. Lemley, en *Journal of Environmental Science and Health*, 2001:261-274) describe un experimento de dos factores diseñado para estudiar la absorción del herbicida trifluralin. Los factores son la concentración inicial de trifluralin y el cociente de entrega Fe₂:H-2O-2. Hubo tres réplicas por cada tratamiento. Los resultados que se presentan en la tabla siguiente son consistentes con las medias y las desviaciones estándar notificadas en el artículo.

Concentración inicial (M)	Cociente de entrega	Absorción (%)		
15	1:0	10.90	8.47	12.43
15	1:1	3.33	2.40	2.67
15	1:5	0.79	0.76	0.84
15	1:10	0.54	0.69	0.57
40	1:0	6.84	7.68	6.79
40	1:1	1.72	1.55	1.82
40	1:5	0.68	0.83	0.89
40	1:10	0.58	1.13	1.28
100	1:0	6.61	6.66	7.43
100	1:1	1.25	1.46	1.49
100	1:5	1.17	1.27	1.16
100	1:10	0.93	0.67	0.80

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba, su distribución nula, y el P -valor.

11. Con referencia al ejercicio 10. Los tratamientos con un cociente de entrega de 1:0 fueron controlados, o en blanco. Después del experimento se descubrió que los niveles evidentes altos de absorción en tales controles se debían en gran parte a la volatilidad del trifluralin. Elimine los tratamientos de control de los datos.

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba, su distribución nula y el P -valor.
- d) Construya una gráfica de interacción. Explique cómo la gráfica ilustra el grado con el que las interacciones están presentes.

12. El artículo “Use of Taguchi Methods and Multiple Regression Analysis for Optimal Process Development of High Energy Electron Beam Case Hardening of Cast Iron” (M. Jean y Y. Tzeng, en *Surface Engineering*, 2003:150-156) describe un experimento factorial diseñado para determinar cómo los factores de un proceso de haces de electrones de alta-energía afectan la dureza en metales. Los resultados para dos factores, cada uno con tres niveles, se presentan en la tabla siguiente. El factor A es la velocidad del viaje en mm/s, y el factor B es el voltaje de aceleración en volts. El resultado es la dureza de Vickers. Hubo seis réplicas para cada tratamiento. En el artículo se estudió un total de siete factores; los dos que aquí se presentan son los que se encontraron más significativos.

A	B	Dureza					
10	10	875	896	921	686	642	613
10	25	712	719	698	621	632	645
10	50	568	546	559	757	723	734
20	10	876	835	868	812	796	772
20	25	889	876	849	768	706	615
20	50	756	732	723	681	723	712
30	10	901	926	893	856	832	841
30	25	789	801	776	845	827	831
30	50	792	786	775	706	675	568

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto de la velocidad de viaje sobre la dureza interpretando los efectos principales de la velocidad de viaje? Si es así, interprete los efectos principales utilizando comparaciones múltiples con un nivel de 5% si es necesario. Si no, explique.
- e) ¿Puede describirse el efecto del voltaje de aceleración en la dureza interpretando los efectos principales de voltaje de aceleración? Si es así, intérpretelos, utilizando comparaciones múltiples con un nivel de 5%. Si no, explique.

13. El artículo “T-Bracing for Stability of Compression Webs in Wood Trusses” (R. Leichti, I. Hofaker, y cols., en *Journal of Structural Engineering*, 2002:374-381) presenta los resultados de los experimentos de la carga crítica de doblaje (en kN) para ensamblajes reforzados en T, la carga se estimó por un método de elemento finito. La tabla siguiente presenta datos en los que los factores son la longitud del miembro de lado y su sistema de unión. Hubo diez réplicas para cada combinación de factores. Los datos son consistentes con las medias y las desviaciones estándar dadas en el artículo.

Sistema de unión	Longitud	Carga crítica de doblaje									
Adhesivo	Un cuarto	7.90	8.71	7.72	8.88	8.55	6.95	7.07	7.59	7.77	7.86
Adhesivo	La mitad	14.07	13.82	14.77	13.39	11.98	12.72	9.48	13.59	13.09	12.09
Adhesivo	Completo	26.80	28.57	24.82	23.51	27.57	25.96	24.28	25.68	21.64	28.16
Clavo	Un cuarto	6.92	5.38	5.38	5.89	6.07	6.37	7.14	6.71	4.36	6.78
Clavo	La mitad	9.67	9.17	10.39	10.90	10.06	9.86	10.41	10.24	9.31	11.99
Clavo	Completo	20.63	21.15	24.75	20.76	21.64	21.47	25.25	22.52	20.45	20.38

- a) Calcule todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor del estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto del método adjunto (clavo contra adhesivo) sobre la carga crítica de doblaje interpretando los efectos principales sobre el sistema de unión? Si es así, interpréte los. Si no, explique.
- e) ¿Puede describirse el efecto de la longitud del miembro de lado sobre la carga crítica de doblaje interpretando los efectos principales de la longitud del miembro de lado? Si es así, interpréte los, utilizando comparaciones múltiples con un nivel de 5% si es necesario. Si no, explique.
14. El artículo referido en el ejercicio 13 también presenta mediciones de los módulos de Young para los miembros de lado de los ensambles reforzados en T. La tabla siguiente presenta los datos en los que los factores son la longitud del miembro de lado y su sistema de unión. Hubo diez réplicas para cada combinación de los factores. Los datos (en KN/mm^2) son consistentes con las medias y las desviaciones estándar dadas en el artículo.

Sistema de unión	Longitud	Módulo de Young									
Adhesivo	Cuarto	9.56	10.67	8.82	8.40	9.23	8.20	10.23	9.58	7.57	8.05
Adhesivo	Mitad	8.74	9.24	10.77	9.10	8.08	11.14	10.00	9.17	9.79	8.13
Adhesivo	Completo	9.84	9.80	8.31	7.37	10.12	9.18	8.93	8.65	7.89	9.07
Clavo	Cuarto	10.24	9.38	9.38	7.48	9.23	9.64	8.45	8.12	8.86	8.07
Clavo	Mitad	9.84	9.34	9.64	8.21	10.43	9.48	7.46	9.51	10.20	9.66
Clavo	Completo	7.96	8.32	8.73	9.37	9.12	7.98	9.84	8.89	10.10	8.07

- a) Calcule todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Es creíble el modelo aditivo? Proporcione el valor de un estadístico de prueba y el P -valor.
- d) ¿Puede describirse el efecto del sistema de unión (clavo contra adhesivo) sobre los módulos de Young al interpretar los efectos principales del sistema de unión? Si es así, interpréte los. Si no, explique.
- e) ¿Puede describirse el efecto de la longitud del miembro de lado sobre los módulos de Young interpretando los efectos principales de la longitud del miembro de lado? Si es así, interpréte los mediante comparaciones múltiples con un nivel de 5%. Si no, explique.
15. Cada uno de tres operarios repesó dos veces algunas placas de silicio. Los resultados se presentan en la tabla siguiente para tres placas. Todas las placas tenían pesos muy cercanos a 54 g, por lo que los pesos fueron notificados en unidades de μg arriba de 54 gramos (basado en “Revelation of a Microbalance Warmup Effect”, J. Buckner, B. Chin y cols., en *Statistical Case Studies for Industrial Process Improvement*, SIAM-ASA, 1997:39-45).

Placa	Operador 1	Operador 2	Operador 3
1	11 15	10 6	14 10
2	210 208	205 201	208 207
3	111 113	102 105	108 111

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Puede determinarse una tabla ANOVA si hay una diferencia en los pesos medidas entre los operadores; si es así, proporcione el valor del estadístico de prueba y el P -valor. Si no, explique por qué no.

16. En el ejercicio 15 se tiene que las mediciones del operario 2 fueron tomadas por la mañana, poco después de que se había encendido la balanza. Una nueva política fue instituida de dejar todo el tiempo prendida la balanza. Entonces, los tres operarios hicieron dos pesadas de tres placas diferentes. Los resultados se presentan en la tabla siguiente.

Placa	Operador 1	Operador 2	Operador 3
1	152 156	156 155	152 157
2	443 440	442 439	435 439
3	229 227	229 232	225 228

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) Compare la tabla ANOVA del inciso a) con la tabla del inciso a) del ejercicio 15. ¿Recomendaría dejar la balanza prendida todo el tiempo? Explique su razonamiento.
17. El artículo “Cellulose Acetate Microspheres Prepared by O/W Emulsification and Solvent Evaporation Method” (K. Soppinath, A Kulkarni y cols., en *Journal of Microencapsulation*, 2001:811-817) describe un estudio de los efectos de las concentraciones de alcohol polivinílico (PVAL) y diclorometano (DCM) sobre la eficiencia de encapsulación en un proceso que produce microesferas que contienen el medicamento ibuprofén. Había tres concentraciones de PVAL (medido en unidades de % w/v) y tres de DCM (en ml). Los resultados que se presentan en la tabla siguiente son consistentes con las medias y las desviaciones estándar presentados en el artículo.

PVAL	DCM = 50			DCM = 40			DCM = 30		
0.5	98.983	99.268	95.149	96.810	94.572	86.718	75.288	74.949	72.363
1.0	89.827	94.136	96.537	82.352	79.156	80.891	76.625	76.941	72.635
2.0	95.095	95.153	92.353	86.153	91.653	87.994	80.059	79.200	77.141

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) Analice las relaciones entre la concentración de PVAL, la concentración de DCM y la eficiencia de encapsulación.

9.4 Diseños de bloque completamente aleatorios

En algunos experimentos hay factores que varían y tienen un efecto en la respuesta, pero esos efectos son irrelevantes para el experimentador. Por ejemplo, una situación que ocurre comúnmente es que resulta imposible terminar un experimento en un día, por lo que las observaciones se tienen que continuar durante varios días. Si las condiciones que pueden afectar el resultado se desvían día tras día, entonces este concepto representa un factor en el experimento, aunque puede haber un nulo interés en calcular su efecto.

Para un ejemplo más específico, imagine que se evalúan tres tipos de fertilizantes con respecto a su efecto sobre la cosecha de fruta en una huerta de naranjas, y que se realizarán tres replicas con un total de nueve observaciones. Un área se divide en nueve parcelas, en tres hileras de tres parcelas cada una. Suponga que hay una caída de agua a lo largo del área de la parcela, por lo que ahora las hileras reciben cantidades diferentes de agua. La cantidad del agua es ahora un factor en el experimento, aunque no hay interés en calcular el efecto de la cantidad de agua sobre la cosecha de naranja.

Si se ignora el factor de agua, un experimento de un solo factor se podría realizar con el fertilizante como el único factor. Cada uno de los tres fertilizantes sería asignado a tres parcelas. En un experimento completamente aleatorio, los tratamientos serían asignados a las parcelas al azar. La figura 9.11 presenta dos arreglos aleatorios posibles. En el arreglo de la izquierda, las parcelas con el fertilizante A tienen más agua que las de los otros dos fertilizantes. En la parcela de la derecha, las parcelas con el fertilizante C tienen más agua. Cuando los tratamientos para un factor son asignados completamente al azar, es probable que no sea distribuido uniformemente sobre los niveles de otro factor.

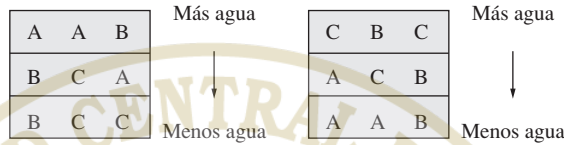


FIGURA 9.11 Dos arreglos posibles para tres fertilizantes, A, B y C, asignados a las nueve parcelas en forma completamente aleatoria. Es probable que las cantidades de agua sean diferentes para los diferentes fertilizantes.

Si la cantidad de agua tiene un efecto insignificante sobre la respuesta, entonces es adecuado el diseño de un solo factor completamente aleatorio. No hay por qué preocuparse por un factor que no afecta la respuesta. Pero ahora suponga que el nivel de agua tiene un impacto importante en la respuesta. Entonces la figura 9.11 muestra que en cualquier otro experimento los efectos estimados de los tratamientos están probablemente fuera de marca, o sesgados, por los niveles diferentes del agua. Los arreglos diferentes de los tratamientos desvían las estimaciones en diferentes direcciones. Si el experimento se repite varias veces, las estimaciones probablemente varíen mucho de repetición en repetición. Por esta razón, el diseño un solo factor completamente aleatorio produce efectos estimados que tienen incertidumbres grandes.

Un mejor diseño para este experimento es uno que contenga dos factores, con el agua como el segundo factor. Debido a que los efectos del agua son irrelevantes, el agua se llama **factor bloqueado**, en vez de un factor de tratamiento. En el experimento de dos factores hay nueve combinaciones de bloque de tratamiento, por lo que corresponde a los tres niveles del tratamiento fertilizante y a los tres niveles de bloque de agua. Con nueve unidades experimentales (las nueve parcelas) es necesario asignar una parcela a cada combinación de fertilizante y agua. La figura 9.12 presenta dos arreglos posibles.

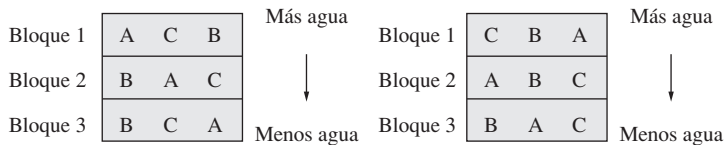


FIGURA 9.12 Dos posibles arreglos para tres fertilizantes, A, B y C, con la restricción que cada fertilizante debe aparecer una vez en cada nivel de agua (bloque). La distribución de niveles de agua es siempre la misma para cada fertilizante.

En el diseño de dos factores cada tratamiento aparece con la misma frecuencia en cada bloque (una vez, en este ejemplo). Por consiguiente, el efecto del factor bloqueado no contribuye a la incertidumbre en la estimación de los efectos principales del factor de tratamiento. Como consecuencia de que cada tratamiento debe salir con la misma frecuencia en cada bloque, la única aleatorización en la asignación de tratamientos para unidades experimentales es el orden en el que salen los tratamientos en cada bloque. Éste no es un diseño completamente aleatorio; es uno en el que los tratamientos están **aleatorizados dentro de los bloques**. Debido a que cada combinación posible de tratamientos y bloques es incluida en el experimento, el diseño está **completo**. Por esta razón el diseño se llama **diseño de bloques completamente aleatorios**.

Estos últimos se pueden construir con varios factores de tratamiento y varios factores de bloques. Se restringirá este análisis al caso donde hay un factor de tratamiento y un factor bloqueado. Los datos de un diseño de bloques completamente aleatorios se analizan con un ANOVA de dos sentidos, del mismo modo que serían los datos de cualquier diseño balanceado de dos factores, completo. Sin embargo, hay una consideración importante. Los únicos efectos de interés son los efectos principales del factor de tratamiento. Para interpretar estos efectos principales, **no debe haber alguna interacción entre el tratamiento y los factores bloqueados**.

Ejemplo

9.19

Se estudia el efecto de tres fertilizantes sobre la cosecha en una huerta de naranjas. Se están utilizando nueve parcelas de tierra, dividida en bloques de tres parcelas cada una. Se usa un diseño de bloques completamente aleatorio, aplicando cada fertilizante una vez en cada bloque. Los resultados, en libras de la fruta cosechada, se presentan en la tabla siguiente, seguida del resultado de MINITAB para el ANOVA de dos sentidos. ¿Se puede concluir que la media de las cosechas difiere entre los fertilizantes? ¿Qué supuesto se hace acerca de las interacciones entre fertilizantes y parcelas? ¿Cómo se calcula la suma de cuadrados del error?

Fertilizante	Parcela 1	Parcela 2	Parcela 3
A	430	542	287
B	367	463	253
C	320	421	207

Two-way ANOVA: Yield versus Block, Fertilizer

Source	DF	SS	MS	F	P
Fertilizer	2	16213.6	8106.778	49.75	0.001
Block	2	77046.9	38523.44	236.4	0.000
Error	4	651.778	162.9444		
Total	8	93912.2			

Solución

El P -valor para el factor fertilizante es 0.001, por lo que se concluye que el fertilizante tiene un efecto sobre la cosecha. Se hace el supuesto de que no hay interacción entre el fertilizan-

te y el factor bloqueado (parcela), por lo que se pueden interpretar los efectos principales de fertilizante. Debido a que hay solamente una observación para cada combinación de tratamiento-bloque (es decir, $K = 1$). La suma de los cuadrados del error (SSE) notificada en el resultado de MINITAB es realmente SSAB, la suma de cuadrados de la interacción, y la media cuadrática del error (MSE) es en realidad MSAB. (Véase el análisis en la p. 676.)

Una mirada más cercana a la tabla ANOVA del ejemplo 9.19 muestra que en este experimento bloquear era necesario para detectar el efecto fertilizante. Con este propósito, considere que el experimento es un experimento de un solo factor. La suma de los cuadrados del error (SSE) sería entonces la suma SSE para el diseño bloqueado más la suma de cuadrados por bloques, o $651.778 + 77\,046.9 = 77\,698.7$. Los grados de libertad del error serían iguales a la suma de los grados de libertad del error en el diseño bloqueado más los grados de libertad por bloques, o $2 + 4 = 6$. El error medio cuadrático (MSE) sería entonces $77\,698.7/6 \approx 12\,950$ en lugar de 162.9444, y el estadístico F para el efecto fertilizante sería entonces al menos de 1, lo que daría como resultado una falla para detectar un efecto.

En general, utilizar un diseño bloqueado reduce los grados de libertad del error, que sólo tiende a reducir la potencia para detectar un efecto. Sin embargo, a menos que el factor bloqueado tenga un muy pequeño efecto sobre la respuesta, esto último será generalmente más que compensado por una reducción en la suma de los cuadrados del error. Dejar de incluir un factor bloqueado que afecta la respuesta puede reducir la potencia enormemente, mientras que incluir un factor bloqueado que no afecta la respuesta reduce la potencia sólo en forma modesta en la mayoría de los casos. Por esta razón es una buena idea utilizar un diseño bloqueado siempre que se considere que es posible que el factor bloqueado está relacionado con la respuesta.

Resumen

- Un diseño de bloques completamente aleatorios de dos factores es un diseño de dos factores balanceado completo en el cual los efectos de un factor (el factor de tratamiento) son relevantes, mientras que los efectos del otro factor (el factor bloqueado) no. El factor bloqueado es incluido para reducir la incertidumbre en las estimaciones del efecto principal del factor de tratamiento.
- Debido a que el objetivo de un diseño de bloques completamente aleatorios es calcular los efectos principales del factor de tratamiento, no debe haber interacción entre el factor de tratamiento y el factor bloqueado.
- Se utiliza un análisis de varianza de dos sentidos para estimar los efectos y realizar pruebas de hipótesis sobre los efectos principales del factor de tratamiento.
- Un diseño de bloques completamente aleatorios proporciona gran ventaja sobre un diseño completamente aleatorio cuando el factor bloqueado afecta fuertemente la respuesta y proporciona una desventaja pequeña cuando el factor bloqueado no tiene poco o nada de efecto. Por tanto, cuando se tiene duda, es una buena idea utilizar un diseño bloqueado.

Ejemplo

9.20

El artículo “Experimental Design for Process Settings in Aircraft Manufacturing” (R. Sauter y R. Lentz, en *Statistical Case Studies: A Collaboration Between Academe and Industry*, SIA-M-ASA, 1998:151-157) describe un experimento en el que se estudió la calidad de huecos perforados en partes de aeronave de metal. Un indicador importante de la calidad de hueco es el “diámetro excedente”, que es la diferencia entre el diámetro poco perforado y el diámetro del hueco. Los diámetros excesivamente pequeños son mejores que los grandes. Suponga que se tiene interés en el efecto de la velocidad rotacional del taladro sobre el diámetro excesivo del hueco. Se perforarán huecos en seis artículos de prueba (cupones), con tres velocidades: 6 000, 10 000 y 15 000 rpm. Al diámetro excedente puede afectarlo no sólo la velocidad del taladro, sino también las propiedades físicas del cupón de prueba. Describa un diseño adecuado para este experimento.

Solución

Un diseño de bloques completamente aleatorios es adecuado, con la velocidad del taladro como el factor de tratamiento y el cupón de prueba como el factor bloqueado. Debido a que se pueden hacer seis observaciones en cada bloque, cada velocidad de taladro se debe usar dos veces en cada bloque. El orden de las velocidades dentro de cada bloque debe elegirse en forma aleatoria.

Ejemplo

9.21

El diseño sugerido en el ejemplo 9.20 ha sido adoptado y se ha realizado el experimento. A continuación se presenta el resultado de MINITAB. ¿El resultado indica alguna violación de los supuestos necesarios? ¿Qué concluye respecto al efecto de la velocidad del taladro sobre el diámetro excedente?

Two-way ANOVA: Excess Diameter versus Block, Speed

Source	DF	SS	MS	F	P
Block	5	0.20156	0.0403117	1.08	0.404
Speed	2	0.07835	0.0391750	1.05	0.370
Interaction	10	0.16272	0.0162717	0.44	0.909
Error	18	0.67105	0.0372806		
Total	35	1.11368			

S = 0.1931

R-Sq = 39.74%

R-Sq(adj) = 0.00%

Solución

En un diseño de bloques completamente aleatorios no debe haber interacción entre el factor de tratamiento y el factor bloqueado, por lo que puede interpretarse el efecto principal del factor de tratamiento. El *P*-valor para las interacciones es 0.909, que es compatible con la hipó-

tesis de no interacciones. Por tanto, no hay señal en el resultado de alguna violación de los supuestos. El P -valor para el efecto principal de la velocidad es 0.370, que no es pequeño. Por tanto, no se puede concluir que al diámetro excedente de hueco lo haya afectado la velocidad de taladro.

El ejemplo 9.22 muestra un diseño apareado (véase la sección 6.8), en el cual se utiliza una prueba t para comparar dos medias poblacionales, es un caso especial de un diseño de bloques aleatorios.

Ejemplo 9.22

Un fabricante de neumáticos quiere comparar el desgaste del dibujo de neumáticos hechos con un material nuevo con el de los hechos con material convencional. Hay diez neumáticos de cada tipo. Cada neumático será montado sobre la rueda frontal de un automóvil de tracción delantera y conducido durante 40 000 millas. Para cada neumático se medirá el desgaste de la huella. Describa un diseño adecuado para este experimento.

Solución

La respuesta es el desgaste de la huella después de 40 000 millas. Hay un factor de interés: el tipo de neumático. Debido a que los automóviles podrían diferir en las cantidades de desgaste que producen, el automóvil también es un factor, pero su efecto carece de interés. Un diseño de bloques completamente aleatorios es adecuado, en el cual un neumático de cada tipo es montado sobre las ruedas de cada automóvil.

Puede observar que el diseño de bloques completamente aleatorios en el ejemplo 9.22 es el mismo que se usa cuando compara dos medias poblacionales con una prueba apareada t , como se analiza en la sección 6.8. El diseño apareado descrito ahí es un caso especial de un diseño de bloques completamente aleatorios, en el cual el factor de tratamiento tiene sólo dos niveles y cada uno sale una vez en cada bloque. En efecto, un análisis de varianza de dos sentidos aplicado a los datos de tal diseño equivale a la prueba t apareada.

Comparaciones múltiples en diseños de bloques completamente aleatorios

Una vez que se ha construido un ANOVA, entonces si la prueba F muestra que el tratamiento de los efectos principales no es el mismo, se puede utilizar un método de comparaciones múltiples para determinar cuáles pares de efectos difieren. Se describe el método de Tukey, que es un caso especial del método de Tukey-Kramer descrito en la sección 9.2. Se usan los grados de libertad y la media cuadrática, que son diferentes si cada tratamiento aparece sólo una vez, o más de una vez, en cada bloque.

Método de Tukey para comparaciones múltiples en diseños de bloques completamente aleatorios

En un diseño de bloques completamente aleatorios, con I niveles de tratamiento, J niveles de bloque, y efectos principales de tratamiento $\alpha_1, \dots, \alpha_I$:

- Si cada tratamiento aparece sólo una vez en cada bloque, entonces la hipótesis nula $H_0 : \alpha_i - \alpha_j = 0$ se rechaza con un nivel de α para cada par de tratamientos i y j para el cual

$$|\bar{X}_{i.} - \bar{X}_{j.}| > q_{I, (I-1)(J-1), \alpha} \sqrt{\frac{\text{MSAB}}{J}}$$

donde MSAB representa la media cuadrática para la interacción.

- Si cada tratamiento aparece $K > 1$ veces en cada bloque, entonces la hipótesis nula $H_0 : \alpha_i - \alpha_j = 0$ se rechaza con un nivel de α para cada par de tratamientos i y j para el cual

$$|\bar{X}_{i..} - \bar{X}_{j..}| > q_{I, IJ(K-1), \alpha} \sqrt{\frac{\text{MSE}}{JK}}$$

donde MSE constituye la media cuadrática del error.

Para mayor información acerca de los diseños de bloques aleatorios se puede consultar un texto acerca del diseño de experimentos, como Montgomery (2001a).

Ejercicios para la sección 9.4

1. El artículo “Methods for Evaluation of Easily-Reducible Iron and Manganese in Paddy Soils” (M. Borges, J. de Mello y cols., en *Communication in Soil Science and Plant Analysis*, 2001:3009-3022) describe un experimento en el cual se midieron los niveles de pH de suelo aluvial. Se aplicaron diferentes niveles de abono a cada suelo. El interés principal se centra en las diferencias entre los suelos, no hay mucho interés en el efecto del abono. Los resultados se presentan en la tabla siguiente.

Tierra	Nivel de abono				
	1	2	3	4	5
A	5.8	5.9	6.1	6.7	7.1
B	5.2	5.7	6.0	6.4	6.8
C	5.5	6.0	6.2	6.7	7.0
D	6.0	6.6	6.7	6.7	7.5

- a) ¿Cuáles son los factores de bloqueo y de tratamiento?
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.

- c) ¿Puede concluir qué suelo tiene niveles diferentes de pH?
- d) ¿Cuáles pares de suelos, si hay, puede concluir que tienen diferentes niveles de pH? Utilice un nivel de 5%.
2. Se hizo un estudio para ver cuál de las cuatro máquinas realiza más rápido cierta tarea. Hay tres operarios; cada uno realiza la tarea dos veces sobre cada máquina. Se emplea un diseño de bloques aleatorios. A continuación se muestra el resultado de MINITAB.

Source	DF	SS	MS	F	P
Machine	(i)	257.678	(ii)	(iii)	0.021
Block	(iv)	592.428	(v)	(vi)	0.000
Interaction	(vii)	(viii)	(ix)	(x)	0.933
Error	(xi)	215.836	17.986		
Total	(xii)	1096.646			

- a) Complete los números que faltan del (i) al (xii) en el resultado.
- b) ¿El resultado indica que se satisfacen los supuestos para el diseño de bloque aleatorios? Explique.
- c) ¿Puede concluir que hay diferencias entre las máquinas? Explique.
3. Se realizó un experimento para determinar el efecto de la concentración del caldo en la producción de cierto microorganismo. Fueron cuatro formulaciones diferentes y tres diferentes concentraciones de caldo. Hubo tres réplicas para cada combinación de la formulación de caldo y de la concentración, cuyo resultado se expresa en ml. La formulación de caldo es un factor bloqueado; el único efecto importante es el efecto de la concentración. Se calcularon las siguientes sumas de los cuadrados: suma de los cuadrados para bloques = 504.7, suma de los cuadrados para tratamientos = 756.7, suma de los cuadrados para las interacciones = 415.3, suma total de los cuadrados = 3 486.1.
- a) Construya una tabla ANOVA. Puede dar rangos para los P -valores.
- b) ¿Los supuestos para un diseño de bloques completamente aleatorios se satisfacen? Explique.
- c) ¿La tabla ANOVA proporciona pruebas de que la concentración de caldo afecta la producción? Explique.
4. Se está considerando emplear tres máquinas diferentes para terminar cierta tarea. Cuatro operarios terminan la tarea sobre cada una de las tres máquinas, cada operario repite la tarea dos veces. El resultado es el tiempo en minutos necesario para terminar la tarea. El único efecto de interés es la máquina; el operario es un factor bloqueado. Se calcularon las siguientes sumas de los cuadrados: suma de los cuadrados para bloques = 48.47, suma de los cuadrados para tratamientos = 30.83, suma de los cuadrados para las interacciones = 93.27, suma total de los cuadrados = 217.36.
- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Los supuestos para diseño de bloques completamente aleatorios se satisfacen? Explique.
- c) ¿La tabla ANOVA proporciona pruebas de que la media del tiempo de terminación difiere entre las máquinas? Explique.
5. El artículo “Genotype-Environment Interactions and Phenotypic Stability Analyses of Linseed in Ethiopia” (W. Adguna y M. Labuschagne, en *Plant Breeding*. 2002:66-71) describe un estudio en que se compararon las producciones de semillas de diez variedades de linaza. Cada variedad fue cultivada en seis parcelas diferentes. Las cosechas, en kilogramos por hectárea, se presentan en la tabla ubicada en la parte superior de la página siguiente.
- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Puede concluir que las variedades tienen diferentes medias de producción?

Variedad	Parcela					
	1	2	3	4	5	6
A	2032	1377	1343	1366	1276	1209
B	1815	1126	1338	1188	1566	1454
C	1739	1311	1340	1250	1473	1617
D	1812	1313	1044	1245	1090	1280
E	1781	1271	1308	1220	1371	1361
F	1703	1089	1256	1385	1079	1318
G	1476	1333	1162	1363	1056	1096
H	1745	1308	1190	1269	1251	1325
I	1679	1216	1326	1271	1506	1368
J	1903	1382	1373	1609	1396	1366

6. El artículo “Sprinkler Technologies, Soil Infiltration, and Runoff” (D. DeBoer y S. Chu, en *Journal of Irrigation and Drainage Engineering*, 2001:234-239) presenta un estudio de la profundidad de corriente (en mm) para varios tipos de aspersor. Se probó cada uno de cuatro aspersores cada cuatro días, con dos réplicas por día (hubo tres réplicas en algunos de días; éstas se omiten). Reviste interés determinar si la profundidad de corriente varía con el tipo de aspersor; la variación de un día a otro parece de interés. Los datos se presentan en la tabla siguiente.

Aspersor	Día 1		Día 2		Día 3		Día 4	
A	8.3	5.5	7.8	4.5	10.7	9.8	10.6	6.6
B	6.5	9.5	3.7	3.6	7.7	10.6	3.6	6.7
C	1.8	1.2	0.5	0.3	1.7	1.9	2.2	2.1
D	0.7	0.8	0.1	0.5	0.1	0.5	0.3	0.5

- a) Identifique el factor bloqueado y el factor de tratamiento.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Los supuestos de un diseño de bloques completamente aleatorios se satisfacen? Explique.
- d) ¿Puede concluir que hay diferencias en la media de la profundidad de corriente entre algunos pares de aspersores? Explique.
- e) ¿Qué pares de aspersores, si hay, puede concluir, con un nivel de 5%, que tienen diferente media de profundidades de corriente?
7. El artículo “Bromate Surveys in French Drinking Waterworks” (B. Legube, B. Parinet y cols., en *Ozone Science and Engineering*, 2002:293-304) presenta mediciones de las concentraciones de bromina (en $\mu\text{g/L}$) en algunas centrales de abastecimiento de agua. Las mediciones hechas 15 veces diferentes en cada una de cuatro centrales de abastecimiento de agua se presentan en la tabla siguiente. (El artículo también presentó algunas mediciones adicionales hechas en algunas otras centrales de abastecimiento de agua.) Es de interés determinar si las concentraciones de bromina varían entre las centrales de abastecimiento de agua; no lo es determinar si las concentraciones varían con el tiempo.

Centrales de abastecimiento de agua	Tiempo														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	29	9	7	35	40	53	38	38	41	34	42	35	38	35	36
B	24	29	21	24	20	25	15	14	8	12	14	35	32	38	33
C	25	17	20	24	19	19	17	23	22	27	17	33	33	39	37
D	31	37	34	30	39	41	34	34	29	33	33	34	16	31	16

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Puede concluir que la concentración de bromina varía entre las centrales de abastecimiento de agua?
- c) ¿Qué pares de las centrales de abastecimiento de agua, si hay, puede concluir, con un nivel de 5%, que tienen concentraciones de bromina diferentes?
- d) Alguien sugiere que estos datos se analizaran con un ANOVA de un sentido, ignorando el factor de tiempo, con 15 observaciones para cada una de las cuatro centrales de abastecimiento de agua. ¿La tabla ANOVA soporta esta sugerencia? Explique.
8. El artículo “Application of Fluorescence Technique for Rapid Identification of IOM Fractions in Source Waters” (T. Marhaba y R. Lippincott, en *Journal of Environmental Engineering*, 2000:1039-1044) presenta mediciones de las concentraciones de carbono orgánico disuelto (en mg/L) en seis ubicaciones (A, B, C, D, E, F) a lo largo del río Millstone, en el centro de Nueva Jersey. Las mediciones se tomaron cuatro veces al año: enero, abril, julio y octubre. Es de interés determinar si las concentraciones varían entre las ubicaciones. Los datos se presentan en la tabla siguiente.

	Enero	Abril	Julio	Octubre
A	3.9	3.7	3.7	4.1
B	4.0	3.5	3.4	5.7
C	4.2	3.4	3.0	4.8
D	4.1	3.3	2.9	4.6
E	4.1	3.4	3.0	3.4
F	4.2	3.5	2.8	4.7

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Puede concluir que la concentración varía entre las ubicaciones?
- c) ¿Qué pares de ubicaciones, si hay, puede concluir, con un nivel de 5%, que tienen concentraciones diferentes?
9. Le han dado la tarea de diseñar un estudio con respecto a los tiempos de vida de cinco tipos diferentes de motor eléctrico. La pregunta inicial que debe responderse es si hay diferencias en la media de vida entre los cinco tipos. Hay 20 motores, cuatro de cada tipo, disponibles para la prueba. Se puede probar un máximo de cinco motores todos los días. La temperatura ambiental es diferente día tras día, esto último puede afectar la vida del motor.
- a) Describa cómo elegiría los cinco motores para hacer pruebas todos los días. ¿Usaría un diseño completamente aleatorio? ¿Usaría cualquier aleatorización en todo?
- b) Si X_{ij} representa la vida media medida de un motor del tipo i probada un día j , exprese el estadístico de prueba para probar la hipótesis nula de tiempos de vida iguales en función del X_{ij} .
10. Una catedrática de ingeniería quiere determinar qué tema encuentran sus alumnos más difícil entre estadística, física y química. Obtiene la calificación del examen final para cuatro estudiantes que tomaron tres cursos el último semestre y quienes estaban en las mismas secciones de cada clase. Los resultados se presentan en la tabla siguiente.

Curso	Estudiante			
	1	2	3	4
Estadística	82	94	78	70
Física	75	70	81	83
Química	93	82	80	70

- a) La catedrática propone un diseño de bloques completamente aleatorios, con los estudiantes como los bloques. Dé una razón de que es probable que no sea adecuado.
- b) Describa las características de los datos en la tabla anterior que indican que no se cumplen los supuestos del diseño de bloques completamente aleatorios.

9.5 Experimentos factoriales 2^p

Cuando un experimentador quiere estudiar varios factores simultáneamente, el número de diferentes tratamientos puede ser muy grande. En estos casos, con frecuencia los experimentos preliminares se realizan en cada factor que tiene sólo dos niveles. Se diseña un nivel como el nivel “alto”, y se diseña otro como el nivel “bajo”. Si hay p factores, entonces hay 2^p tratamientos diferentes. Tales experimentos se llaman **experimentos factoriales 2^p** . Con frecuencia, el propósito de un experimento de este tipo es determinar qué factores tienen un efecto importante sobre el resultado. Una vez que se han determinado, se pueden realizar experimentos más elaborados, en los cuales los factores que son importantes varían en algunos niveles. Se iniciará por describir los experimentos factoriales 2^3 .

Notación para los experimentos factoriales 2^3

En experimentos factoriales 2^3 hay tres factores y $2^3 = 8$ tratamientos. El **efecto principal** de un factor se define como la diferencia entre la respuesta media cuando el factor está en su nivel alto y la respuesta media cuando el factor está en su nivel bajo. Es común denotar los efectos principales por A , B y C como con cualquier experimento factorial, puede haber interacciones entre los factores. Con tres factores hay tres interacciones de dos sentidos, una para cada par de factores, y una interacción de tres sentidos. Las interacciones de dos sentidos por AB , AC y BC , y la interacción de tres sentidos se denota con ABC . Los tratamientos se denotan tradicionalmente con letras minúsculas, con una letra que indica que un factor está en su nivel alto. Por ejemplo, ab denota el tratamiento en el que los primeros dos factores están en su nivel alto y el tercer factor en su nivel bajo. El símbolo “1” se utiliza para denotar el tratamiento en el que todos los factores están a sus niveles bajos.

Estimación de los efectos en un experimento factorial 2^3

Suponga que hay K réplicas para cada tratamiento en un experimento factorial 2^3 . Para cada tratamiento, la media de la celda representa el promedio de las K observaciones para ese tratamiento. Las fórmulas para las estimaciones del efecto se pueden obtener fácilmente de la **tabla de signos 2^3** , que se presenta como la tabla 9.7 en la página 694. Los signos se colocan en la tabla como se indica a continuación. Para los efectos principales A , B , C el signo es $+$ para los tratamientos en los que el factor está en su nivel alto, y $-$ para los tratamientos donde el factor está en su nivel bajo. En este contexto, para el efecto principal A , el signo es $+$ para los tratamientos a , ab , ac y abc y $-$ para el resto. Para las interacciones, los signos se calculan tomando el producto de los signos en las columnas de los efectos principales correspondientes. Por ejemplo, los signos para la interacción de dos sentidos AB son los productos de los signos en las columnas A y B , y los signos para la interacción de tres sentidos ABC son los productos de los signos en las columnas A y B y C .

La estimación de los efectos principales y de las interacciones se hace empleando la tabla de los signos. Se muestra cómo estimar el efecto principal del factor A . El factor A está en su nivel alto en los renglones de la tabla donde hay un signo “ $+$ ” en la columna A . Cada una de las medias de la celda $\bar{X}_a$, $\bar{X}_{ab}$, $\bar{X}_{ac}$ y $\bar{X}_{abc}$ es una respuesta promedio para las corridas hechas con el factor A como en su nivel alto. Se estima que la respuesta media para el factor A en su nivel alto es el promedio de estas medias de celda.

TABLA 9.7 Tabla de signos para un experimento factorial 2^3

Tratamiento	Media de la celda	A	B	C	AB	AC	BC	ABC
1	$\bar{X}_1$	—	—	—	+	+	+	—
a	$\bar{X}_a$	+	—	—	—	—	+	+
b	$\bar{X}_b$	—	+	—	—	+	—	+
ab	$\bar{X}_{ab}$	+	+	—	+	—	—	—
c	$\bar{X}_c$	—	—	+	+	—	—	+
ac	$\bar{X}_{ac}$	+	—	+	—	+	—	—
bc	$\bar{X}_{bc}$	—	+	+	—	—	+	—
abc	$\bar{X}_{abc}$	+	+	+	+	+	+	+

Respuesta media estimada para A en su nivel alto = $\frac{1}{4}(\bar{X}_a + \bar{X}_{ab} + \bar{X}_{ac} + \bar{X}_{abc})$

De manera similar, cada renglón con signo “—” en la columna A representa un tratamiento con un factor A puesto en su nivel bajo. Se estima la respuesta media para el factor A en su nivel bajo como el promedio de las medias de celda en estos renglones.

Respuesta media estimada para A en su nivel bajo = $\frac{1}{4}(\bar{X}_1 + \bar{X}_b + \bar{X}_c + \bar{X}_{bc})$

La estimación del efecto principal del factor A es la diferencia en la respuesta media estimada entre sus niveles alto y bajo.

Efecto estimado de A = $\frac{1}{4}(-\bar{X}_1 + \bar{X}_a - \bar{X}_b + \bar{X}_{ab} - \bar{X}_c + \bar{X}_{ac} - \bar{X}_{bc} + \bar{X}_{abc})$

La cantidad dentro de los paréntesis se llama **contraste** del factor A. Se calcula al sumar y restar las medias de celda, utilizando los signos en la columna adecuada de la tabla de signos. Observe que el número de signos “más” es el mismo como el número de signos “menos”, por lo que la suma de los coeficientes es igual a 0. La estimación del efecto se obtiene al dividir el contraste entre la mitad del número de los tratamientos, que es $2^3/2$, o 4. Las estimaciones de otros efectos principales y las interacciones se calculan en una forma similar. Para ilustrar lo anterior se presentan las estimaciones del efecto para el efecto principal C y la interacción de dos sentidos AB:

Estimación de efecto C = $\frac{1}{4}(-\bar{X}_1 - \bar{X}_a - \bar{X}_b - \bar{X}_{ab} + \bar{X}_c + \bar{X}_{ac} + \bar{X}_{bc} + \bar{X}_{abc})$

Estimación de la interacción AB = $\frac{1}{4}(\bar{X}_1 - \bar{X}_a - \bar{X}_b + \bar{X}_{ab} + \bar{X}_c - \bar{X}_{ac} - \bar{X}_{bc} + \bar{X}_{abc})$

Resumen

El **contraste** para cualquier efecto principal o interacción se obtiene mediante la suma y resta de las medias de la celda, así como utilizando los signos en la columna adecuada de la tabla de signos.

Para un experimento factorial 2³,

$$\text{Efecto estimado} = \frac{\text{contraste}}{4} \quad (9.65)$$

Ejemplo 9.23

Se realizó un experimento factorial 2³ para calcular los efectos de tres factores en la producción de una reacción química. Los factores fueron A: concentración de catalizador (bajo o alto), B: reactivo (formulación usual o nueva formulación), y C: tasa de agitación (lenta o rápida). Se obtuvieron tres réplicas para cada tratamiento. Las producciones se presentan en la tabla siguiente y se miden como un porcentaje de un máximo teórico. Estime todos los efectos y las interacciones.

Tratamiento	Producción			Media de la celda
1	71.67	70.55	67.40	69.8733
a	78.46	75.42	81.77	78.5500
b	77.14	78.25	78.33	77.9067
ab	79.72	76.17	78.41	78.1000
c	72.65	71.03	73.54	72.4067
ac	80.10	73.91	74.81	76.2733
bc	80.20	73.49	74.86	76.1833
abc	75.58	80.28	71.64	75.8333

Solución

Mediante la tabla de signos (tabla 9.7) se determinan las sumas adecuadas y las diferencias de las medias de celda. Se presentan los cálculos para el efecto principal A, la interacción de dos sentidos BC, y la interacción de tres sentidos ABC:

$$\begin{aligned} \text{Estimación de efecto A} &= \frac{1}{4}(-69.8733 + 78.5500 - 77.9067 + 78.1000 \\ &\quad - 72.4067 + 76.2733 - 76.1833 + 75.8333) = 3.0967 \end{aligned}$$

$$\begin{aligned} \text{Interacción estimada BC} &= \frac{1}{4}(69.8733 + 78.5500 - 77.9067 - 78.1000 \\ &\quad - 72.4067 - 76.2733 + 76.1833 + 75.8333) = -1.0617 \end{aligned}$$

$$\begin{aligned} \text{Interacción estimada ABC} &= \frac{1}{4}(-69.8733 + 78.5500 + 77.9067 - 78.1000 \\ &\quad + 72.4067 - 76.2733 - 76.1833 + 75.8333) = 1.0667 \end{aligned}$$

En la tabla siguiente se presentan todos los efectos estimados, redondeados con la misma precisión que los datos:

Variable	Efecto
<i>A</i>	3.10
<i>B</i>	2.73
<i>C</i>	−0.93
<i>AB</i>	−3.18
<i>AC</i>	−1.34
<i>BC</i>	−1.06
<i>ABC</i>	1.07

Para cada efecto se puede probar la hipótesis nula de que el efecto es igual a 0. Cuando ésta es rechazada se tiene evidencia de que los factores implicados afectan en realidad el resultado. Para probar estas hipótesis nulas se construye una tabla ANOVA que contiene las sumas de los cuadrados adecuadas. Las pruebas que se presentan serán válidas si el número de réplicas es para cada tratamiento y deber ser al menos 2. Además, las observaciones en cada tratamiento deben constituir una muestra aleatoria de una población normal, y todas las poblaciones deben tener la misma varianza.

Se calcula la suma de los cuadrados del error (SSE) al sumar las sumas de las desviaciones al cuadrado de las medias muestrales para todos los tratamientos. Con el fin de expresar esto en una ecuación, sean $s_1^2, \dots, s_8^2$ las variancias muestrales de las observaciones en cada uno de los ocho tratamientos, y K el número de réplicas por tratamiento. Entonces

$$SSE = (K - 1) \sum_{i=1}^8 s_i^2 \quad (9.66)$$

Cada efecto principal e interacción también tienen su propia suma de los cuadrados. Éstas son fáciles calcular. La suma de los cuadrados para cualquier efecto o interacción se calcula elevando al cuadrado su contraste, multiplicando por el número de réplicas K , y dividiendo entre el número total de tratamientos: $2^3 = 8$.

$$\text{Suma de los cuadrados para un efecto} = \frac{K(\text{contraste})^2}{8} \quad (9.67)$$

Cuando se utiliza la ecuación (9.67) es mejor conservar, en tanto sea posible, los mismos dígitos que en las estimaciones de los efectos con el fin de obtener la máxima precisión en la suma de los cuadrados. Para la presentación en una tabla, las estimaciones de los efectos y las sumas de los cuadrados se pueden redondear con la misma precisión que los datos.

Las sumas de los cuadrados para los efectos y las interacciones tienen un grado de libertad cada una. La suma de los cuadrados del error tiene $8(K - 1)$ grados de libertad. El método para calcular la media cuadrática y el estadístico F es el mismo que el que se presentó en la sección 9.3 para una tabla ANOVA de dos sentidos. Cada media cuadrática es igual a su

suma de los cuadrados dividida entre sus grados de libertad. El estadístico de prueba para probar la hipótesis nula de que un efecto o interacción es igual a 0 se calcula dividiendo la media cuadrática para la estimación del efecto entre la media cuadrática para el error. Cuando la hipótesis nula es verdadera, el estadístico de prueba tiene una distribución $F_{1, 8(K-1)}$.

Ejemplo

9.24

Con referencia al ejemplo 9.23, construya una tabla ANOVA. Respecto a cada efecto e interacción pruebe la hipótesis nula de que es igual a 0. ¿Cuáles factores, si hay, parecen más probables de tener un efecto en el resultado?

Solución

La tabla ANOVA se muestra a continuación. Las sumas de los cuadrados para los efectos y las interacciones se calcularon utilizando la ecuación (9.67). La suma de los cuadrados del error fue calculada aplicando la ecuación (9.66) a los datos del ejemplo 9.23. Cada estadístico F es el cociente de la media cuadrática con la media cuadrática del error. Cada estadístico F tiene 1 y 16 grados de libertad.

Fuente	Efecto	Suma de cuadrados	gl	Media cuadrática	F	P
A	3.10	57.54	1	57.54	7.34	0.015
B	2.73	44.72	1	44.72	5.70	0.030
C	-0.93	5.23	1	5.23	0.67	0.426
AB	-3.18	60.48	1	60.48	7.71	0.013
AC	-1.34	10.75	1	10.75	1.37	0.259
BC	-1.06	6.76	1	6.76	0.86	0.367
ABC	1.07	6.83	1	6.83	0.87	0.365
Error		125.48	16	7.84		
Total		317.78	23			

Los efectos principales de los factores A y B, así como la interacción de AB, tienen P -valores bastante pequeños. Ello indica que estos efectos no son iguales a 0 y que los factores A y B afectan el resultado. No hay evidencias de que el efecto principal o factor C, o alguna de sus interacciones, difieran de 0. Experimentos adicionales se podrían concentrar en los factores A y B. Quizá se realizaría un ANOVA de dos sentidos, con cada uno de los factores A y B evaluados en diferentes niveles, para obtener información más detallada acerca de sus efectos en el resultado.

Interpretación del resultado computacional

En la práctica, los análisis de diseños factoriales se realizan generalmente en una computadora, utilizando un software como MINITAB. El siguiente resultado de MINITAB presenta los resultados del análisis descrito en los ejemplos 9.23 y 9.24.

Factorial Fit: Yield versus A, B, C

Estimated Effects and Coefficients for Yield (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		75.641	0.5716	132.33	0.000
A	3.097	1.548	0.5716	2.71	0.015
B	2.730	1.365	0.5716	2.39	0.030
C	-0.933	-0.467	0.5716	-0.82	0.426
A*B	-3.175	-1.587	0.5716	-2.78	0.013
A*C	-1.338	-0.669	0.5716	-1.17	0.259
B*C	-1.062	-0.531	0.5716	-0.93	0.367
A*B*C	1.067	0.533	0.5716	0.93	0.365

S = 2.80040 R-Sq = 60.51% R-Sq(adj) = 43.24%

Analysis of Variance for Yield (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	107.480	107.480	35.827	4.57	0.017
2-Way Interactions	3	77.993	77.993	25.998	3.32	0.047
3-Way Interactions	1	6.827	6.827	6.827	0.87	0.365
Residual Error	16	125.476	125.476	7.842		
Pure Error	16	125.476	125.476	7.842		
Total	23	317.776				

La tabla en la parte superior del resultado de MINITAB presenta las estimaciones de los efectos y coeficientes. La frase “coded units” significa que se usan los valores 1 y -1 en lugar de los valores verdaderos y con ellos se representan los niveles altos y bajos de cada factor. Los efectos estimados se listan en la columna “Effect”. En la columna siguiente están las estimaciones de los **coeficientes**, cada una de las cuales es igual a la mitad del efecto correspondiente. Mientras que el efecto representa la diferencia en la respuesta media entre los niveles altos y bajos de un factor, el coeficiente representa la diferencia entre la respuesta media al nivel alto y la respuesta de la gran media, que es cuando mucho de la mitad. El coeficiente “Constant” es la media de todas las observaciones; es decir, es la gran media muestral. Cada estimación del coeficiente tiene la misma desviación estándar, ello se muestra en la columna “SECoef”.

MINITAB usa el estadístico t de Student, en lugar del estadístico F , para probar las hipótesis de que los efectos son iguales cero. La columna “T” presenta el valor del estadístico t de Student, que es igual al cociente de la estimación del coeficiente (Coef) y su desviación estándar. Bajo la hipótesis nula, el estadístico t tiene la distribución t de Student con $2^p(K - 1)$ grados de libertad. Los P -valores se presentan en la columna “P”. La prueba t que realiza MI-

NITAB equivale a la prueba F descrita en el ejemplo 9.24. El estadístico $t_{8(K-1)}$ se puede calcular tomando la raíz cuadrada del estadístico $F_{1, 8(K-1)}$ y aplicando el signo de la estimación del efecto. Los P -valores son idénticos.

A continuación se tratará el análisis de la tabla de varianzas. La columna “DF” presenta los grados de libertad. Las columnas “SeqSS” (suma de los cuadrados secuencial) y “AdjSS” (suma de los cuadrados ajustada) serán idénticas en todos los ejemplos que se considerarán y que contendrán sumas de los cuadrados. La columna “Adj MS” contiene medias cuadráticas, o sumas de cuadrados divididas entre sus grados de libertad. Ahora se explicarán los renglones que implican al error. El renglón “PureError” se refiere a la suma de los cuadrados del error (SSE) (ecuación 9.66). Hay $8(K-1) = 16$ grados de libertad (DF) para el error puro. La suma de los cuadrados para este último se encuentra en cada una de las siguientes dos columnas, es la suma de los cuadrados del error (SSE). Bajo la columna “Adj MS” está la media cuadrática para el error. El renglón arriba del renglón del error puro se denomina “Residual Error”. La suma de los cuadrados para el error de residuos es igual a la suma de cuadrados para el error puro, más las sumas de los cuadrados para cualquiera de los efectos principales o interacciones que no se incluyen en el modelo. Los grados de libertad por la suma de los cuadrados del error de residuos es igual a los grados de libertad para el error puro, más los grados de libertad (uno de cada uno) para cada efecto principal o interacción no incluidos en el modelo. Debido a que en este ejemplo todos los efectos principales e interacciones están incluidos en el modelo, la suma de los cuadrados del error de residuos y sus grados de libertad son iguales a las cantidades correspondientes para el error puro. El renglón “Total” contiene la suma de los cuadrados (SST). La suma total de los cuadrados y sus grados de libertad son iguales a las sumas de las cantidades correspondientes para todos los efectos, las interacciones y el error de residuos.

Yendo a la parte superior de la tabla, el primer renglón es “Main Effects”. Hay tres grados de libertad para los efectos principales, porque hay tres efectos principales (A , B , y C) con un grado de libertad cada uno. La suma de los cuadrados secuencial es la suma de las sumas de los cuadrados para cada uno de los tres efectos principales. La media cuadrática (Adj MS) es la suma de los cuadrados dividida entre sus grados de libertad. La columna “F” presenta el estadístico F para probar la hipótesis nula de que todos los efectos principales son iguales al cero. El valor del estadístico F (4.57) es igual al cociente de la media cuadrática para los efectos principales (35.827) y la media cuadrática para el error (puro) (7.842). Los grados de libertad para el estadístico F son 3 y 16, por lo que corresponden a los grados de libertad para las dos medias cuadráticas. La columna “P” presenta el P -valor para la prueba F . En este caso el P -valor es 0.017, que indica que no todos los efectos principales son cero.

Los renglones “2-Way Interactions” y “3-Way Interactions” son similares al renglón para los efectos principales. El P -valor para las interacciones de dos sentidos es 0.047, que es una evidencia razonablemente fuerte de que al menos algo de las interacciones de dos sentidos no es igual a cero. Debido a que sólo hay una interacción de tres sentidos ($A * B * C$), el P -valor en el renglón “3-Way Interactions” es igual (0.365) como el P -valor en la tabla en la parte superior del resultado de MINITAB para $A * B * C$.

Recuerde que las pruebas de hipótesis se realizan bajo el supuesto de que todas las observaciones tienen la misma desviación estándar σ . La cantidad “S” es la estimación de σ y es igual a la raíz cuadrada de la media al cuadrado del error (MSE). Las cantidades “R-sq” y “R-sq(adj)” son los coeficientes de determinación R^2 y del R^2 ajustado, respectivamente, y se calculan con métodos similares a los del ANOVA de un sentido.

Estimación de efectos en un experimento factorial 2^p

Mediante una tabla de signos se pueden obtener las fórmulas para calcular las estimaciones del efecto en cualquier experimento factorial 2^p. El método es similar al caso 2³, los tratamientos están listados en una columna. El signo para cualquier efecto principal es + en los renglones que corresponden a los tratamientos donde el factor está en su nivel alto, y – en renglones que corresponden a los tratamientos donde el factor está en su nivel bajo. Los signos para las interacciones se encuentran al multiplicar los signos que corresponden a los factores en la interacción. La estimación para cualquier efecto o interacción se encuentra sumando y restando las medias de celda para los tratamientos, utilizando los signos en las columnas adecuadas, para calcular un contraste. El contraste es dividido entre la mitad del número de tratamientos, o 2^{p-1}, y obtener así la estimación del efecto.

Resumen	
Para un experimento factorial 2 ^p :	
$\text{Efecto estimado} = \frac{\text{contraste}}{2^{p-1}} \tag{9.68}$	

Como ejemplo, la tabla 9.8 presenta una tabla de signos para un experimento factorial 2⁵. Se enlistan los signos para los efectos principales y las interacciones seleccionadas.

Las sumas de cuadrados se calculan con un método similar al del experimento 2³. Para calcular la suma de los cuadrados de error (SSE), sean s₁, . . . , s_{2^p} las variancias de muestra de las observaciones en cada uno de los 2^p tratamientos. Entonces

$$SSE = (K - 1) \sum_{i=1}^{2^p} s_i^2$$

Los grados de libertad para el error son 2^p(K – 1), dónde K es el número de réplicas por tratamiento. La suma de los cuadrados para cada efecto e interacción es igual al cuadrado del contraste, multiplicado por el número de réplicas K y dividido entre el número de tratamientos 2^p. Las sumas de los cuadrados para los efectos y las interacciones tienen un grado de libertad cada una.

$$\text{Suma de los cuadrados para un efecto} = \frac{K(\text{contraste})^2}{2^p} \tag{9.69}$$

Los estadísticos F para los efectos principales y las interacciones se calculan dividiendo la suma de los cuadrados para el efecto entre la media cuadrática para el error. La distribución nula del estadístico F es F_{1, 2^p(K – 1)}.

Experimentos factoriales sin réplica

Cuando el número de p factores es grande con frecuencia no es viable realizar más de una réplica por cada tratamiento. En este caso no es posible calcular SSE, por lo que las pruebas de hipótesis antes descritas no se pueden efectuar. Si es razonable suponer que algunas de las interacciones de más alto-orden sean iguales a 0, entonces las sumas de los cuadrados para esas interacciones pueden sumarse y tratarse de la misma manera que una suma de los cuadrados

del error. Entonces los efectos principales y las interacciones de orden más bajo se pueden probar.

TABLA 9.8 Tabla de signos para los efectos principales e interacciones seleccionadas para un experimento factorial 2⁵

Tratamiento	A	B	C	D	E	AB	CDE	ABDE	ABCDE
1	—	—	—	—	—	+	—	+	—
a	+	—	—	—	—	—	—	—	+
b	—	+	—	—	—	—	—	—	+
ab	+	+	—	—	—	+	—	+	—
c	—	—	+	—	—	+	+	+	+
ac	+	—	+	—	—	—	+	—	—
bc	—	+	+	—	—	—	+	—	—
abc	+	+	+	—	—	+	+	+	+
d	—	—	—	+	—	+	+	—	+
ad	+	—	—	+	—	—	+	+	—
bd	—	+	—	+	—	—	+	+	—
abd	+	+	—	+	—	+	+	—	+
cd	—	—	+	+	—	+	—	—	—
acd	+	—	+	+	—	—	—	+	+
bcd	—	+	+	+	—	—	—	+	+
abcd	+	+	+	+	—	+	—	—	—
e	—	—	—	—	+	+	+	—	+
ae	+	—	—	—	+	—	+	+	—
be	—	+	—	—	+	—	+	+	—
abe	+	+	—	—	+	+	—	—	+
ce	—	—	+	—	+	—	—	—	—
ace	+	—	+	—	+	—	—	+	+
bce	—	+	+	—	+	—	—	+	+
abce	+	+	+	—	+	+	—	—	—
de	—	—	—	+	—	+	—	+	—
ade	+	—	—	+	—	—	—	—	+
bde	—	+	—	+	—	—	—	—	+
abde	+	+	—	+	—	+	—	+	—
cde	—	—	+	+	+	+	+	+	+
acde	+	—	+	+	+	—	+	—	—
bcde	—	+	+	+	+	—	+	—	—
abcde	+	+	+	+	+	+	+	+	+

Ejemplo

9.25

Mediante un experimento factorial 2⁵ se estimaron los efectos de cinco factores en la calidad de focos fabricados por cierto proceso. Los factores eran A: planta (1 o 2), B: tipo de máquina (velocidad baja o alta), C: turno (día o noche), D: material del alambre de plomo (usado o nuevo) y E: método de cargar materiales en el ensamblador (manual o automático). Se obtuvo una réplica para cada tratamiento. La tabla 9.9 presenta los resultados. Calcule las estimaciones de los efectos principales y las interacciones, y sus sumas de los cuadrados. Suponga que las interacciones de tercero, cuarto y quinto órdenes son despreciables, y haga las sumas de los cuadrados para utilizarlas como sustituto de una suma de los cuadrados del error. Utilice este sustituto para probar hipótesis referidas a los efectos principales y a las interacciones de segundo orden.

TABLA 9.9

Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado
1	32.07	d	35.64	e	25.10	de	40.60
a	39.27	ad	35.91	ae	39.25	ade	37.57
b	34.81	bd	47.75	be	37.77	bde	47.22
ab	43.07	abd	51.47	abe	46.69	abde	56.87
c	31.55	cd	33.16	ce	32.55	cde	34.51
ac	36.51	acd	35.32	ace	32.56	acde	36.67
bc	28.80	bcd	48.26	bce	28.99	bcde	45.15
abc	43.05	abcd	53.28	abce	48.92	abcde	48.72

TABLA 9.10

Variable	Efecto	Suma de cuadrados	Variable	Efecto	Suma de cuadrados
A	6.33	320.05	ABD	-0.29	0.67
B	9.54	727.52	ABE	0.76	4.59
C	-2.07	34.16	ACD	0.11	0.088
D	6.70	358.72	ACE	-0.69	3.75
E	0.58	2.66	ADE	-0.45	1.60
AB	2.84	64.52	BCD	0.76	4.67
AC	0.18	0.27	BCE	-0.82	5.43
AD	-3.39	91.67	BDE	-2.17	37.63
AE	0.60	2.83	CDE	-1.25	12.48
BC	-0.49	1.95	ABCD	-2.83	63.96
BD	4.13	136.54	ABCE	0.39	1.22
BE	0.65	3.42	ABDE	0.22	0.37
CD	-0.18	0.26	ACDE	0.18	0.24
CE	-0.81	5.23	BCDE	-0.25	0.52
DE	0.24	0.46	ABCDE	-1.73	23.80
ABC	1.35	14.47			

Solución

Los efectos se calculan utilizando las reglas y al sumar y restar las observaciones dadas en la tabla de signos y las sumas de los cuadrados mediante la ecuación (9.69). Véase la tabla 9.10.

Observe que ninguna de las interacciones de tercer, cuarto o quinto orden están entre los efectos mayores. Si alguna de ellas estuviera no sería inteligente combinar sus sumas de los cuadrados. En consecuencia, se suman las sumas de los cuadrados de las interacciones tercera, cuarta y quinta. Los resultados se presentan en el siguiente resultado de MINITAB.

Factorial Fit: Response versus A, B, C, D, E

Estimated Effects and Coefficients for
Response (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		39.658	0.5854	67.74	0.000
A	6.325	3.163	0.5854	5.40	0.000

B	9.536	4.768	0.5854	8.14	0.000
C	-2.066	-1.033	0.5854	-1.76	0.097
D	6.696	3.348	0.5854	5.72	0.000
E	0.576	0.288	0.5854	0.49	0.629
A*B	2.840	1.420	0.5854	2.43	0.027
A*C	0.183	0.091	0.5854	0.16	0.878
A*D	-3.385	-1.693	0.5854	-2.89	0.011
A*E	0.595	0.298	0.5854	0.51	0.618
B*C	-0.494	-0.247	0.5854	-0.42	0.679
B*D	4.131	2.066	0.5854	3.53	0.003
B*E	0.654	0.327	0.5854	0.56	0.584
C*D	-0.179	-0.089	0.5854	-0.15	0.881
C*E	-0.809	-0.404	0.5854	-0.69	0.500
D*E	0.239	0.119	0.5854	0.20	0.841

S = 3.31179 R-Sq = 90.89% R-Sq(adj) = 82.34%

Analysis of Variance for Response (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	1443.1	1443.1	288.62	26.31	0.000
2-Way Interactions	10	307.1	307.1	30.71	2.80	0.032
Residual Error	16	175.5	175.5	10.97		
Total	31	1925.7				

Las estimaciones no han cambiado para los efectos principales o las interacciones de dos sentidos. La suma de los cuadrados del error de residuos (175.5) en la tabla de análisis de varianza se encuentra al sumar los cuadrados para todas las interacciones de orden más alto que se eliminaron del modelo. El número de grados de libertad (16) es igual a la suma de los grados de libertad (uno para cada uno) para las 16 interacciones de orden más alto. No hay ninguna suma de los cuadrados del error puro (SSE), porque sólo hay una réplica por tratamiento. La suma de los cuadrados del error de residuos se utiliza como un sustituto de SSE para calcular todas las cantidades que requieren una suma de los cuadrados del error.

Se concluye del resultado que los factores A, B, y D probablemente afecten el resultado. También parece que lo hacen las interacciones entre algunos pares de estos factores. Podría ser adecuado planear experimentos adicionales concentrados en los factores A, B, y D.

Uso de gráficas de probabilidad para detectar efectos grandes

Un método informal, sugerido para ayudar a determinar cuáles de los efectos es grande, consiste en trazar el efecto y la interacción estimados en una gráfica de probabilidad normal. Si, en efecto, ninguno de los factores afecta el resultado, entonces el efecto y la interacción estimados forman una muestra aleatoria simple de una población normal y se deben encontrar aproximadamente en una línea recta. En muchos casos, la mayoría de las estimaciones se encontrarán aproximadamente en una recta, y unas pocas estarán lejos de la recta. Los efectos principales y las interacciones cuyas estimaciones se encuentran lejos de la recta tienen más probabilidad de ser importantes. La figura 9.13 presenta una gráfica de probabilidad normal

del efecto principal y la interacción estimada a partir de los datos del ejemplo 9.25. Es claro, a partir de la gráfica, que los efectos principales de los factores A , B y D , y las interacciones AB y BD , resaltan del resto.

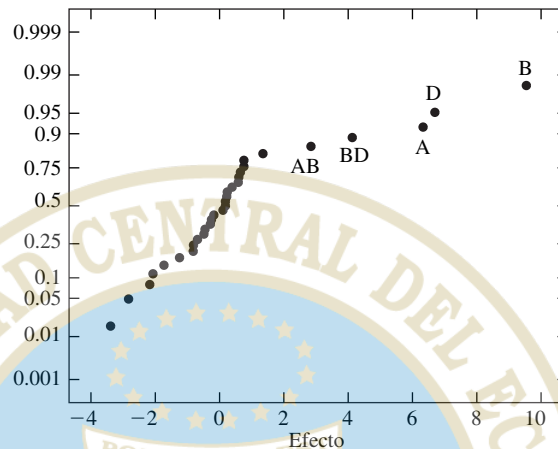


FIGURA 9.13 Gráfica de probabilidad normal de las estimaciones del efecto de los datos del ejemplo 9.25. Los efectos principales de los factores A , B y D son mayores que el resto.

Experimentos factoriales muy pequeños

Cuando el número de los factores es grande tal vez no sea factible realizar incluso una réplica para cada tratamiento. En estos casos las observaciones pueden tomarse sólo para alguna fracción de los tratamientos. Si éstos se eligen correctamente, aun es posible obtener la información acerca de los factores.

Cuando cada factor tiene dos niveles, la fracción debe ser siempre una potencia de 2; es decir, la mitad, un cuarto, etcétera. Un experimento en el que se usa la mitad de los tratamientos se llama **media-réplica**; si se usa un cuarto de los tratamientos, es un **cuarto-réplica**, etcétera. Un experimento 2^p de media-réplica con frecuencia se denota por 2^{p-1} , para indicar que mientras que hay p factores, solamente hay 2^{p-1} tratamientos que deben considerarse. De forma similar, un cuarto-réplica con frecuencia se denota mediante 2^{p-2} . Se concentrará la atención en experimentos de media-réplica.

Se presenta un método para elegir un media-réplica de un experimento 2^5 . Tal experimento tendrá 16 tratamientos elegidos de los 32 en el experimento 2^5 . La elección de 16 tratamientos se inicia con una tabla de signos para un diseño 2^4 que indica los signos para los efectos principales y la interacción de más alto orden. Esto último se presenta en la tabla 9.11.

La tabla 9.11 tiene el número correcto de tratamientos (16), pero sólo cuatro factores. Con el fin de transformarla para un diseño 2^5 de media-réplica, se debe presentar un quinto factor, E . Para hacer esto se reemplaza la interacción de más alto orden con E . Éste establece los signos para el efecto principal de E . Entonces en cada renglón donde el signo para E es $+$, se agrega la letra e al tratamiento. Esto indica que el factor E se establece en su nivel más

TABLA 9.11 Tabla de signos para los efectos principales y para las interacciones de cuatro sentidos en un experimento factorial 2⁴

Tratamiento	A	B	C	D	ABCD
1	—	—	—	—	+
a	+	—	—	—	—
b	—	+	—	—	—
ab	+	+	—	—	+
c	—	—	+	—	—
ac	+	—	+	—	+
bc	—	+	+	—	+
abc	+	+	+	—	—
d	—	—	—	+	—
ad	+	—	—	+	+
bd	—	+	—	+	+
abd	+	+	—	+	—
cd	—	—	+	+	+
acd	+	—	+	+	—
bcd	—	+	+	+	—
abcd	+	+	+	+	+

alto para ese tratamiento. Donde el signo de E es —, el factor E se establece en su nivel bajo. El diseño resultante se llama **fracción principal** del diseño 2⁵. La tabla 9.12 presenta los signos para los efectos principales y las interacciones seleccionadas de este diseño.

TABLA 9.12 Tabla de signos para los efectos principales y las interacciones seleccionadas para un experimento factorial 2⁵

Tratamiento	A	B	C	D	E = ABCD	AB	CDE	ACDE
e	—	—	—	—	+	+	+	—
a	+	—	—	—	—	—	—	—
b	—	+	—	—	—	—	—	+
abe	+	+	—	—	+	+	+	+
c	—	—	+	—	—	+	+	—
ace	+	—	+	—	+	—	—	—
bce	—	+	+	—	+	—	—	+
abc	+	+	+	—	—	+	+	+
d	—	—	—	+	—	+	+	—
ade	+	—	—	+	+	—	—	—
bde	—	+	—	+	+	—	—	+
abd	+	+	—	+	—	+	+	+
cde	—	—	+	+	+	+	+	—
acd	+	—	+	+	—	—	—	—
bcd	—	+	+	+	—	—	—	+
abcde	+	+	+	+	+	+	+	+

Hay un precio que pagar por utilizar solamente la mitad de los tratamientos. Con este propósito, observe que en la tabla 9.12 la interacción AB tiene los mismos signos que la interacción CDE y la interacción $ACDE$ tiene los mismos signos que el efecto principal B . Cuando dos efectos tienen los mismos signos, se dice que son **seudónimos**. En efecto, los efectos principales y las interacciones en una fracción media forman pares en los cuales cada miembro de la pareja es seudónimo del otro. Los pares de seudónimos para esta fracción media de diseño 2^5 son

$$\begin{array}{ccccc} \{A, BCDE\} & \{B, ACDE\} & \{C, ABDE\} & \{D, ABCE\} & \{E, ABCD\} \\ \{AB, CDE\} & \{AC, BDE\} & \{AD, BCE\} & \{AE, BCD\} & \{BC, ADE\} \\ \{BD, ACE\} & \{BE, ACD\} & \{CD, ABE\} & \{CE, ABD\} & \{DE, ABC\} \end{array}$$

Cuando dos efectos son seudónimos sus estimaciones del efecto son las mismas, ya que implican los mismos signos. En efecto, cuando se utiliza la fracción principal de un diseño, la estimación de cualquier efecto representa la suma de ese efecto y de su seudónimo. Por tanto, para la fracción principal de un diseño 2^5 , cada estimación del efecto principal representa realmente la suma del efecto principal más su interacción seudónima de cuatro sentidos, y cada estimación de la interacción de dos sentidos representa la suma de la interacción de dos sentidos más su interacción seudónima de tres sentidos.

En muchos casos es razonable suponer que las interacciones de más alto orden son pequeñas. En la media-réplica 2^5 , si las interacciones de cuatro sentidos son despreciables, la estimación del efecto principal será exacta. Si además las interacciones de tres sentidos son despreciables, la estimación de la interacción de dos sentidos también es exacta.

En un diseño fraccional sin replicación con frecuencia no hay ninguna buena manera de calcular una suma de los cuadrados del error; por tanto, ninguna forma rigurosa de probar las hipótesis de que los efectos sean iguales a 0. En muchos casos, el propósito de un diseño fraccional sólo es identificar algunos factores que parecen tener mayor impacto en el resultado. Esta información se puede utilizar para diseñar experimentos más elaborados con el fin de investigar estos factores. Con este propósito puede ser suficiente sólo elegir aquellos factores cuyos efectos o interacciones de dos sentidos son inusualmente grandes, sin realizar pruebas de hipótesis. Esto último es posible enlistando las estimaciones en orden decreciente, y luego ver si hay algunos que sean notablemente mayores que el resto. Otro método es trazar las estimaciones del efecto y de la interacción en una gráfica de probabilidad normal, como se analizó previamente.

Ejemplo

9.26

En un sistema de emulsión de membrana líquida, una emulsión (fase interna) se dispersa en un medio externo líquido que contiene un contaminante. Éste se elimina del líquido externo mediante transferencia de masa en la emulsión. El escape de la fase interna ocurre cuando partes del material extraído se derraman en el líquido externo. En el artículo “Leakage and Swell in Emulsion Liquid Membrane Systems: Batch Experiments” (R. Pfeiffer, W. Navidi y A. Bunge, en *Separation Science and Technology*, 2003:519-539) se estudiaron los efectos de cinco factores para determinar el efecto sobre la fuga en cierto sistema. Los cinco factores fueron A : concentración del agente tensoactivo, B : concentración de hidróxido de litio de la

fase interna, *C*: fase de membrana, *D*: fracción de volumen de fase interna y *E*: tasa de agitación de extracción del recipiente. Se utilizó un diseño 2⁵ de media fracción. Los datos se presentan en la tabla siguiente (en el experimento verdadero, cada punto en realidad representaba el promedio de dos mediciones). La fuga se mide en unidades porcentuales. Suponga que las interacciones de tercero, cuarto y quinto órdenes son despreciables. Estime los efectos principales y las interacciones de dos sentidos. ¿Cuáles, si hay, destacan por ser notablemente mayores que el resto?

Tratamiento	Fuga	Tratamiento	Fuga	Tratamiento	Fuga	Tratamiento	Fuga
<i>e</i>	0.61	<i>c</i>	0.35	<i>d</i>	2.03	<i>cde</i>	1.45
<i>a</i>	0.13	<i>ace</i>	0.075	<i>ade</i>	0.64	<i>acd</i>	0.31
<i>b</i>	2.23	<i>bce</i>	7.31	<i>bde</i>	11.72	<i>bcd</i>	1.33
<i>abe</i>	0.095	<i>abc</i>	0.080	<i>abd</i>	0.56	<i>abcde</i>	6.24

Solución

Utilizando la tabla de signos (tabla 9.12) se calculan las estimaciones para los efectos principales y las interacciones de dos sentidos, los cuales se muestran en la tabla siguiente.

Variable	Efecto	Variable	Efecto
<i>A</i>	-2.36	<i>AE</i>	-1.15
<i>B</i>	3.00	<i>BC</i>	0.20
<i>C</i>	-0.11	<i>BD</i>	0.86
<i>D</i>	1.68	<i>BE</i>	2.65
<i>E</i>	2.64	<i>CD</i>	-1.30
<i>AB</i>	-1.54	<i>CE</i>	0.61
<i>AC</i>	1.43	<i>DE</i>	1.32
<i>AD</i>	0.17		

Observe que no se considera realizar cálculos para las sumas de los cuadrados para las estimaciones, porque no se tiene SSE con el cual compararlos. Para determinar informalmente qué efectos pueden merecer investigación adicional, se clasifican las estimaciones en orden de sus valores absolutos; *B*: 3.00, *BE*: 2.65, *E*: 2.64, *A*: -2.36, *D*: 1.68, y así sucesivamente. Parece razonable decidir que hay bastante separación entre los efectos *A* y *D*; por tanto, se tiene que los factores *A*, *B* y *E* son probablemente los más importantes.

Ejercicios para la sección 9.5

1. Construya una tabla de signos para la fracción principal respecto de un diseño 2⁴. Después indique todos los pares de pseudónimos.
2. Dé un ejemplo de un experimento factorial en el cual la falla de aleatoriedad puede producir los resultados incorrectos.
3. Un estudio en una planta piloto investigó los efectos de tres factores en la producción (en gramos) de monómero para la formulación adhesiva. Los tres factores fueron *A*: temperatura, *B*: nivel catalizador y *C*: tipo de catalizador. Los niveles de los factores fueron *A*: 160°C(-1) y 180°C(+1); *B*: 20%(-1) y 40%(+1); *C*: distribuidor E(-1) y distribuidor J(+1). Se realizaron dos réplicas de un diseño factorial 2³, con los resultados que se presentan en la tabla siguiente.

A	B	C	Producciones	Media de la producción
-1	-1	-1	58, 60	59
1	-1	-1	74, 70	72
-1	1	-1	43, 51	47
1	1	-1	75, 73	74
-1	-1	1	56, 60	58
1	-1	1	74, 78	76
-1	1	1	46, 44	45
1	1	1	78, 80	79

- a) Calcule las estimaciones de los efectos principales y las interacciones, y la suma de los cuadrados y el P -valor para cada uno.
- b) ¿Cuáles efectos principales e interacciones son importantes?
- c) Las otras cosas siguen igual, ¿la media de la producción será más alta cuando la temperatura sea de 160°C o de 180°C? Explique.
4. El artículo “Efficient Pyruvate Production by a Multi-Vitamin Auxotroph of *Torulopsis glabrata*: Key Role and Optimization of Vitamin Levels” (Y. Li, J. Chen y cols., en *Applied Microbiology and Biotechnology*, 2001:680-685) investiga los efectos de los niveles de algunas vitaminas en un cultivo de células en la producción (en g/L) de piruvato, un ácido orgánico útil. Los datos en la tabla siguiente se presentan con dos réplicas de un diseño 2^3 . Los factores son A: el ácido nicotínico, B: tiamina y C: biotina. (Se han eliminado dos factores estadísticamente insignificantes. En el artículo, cada factor fue probado en cuatro niveles; se han reducido éstos a dos.)

A	B	C	Producciones	Media de la producción
-1	-1	-1	0.55, 0.49	0.520
1	-1	-1	0.60, 0.42	0.510
-1	1	-1	0.37, 0.28	0.325
1	1	-1	0.30, 0.28	0.290
-1	-1	1	0.54, 0.54	0.540
1	-1	1	0.54, 0.47	0.505
-1	1	1	0.44, 0.33	0.385
1	1	1	0.36, 0.20	0.280

- a) Calcule las estimaciones de los efectos principales y las interacciones, y la suma de los cuadrados y el P -valor para cada una.
- b) ¿Es adecuado el modelo aditivo?
- c) ¿Qué conclusiones acerca de los factores se pueden obtener de estos resultados?
5. El artículo citado en el ejercicio 4 también investigó los efectos de los factores en el consumo de glucosa (en g/L). Se proporciona una sola medición para cada combinación de factores (en el artículo hubo cierta replicación). Los resultados se presentan en la tabla siguiente.

A	B	C	Consumo de glucosa
-1	-1	-1	68.0
1	-1	-1	77.5
-1	1	-1	98.0
1	1	-1	98.0
-1	-1	1	74.0
1	-1	1	77.0
-1	1	1	97.0
1	1	1	98.0

- a) Calcule las estimaciones de los efectos principales y las interacciones.
- b) ¿Se puede calcular una suma de los cuadrados del error? Explique.
- c) ¿Alguna de las interacciones se encuentra entre los efectos mayores? Si es así, ¿cuál es?
- d) Suponga que se conoce de experiencias pasadas donde el modelo aditivo vale. Sume las sumas de los cuadrados para las interacciones, y utilice ese resultado en lugar de una suma de los cuadrados del error para probar las hipótesis de que los efectos principales sean iguales a 0.
6. Se diseñó y se construyó un decantador para la separación de sólidos de jabón de una suspensión con base de agua. El propósito del dispositivo es reducir el contenido de agua de la suspensión de jabón para que éste se pueda procesar aún más en barras y el líquido (una solución salina) se pueda reutilizar. Se realizó un estudio preliminar en el que se variaron tres factores. Éstos fueron A: temperatura de la operación (medida en °C), B: tasa de circulación del dispositivo (medida como el tiempo de residencia, en segundos) y C: atasco de sólidos en el dispositivo (medido como el tiempo de residencia, en segundos). El resultado es la fracción sólida en la suspensión. El experimento se realizó una vez para cada combinación de factores. Los resultados se presentan en la tabla siguiente.

Temperatura (A)	Circulación (B)	Atasco de sólidos (C)	Fracción de sólidos en la suspensión
50	170	150	33
70	170	150	44
50	300	150	47
70	300	150	36
50	170	500	51
70	170	500	43
50	300	500	54
70	300	500	56

- a) Calcule las estimaciones de los efectos principales y de las interacciones.
- b) ¿Se puede calcular una suma de los cuadrados del error? Explique.
- c) ¿Alguna de las interacciones se encuentra entre los efectos mayores? Si es así, ¿cuál es?
- d) Alguien afirma que el modelo aditivo vale. ¿Los resultados tienden a respaldar esta declaración? Explique.
7. La medición de la resistencia de una placa cubierta con tungsteno que se utiliza en la fabricación de circuitos integrados puede estar afectada por varios factores, incluyendo A: cuál de los dos tipos de casete se utiliza para fijar la placa, B: si la placa está cargada hacia arriba o sobre la parte inferior del casete, y C: si se utilizó la parte del frente o trasera de la estación del casete. Los datos de un experimento factorial 2³ con una réplica se presentan en la tabla siguiente. (Basado en el artículo “Prometrix RS35e Gauge Study in Five Two-Level Factors and One Three-Level Factor”, J. Buckner, B. Chin y J. Henri, en *Statistical Case Studies for Industrial Process Improvement*, SIAM-ASA, 1997:9-18.)

A	B	C	Resistencia (mΩ)
-1	-1	-1	85.04
1	-1	-1	84.49
-1	1	-1	82.15
1	1	-1	86.37
-1	-1	1	82.60
1	-1	1	85.14
-1	1	1	82.87
1	1	1	86.44

- a) Calcule las estimaciones de los efectos principales y de las interacciones.
- b) ¿Se puede calcular una suma de los cuadrados del error? Explique.
- c) Trace las estimaciones en una gráfica de probabilidad normal. ¿La gráfica muestra que algunos de los factores influyen en la resistencia? Si es así, ¿cuál es?
8. En un diseño 2^p con una réplica por tratamiento a veces ocurre que falta la observación para uno de los tratamientos, debido al error experimental o a alguna otra causa. Cuando esto último ocurre, un enfoque es reemplazar el valor faltante con el que tiene la interacción de más alto orden igual a 0. Considere el ejercicio 7 y suponga que falta la observación para el tratamiento donde A , B y C están en su nivel bajo (-1).
- a) ¿Qué valor para esta observación hace que la interacción de tres sentidos sea igual a 0?
- b) Utilizando este valor, calcule las estimaciones para los efectos principales y las interacciones.
9. Las consideraciones de seguridad son importantes en el diseño de automóviles. El artículo “An Optimum Design Methodology Development Using a Statistical Technique for Vehicle Occupant Safety” (J. Hong, M. Mun, y S. Song, en *Proceedings of the Institution of Mechanical Engineers*, 2001:795-801) presenta los resultados de un estudio de simulación de un ocupante. La variable del resultado es la aceleración del pecho (en g) 3 ms después del impacto. Se consideraron cuatro factores. Éstos fueron A : característica de abertura de la bolsa de aire, B : tiempo en que se dispara el inflador de la bolsa de aire, C : tasa de flujo de masa del inflador de la bolsa de aire y D : relación de tensión-tirantez de espuma de rodilla. Los resultados (parte de un estudio mayor) se presentan en la tabla siguiente. Hay una réplica por tratamiento.

Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado
1	85.2	c	66.0	d	85.0	cd	62.6
a	79.2	ac	69.0	ad	82.0	acd	65.4
b	84.3	bc	68.5	bd	84.7	bcd	66.3
ab	89.0	abc	76.4	abd	82.2	$abcd$	69.0

- a) Calcule las estimaciones de los efectos principales y de las interacciones.
- b) Si usted fuera a diseñar un estudio de seguimiento, ¿en qué factor o factores se concentraría? Explique.
10. El artículo estudio “Experimental Study of Workpiece-Level Variability in Blind via Electroplating” (G. Poon, J. Chan y D. Williams, en *Proceedings of the Institution of Mechanical Engineers*, 2001:521-530) describe un experimento factorial con que determinó cuáles de los diferentes factores influyen en la variabilidad del espesor de un depósito de ácido de cobre electrodepositado. Los resultados del experimento se presentan aquí como un diseño completo 2^5 ; se han omitido dos factores estadísticamente insignificantes. Hubo una réplica por tratamiento. Los factores son A : concentración de ácido sulfúrico (g/L), B : concentración de sulfato de cobre (g/L), C : densidad de corriente promedio (mA/cm²), D : separación de electrodo (cm) y E : distancia entre los huecos (cm). El resultado es la variabilidad en el espesor (en μm), medida al calcular la desviación estándar de los espesores medidos en diferentes puntos en un tablero de prueba, después de que se han eliminado las mediciones más grandes y más pequeñas. Los datos se presentan en la tabla siguiente.

Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado	Tratamiento	Resultado
1	1.129	d	1.760	e	1.224	de	1.674
a	0.985	ad	1.684	ae	1.092	ade	1.215
b	1.347	bd	1.957	be	1.280	bde	1.275
ab	1.151	abd	1.656	abe	1.381	$abde$	1.446
c	2.197	cd	2.472	ce	1.859	cde	2.585
ac	1.838	acd	2.147	ace	1.865	$acde$	2.587
bc	1.744	bcd	2.142	bce	1.867	$bcde$	2.339
abc	2.101	$abcd$	2.423	$abce$	2.005	$abcde$	2.629

- a) Calcule las estimaciones de los efectos principales y de las interacciones.
- b) Si usted fuera a diseñar un experimento de seguimiento, ¿en qué factores se concentraría? ¿Por qué?

11. El artículo “Factorial Design for Column Flotation of Phosphate Wastes” (N. Abdel-Khalek, en *Particulate Science and Technology*, 2000:57-70) describe un diseño factorial 2^3 para investigar el efecto de velocidad aérea superficial (*A*), concentración más espumosa (*B*) y velocidad de agua de lavado superficial (*C*) en la recuperación porcentual del P_2O_5 . Hubo dos réplicas. Los datos se presentan en la tabla siguiente.

<i>A</i>	<i>B</i>	<i>C</i>	Recuperación porcentual	
-1	-1	-1	56.30	54.85
1	-1	-1	70.10	72.70
-1	1	-1	65.60	63.60
1	1	-1	80.20	78.80
-1	-1	1	50.30	48.95
1	-1	1	65.30	66.00
-1	1	1	60.53	59.50
1	1	1	70.63	69.86

- a) Calcule las estimaciones de los efectos principales y las interacciones, junto con sus sumas de los cuadrados y *P*-valores.
- b) ¿Cuales factores parecen más importantes? ¿Los factores importantes interactúan? Explique.

12. El artículo “An Application of Fractional Factorial Designs” (M. Kilgo, en *Quality Engineering*, 1988:19-23) describe un diseño 2^{5-1} (media-réplica de un diseño 2^5) que implica el uso de dióxido de carbono (CO_2) en la alta presión para extraer aceite de los cacahuates. Los resultados fueron la solubilidad del aceite de cacahuete en el CO_2 (en mg de aceite/litro de CO_2) y la producción de aceite de cacahuete (en %). Los cinco factores fueron *A*: presión del CO_2 , *B*: temperatura del CO_2 , *C*: humedad del cacahuete, *D*: tasa de circulación del CO_2 y *E*: tamaño de partícula de cacahuete. Los resultados se presentan en la tabla siguiente.

Tratamiento	Solubilidad	Producción	Tratamiento	Solubilidad	Producción
<i>e</i>	29.2	63	<i>d</i>	22.4	23
<i>a</i>	23.0	21	<i>ade</i>	37.2	74
<i>b</i>	37.0	36	<i>bde</i>	31.3	80
<i>abe</i>	139.7	99	<i>abd</i>	48.6	33
<i>c</i>	23.3	24	<i>cde</i>	22.9	63
<i>ace</i>	38.3	66	<i>acd</i>	36.2	21
<i>bce</i>	42.6	71	<i>bcd</i>	33.6	44
<i>abc</i>	141.4	54	<i>abcde</i>	172.6	96

- a) Suponiendo que las interacciones de tercer orden y superiores son despreciables, calcule las estimaciones de los efectos principales y de las interacciones para la solubilidad resultante.
- b) Trace las estimaciones en una gráfica de probabilidad normal. ¿La gráfica muestra que algunos de los factores influyen en la solubilidad? Si es así, ¿cuáles son?
- c) Suponiendo que las interacciones de tercer orden y superior sean insignificantes, calcule las estimaciones de los efectos principales y de las interacciones para la producción resultante.
- d) Trace las estimaciones en una gráfica de probabilidad normal. La gráfica muestra que algunos de los factores influyen en la producción? Si es así, ¿cuáles son?

13. En un diseño 2^{5-1} (como en el ejercicio 12), ¿qué representa en realidad la estimación del efecto principal del factor *A*?

- i) El efecto principal de A .
- ii) La suma del efecto principal de A y la interacción de $BCDE$.
- iii) La diferencia entre el efecto principal de A y la interacción de $BCDE$.
- iv) La interacción entre A y $BCDE$.

Ejercicios adicionales para el capítulo 9

- El artículo “Gypsum Effect on the Aggregate Size and Geometry of Three Sodic Soils Under Reclamation” (I. Lebron, D. Suárez y T. Yoshida, en *Journal of the Soil Science Society of America*, 2002:92-98) describe un experimento en el que se ha agregado yeso en diferentes cantidades a muestras de tierra antes del filtrado. Uno de los resultados relevantes fue el pH de la tierra. El yeso se agregó en cuatro cantidades diferentes. Tres muestras de tierra recibieron una cantidad agregada. Las mediciones de pH de las muestras se presentan en la tabla siguiente.

Yeso (g/kg)	pH		
0.00	7.88	7.72	7.68
0.11	7.81	7.64	7.85
0.19	7.84	7.63	7.87
0.38	7.80	7.73	8.00

¿Puede concluir que el pH difiere de la cantidad de yeso agregado? Proporcione el valor del estadístico de prueba y el P -valor.

- El artículo mencionado en el ejercicio 1 también considera el efecto del yeso en la conductancia eléctrica (en dS m^{-1}) de la tierra. Dos tipos de tierra fueron tratados, cada uno, con tres cantidades diferentes de yeso, con dos réplicas para cada combinación de tierra-yeso. Los datos se presentan en la tabla siguiente.

Yeso (g/kg)	Tipo de tierra			
	las Animas		Madera	
0.00	1.52	1.05	1.01	0.92
0.27	1.49	0.91	1.12	0.92
0.46	0.99	0.92	0.88	0.92

- ¿Hay evidencias convincentes de una interacción entre la cantidad de yeso y el tipo de tierra?
 - ¿Puede concluir que la conductividad difiere entre los tipos de tierra?
 - ¿Puede concluir que la conductividad difiere con la cantidad de yeso agregado?
- La penicilina la produce el hongo *Penicillium*, que se cultiva en un caldo cuyo contenido de azúcar debe ser cuidadosamente controlado. Se tomaron algunas muestras de caldo en cada uno de tres días sucesivos, y se midió la cantidad de azúcar disuelta (en mg/ml) en cada muestra. Los resultados fueron:

Día 1:	4.8	5.1	5.1	4.8	5.2	4.9	5.0	4.9	5.0	4.8	4.8	5.1	5.0
Día 2:	5.4	5.0	5.0	5.1	5.2	5.1	5.3	5.2	5.2	5.1	5.4	5.2	5.4
Día 3:	5.7	5.1	5.3	5.5	5.3	5.5	5.1	5.6	5.3	5.2	5.5	5.3	5.4

¿Puede concluir que la media de la concentración de azúcar difiere entre los tres días?

4. El siguiente resultado de MINITAB es un ANOVA de dos sentidos. Algo no funcionó con la impresora, y algunos de los números no se imprimieron.

Two-way Analysis of Variance

Analysis of Variance

Source	DF	SS	MS	F	P
Row	3	145.375	(d)	(g)	(j)
Column	2	15.042	(e)	(h)	(k)
Interaction	6	(b)	4.2000	(i)	(l)
Error	(a)	(c)	(f)		
Total	23	217.870			

Complete los números que faltan en la tabla del (a) al (l). Puede dar los rangos para los *P*-valores.

5. Mediante un experimento se determinó si a diferentes clases de chocolate les toma periodos diferentes disolverse. Cuarenta personas fueron divididas en cinco grupos. A cada grupo se le asignó cierto tipo de chocolate. Cada persona disolvió una pieza de chocolate, y se registró el tiempo de disolución (en segundos). Para comparar, cada persona en cada grupo también disolvió una pieza de golosina de dulce de azúcar y mantequilla; estos artículos eran idénticos para todos los grupos. Los datos, que incluyen el grupo, el tiempo de disolución tanto para el chocolate como para el dulce de azúcar y mantequilla, la diferencia entre los tiempos de disolución, y el cociente de los tiempos de disolución, se presentan en la tabla siguiente. Observe que el diseño es ligeramente no balanceado; el grupo 3 tiene nueve personas y el grupo 5 tiene solamente siete.

Grupo	Chocolate	Dulce de azúcar y mantequilla	Diferencia	Cociente
1	135	60	75	2.25
1	865	635	230	1.36
1	122	63	59	1.94
1	110	75	35	1.47
1	71	37	34	1.92
1	81	58	23	1.40
1	2 405	1 105	1 300	2.18
1	242	135	107	1.79
2	42	38	4	1.11
2	30	30	0	1.00
2	104	110	-6	0.95
2	124	118	6	1.05
2	75	40	35	1.88
2	80	91	-11	0.88
2	255	121	134	2.11
2	71	71	0	1.00
3	51	53	-2	0.96
3	47	40	7	1.18
3	90	155	-65	0.58
3	65	90	-25	0.72
3	27	33	-6	0.82

Continúa en la página 714

Grupo	Chocolate	Dulce de azúcar y mantequilla	Diferencia	Cociente
3	105	68	37	1.54
3	90	72	18	1.25
3	54	52	2	1.04
3	93	77	16	1.21
4	48	30	18	1.60
4	85	55	30	1.55
4	82	50	32	1.64
4	46	22	24	2.09
4	64	46	18	1.39
4	125	45	80	2.78
4	69	30	39	2.30
4	73	44	29	1.66
5	105	45	60	2.33
5	99	58	41	1.71
5	45	23	22	1.96
5	137	64	73	2.14
5	170	105	65	1.62
5	153	93	60	1.65
5	49	28	21	1.75

- a) Para probar si hay diferencias en la media de los tiempos de disolución para los diferentes tipos de chocolate, alguien sugiere realizar un ANOVA de un sentido, utilizando el tiempo de disolución para los datos del chocolate. ¿Estos datos parecen satisfacer los supuestos para un ANOVA de un sentido? Explique.
- b) Otra persona sugiere utilizar las diferencias (chocolate — dulce de azúcar y mantequilla). ¿Estos datos parecen satisfacer los supuestos para un ANOVA de un sentido? Explique.
- c) Realice un análisis de varianza de un sentido utilizando los cocientes. ¿Puede concluir que la media del cociente del tiempo de disolución difiere para clases de chocolate diferentes?
6. El artículo “Stability of Silico-Ferrite of Calcium and Aluminum (SFCA) in Air-Solid Solution Limits Between 1 240°C and 1 390°C and Phase Relationships within the Fe_2O_3 -CaO- Al_2O_3 -SiO₂ (FCAS) System” (T. Patrick y M. Pownceby, en *Metallurgical and Materials Transactions B*, 2002:79-90) investiga las propiedades de silicio-ferritas de calcio y aluminio (SFCA). Los datos en la tabla siguiente presentan el cociente de los pesos Fe_2O_3 y CaO para elementos SFCA con algunos porcentajes de peso diferentes de Al_2O_3 y C_4S_3 .

Al_2O_3 (%)	C_4S_3	$\text{Fe}_2\text{O}_3/\text{CaO}$			
1.0	Bajo (3-6%)	7.25	6.92	6.60	6.31
1.0	Medio (7-10%)	6.03	5.78	5.54	5.31
1.0	Alto (11-14%)	5.10	4.90	4.71	4.53
5.0	Bajo (3-6%)	6.92	6.59	6.29	6.01
5.0	Medio (7-10%)	5.74	5.26	5.04	4.84
5.0	Alto (11-14%)	4.84	4.65	4.47	4.29
10.0	Bajo (3-6%)	6.50	6.18	5.89	5.63
10.0	Medio (7-10%)	5.37	5.14	4.92	4.71
10.0	Alto (11-14%)	4.52	4.33	4.16	3.99

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) ¿Los datos indican que hay interacciones entre el porcentaje de peso de Al_2O_3 y el porcentaje de peso C_4S_3 ? Explique.

- d) ¿Los datos demuestran convincentemente que el cociente $\text{Fe}_2\text{O}_3/\text{CaO}$ depende del porcentaje de peso de Al_2O_3 ? Explique.
- e) ¿Los datos demuestran convincentemente que el cociente $\text{Fe}_2\text{O}_3/\text{CaO}$ depende del porcentaje de peso de C_4S_3 ? Explique.
7. Se puede fabricar un componente de acuerdo con cualquiera de dos diseños, uno con material más costoso o uno con material menos costoso. Varios componentes se fabrican con cada combinación de diseño y material, y los tiempos de vida de cada uno se miden (en horas). Mediante un análisis de varianza de dos sentidos se estimaron los efectos de diseño y material en el tiempo de vida de un componente. Las estimaciones de la media de las celdas y de los efectos principales se presentan en la tabla siguiente.

Medias de la celda			Efectos principales	
	Diseño 1	Diseño 2		
Más caros	118	120	Más caros	14
Menos caros	60	122	Menos caros	-14
			Diseño 1	-16
			Diseño 2	16

Tabla ANOVA

Fuente	GL	SS	MS	F	P
Material	1	2 352.0	2 352.0	10.45	0.012
Diseño	1	3 072.0	3 072.0	13.65	0.006
Interacción	1	2 700.0	2 700.0	12.00	0.009
Error	8	1 800.0	225.00		
Total	11	9 924.0			

El ingeniero de proceso recomienda que se debe utilizar el diseño 2 junto con el material más costoso. Argumenta que los efectos principales son positivos, por lo que la utilización de esta combinación dará como resultado un tiempo de vida mayor del componente. ¿Está de acuerdo con la recomendación? ¿Por qué sí o no?

8. El artículo “Case Study Based Instruction of DOE and SPC” (J. Brady y T. Allen, en *The American Statistician*, 2002:312-315) presenta el resultado de un experimento factorial 2^{4-1} para investigar los efectos de cuatro factores en la producción de un proceso que fabrica tableros de circuitos impresos. Los factores fueron *A*: la alimentación eléctrica de transistor produjo (límite de especificación superior o más bajo), *B*: enfoque de soporte de transistor (atornillado o soldado), *C*: tipo de absorbedor de calor del transistor (configuración actual o alternativa) y *D*: posición de tornillo en el ajustador de frecuencia (media o dos vueltas). Los resultados se presentan en la tabla siguiente. La producción es un porcentaje de un máximo teórico.

A	B	C	D	Producción
-1	-1	-1	-1	79.8
1	-1	-1	1	69.0
-1	1	-1	1	72.3
1	1	-1	-1	71.2
-1	-1	1	1	91.3
1	-1	1	-1	95.4
-1	1	1	-1	92.7
1	1	1	1	91.5

- a) Estime los efectos principales de cada uno de los cuatro factores.
- b) Suponiendo que todas las interacciones sean despreciables, combine las sumas de los cuadrados para la interacción que se utilizarán en vez de una suma de los cuadrados del error.
- c) ¿Cuál de los cuatro factores, si hay, puede concluir que afecta la producción? ¿Cuál es el P -valor de la prueba importante?
9. El artículo “Combined Analysis of Real-Time Kinematic GPS Equipment and Its Users for Height Determination” (W. Featherstone y M. Stewart, en *Journal of Surveying Engineering*, 2001:31-51) presenta un estudio de la precisión de los equipos para medir la altura del sistema de posicionamiento global (GPS, por sus siglas en inglés). Se estudiaron tres tipos de equipos y se utilizó cada uno en la realización de mediciones en cuatro estaciones base diferentes (en el artículo se incluye una quinta estación, para que los resultados difirieran considerablemente de las otras cuatro). Se hicieron 60 mediciones con cada pieza del equipo en cada base. Las medias y las desviaciones estándar de los errores de medición (en mm) se presentan en la tabla siguiente para cada combinación de tipo de equipo y estación base.

	Instrumento A		Instrumento B		Instrumento C	
	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar
Base 0	3	15	-24	18	-6	18
Base 1	14	26	-13	13	-2	16
Base 2	1	26	-22	39	4	29
Base 3	8	34	-17	26	15	18

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) La pregunta relevante es si la media del error es diferente entre los instrumentos. No es de interés determinar si el error es diferente entre las estaciones base. Por esta razón, un topógrafo sugiere tratar esto último como un diseño de bloques completamente aleatorios, con las estaciones base como los bloques. ¿Esto es adecuado? Explique.
10. Productores de azúcar de arce de Vermont patrocinaron un programa de prueba para determinar el beneficio de un potencial nuevo régimen fertilizante. Se eligió una muestra aleatoria de 27 árboles de arce en Vermont y se trató con uno de los tres niveles de fertilizante sugeridos por el químico del productor. En esta configuración experimental, nueve árboles (tres de cada una de las tres zonas climáticas) se trataron con cada nivel de fertilizante y se midió la cantidad de savia producida (en ml) por los árboles en la siguiente estación. Los resultados se presentan en la tabla siguiente.

	Zona sur			Zona central			Zona norte		
Fertilizante bajo	76.2	80.4	74.2	79.4	87.9	86.9	84.5	85.2	80.1
Fertilizante medio	87.0	95.1	93.0	98.2	94.7	96.2	88.4	90.4	92.2
Fertilizante alto	84.2	87.5	83.1	90.3	89.9	93.2	81.4	84.7	82.2

- a) Calcule los efectos principales de los niveles de fertilizante y la zona climática, y sus interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- c) Pruebe la hipótesis de que no hay interacción entre los niveles de fertilizante y la zona climática.
- d) Pruebe la hipótesis de que no hay ninguna diferencia en la producción de savia respecto los tres niveles de fertilizante.
11. Un ingeniero civil está interesado en algunos diseños para un canal de drenaje que se utiliza en la desviación de aguas desbordadas de los alrededores de una ciudad. Los tiempos de drenaje de una represa conectada a cada uno de los cinco diseños de canal diferentes, obtenidos a partir de una serie de experimentos usando condiciones de flujo inicial similares, se presentan en la tabla siguiente.

Tipo de canal	Tiempo de drenaje (min)			
1	41.4	43.4	50.0	41.2
2	37.7	49.3	52.1	37.3
3	32.6	33.7	34.8	22.5
4	27.3	29.9	32.3	24.8
5	44.9	47.2	48.5	37.1

- a) ¿Puede concluir que hay una diferencia en la media de los tiempos de drenaje para los diferentes diseños de canal?
- b) ¿Cuáles pares de diseños, si hay, puede concluir que difieren en sus tiempos de drenaje?

12. Se estudió un proceso que fabrica vinilo para cubiertas de asientos de automóvil. Se variaron tres factores: el cociente de cierto plastificante (*A*), la tasa de extrusión (*B*), y la temperatura de secado (*C*). El resultado relevante era el espesor del vinilo (en milésimas de pulgada). Se empleó un diseño factorial 2^3 con cuatro réplicas. Los resultados se presentan en la tabla siguiente. (Basado en el artículo “Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables”, S. Kowalski, J. Cornell y G. Vining, en *Technometrics*, 2002:72-79.)

A	B	C	Espesor			
-1	-1	-1	7	5	6	7
1	-1	-1	6	5	5	5
-1	1	-1	8	8	4	6
1	1	-1	9	5	6	9
-1	-1	1	7	6	5	5
1	-1	1	7	7	11	10
-1	1	1	6	4	5	8
1	1	1	8	11	11	9

- a) Estime todos los efectos principales y las interacciones.
- b) Construya una tabla ANOVA. Puede dar los rangos para los *P*-valores.
- c) ¿El modelo aditivo es adecuado? Explique.
- d) ¿Qué conclusiones acerca de los factores puede obtener a partir de estos resultados?
13. En el artículo “Occurrence and Distribution of Ammonium in Iowa Groundwater” (K. Schilling, en *Water Environment Research*, 2002:177-186) se midieron concentraciones de amoníaco (en mg/L) en un gran número de pozos en el estado de Iowa. Éstos incluían cinco tipos de pozos base. El número de pozos de cada tipo, junto con la media y la desviación estándar de las concentraciones en esos pozos, se presentan en la tabla siguiente.

Tipo de pozo	Tamaño muestral	Media	Desviación estándar
Cretaceous	53	0.75	0.90
Mississippian	57	0.90	0.92
Devonian	66	0.68	1.03
Silurian	67	0.50	0.97
Cambrian-Ordovician	51	0.82	0.89

¿Puede concluir que la media de la concentración difiere entre los cinco tipos de pozos?

14. El artículo “Enthalpies and Entropies of Transfer of Electrolytes and Ions from Water to Mixed Aqueous Organic Solvents” (G. Hefter, Y. Marcus y W. Waghorne, en *Chemical Reviews*, 2002:2773-2836) presenta mediciones de los cambios de entropía y

entalpía para muchas sales en una variedad de condiciones. La tabla siguiente presenta los resultados de la transferencia de entalpía (en kJ/mol) de agua a agua + metanol de NaCl (sal de mesa) para varias concentraciones de metanol. Se realizaron cuatro mediciones independientes en cada concentración.

Concentración (%)	Entalpía			
5	1.62	1.60	1.62	1.66
10	2.69	2.66	2.72	2.73
20	3.56	3.45	3.65	3.52
30	3.35	3.18	3.40	3.06

- a) ¿Es creíble que la entalpía sea la misma en todas las concentraciones? Explique.
 b) ¿Cuáles pares de concentraciones, si hay, puede concluir que tienen entalpías diferentes?

15. Considere el ejercicio 11.

- a) Calcule la cantidad $s = \sqrt{\text{MSE}}$, la estimación de la desviación estándar del error σ .
 b) Suponga que s sea la desviación estándar del error, determine el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.90 y así detectar una diferencia máxima de 10 en las medias del tratamiento con un nivel de 5%.
 c) Utilizando una estimación más conservadora de $1.5s$ como la desviación estándar del error, determine el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.90 y así detectar una diferencia máxima de 10 en las medias de tratamiento con un nivel de 5%.

16. Considere el ejercicio 14.

- a) Calcule la cantidad $s = \sqrt{\text{MSE}}$, la estimación de la desviación estándar del error σ .
 b) Suponga que s sea la desviación estándar del error, encuentre el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.80 y así detectar una diferencia máxima de 0.1 en las medias de tratamiento con un nivel de 5%.
 c) Utilizando una estimación más conservadora de $1.5s$ como la desviación estándar del error, determine el tamaño de muestra necesario en cada tratamiento para proporcionar una potencia de 0.80 y así detectar una diferencia máxima de 0.1 en las medias de tratamiento con un nivel de 5%.

17. El artículo "Factorial Experiments in the Optimization of Alkaline Wastewater Pretreatment" (M. Prisciandaro, A. Del Borghi y F. Veglio, en *Industrial Engineering and Chemistry Research*, 2002:5034-5041) presenta los resultados de algunos experimentos para investigar los métodos de tratamientos de agua de residuos alcalinos. Se diseñó un experimento 2^4 sin réplicas. Los cuatro factores fueron A: concentración de ácido sulfúrico, B: temperatura, C: tiempo y D: cloruro de calcio de la concentración. La variable resultado es la cantidad de precipitados en kg/m^3 . Los resultados se presentan en la tabla siguiente.

A	B	C	D	Resultado	A	B	C	D	Resultado
-1	-1	-1	-1	6.4	-1	-1	-1	1	11.9
1	-1	-1	-1	12.9	1	-1	-1	1	13.1
-1	1	-1	-1	8.6	-1	1	-1	1	12.1
1	1	-1	-1	12.9	1	1	-1	1	16.0
-1	-1	1	-1	7.4	-1	-1	1	1	12.4
1	-1	1	-1	12.0	1	-1	1	1	16.5
-1	1	1	-1	10.7	-1	1	1	1	15.3
1	1	1	-1	15.0	1	1	1	1	18.3

- a) Calcule todos los efectos principales y las interacciones.
- b) ¿Qué efectos parecen ser mayores que los otros?
- c) Suponga que todas las interacciones de tercer y superior órdenes son iguales a 0, y sume las sumas de los cuadrados. Utilice el resultado en lugar de una suma de los cuadrados del error para calcular los estadísticos F y los P -valores de los efectos principales. ¿Cuáles factores puede concluir que tienen un efecto sobre el resultado?
- d) El artículo describe algunas réplicas del experimento, en las cuales el error medio cuadrático se encuentra de 1.04, con cuatro grados de libertad. Utilizando este valor, calcule los estadísticos F y los P -valores de todos los efectos principales y las interacciones.
- e) ¿Los resultados del inciso d) ayudan a justificar el supuesto de que las interacciones de tercer y superior órdenes son iguales a 0? Explique.
- f) Utilizando los resultados del inciso d), ¿cuáles factores puede concluir que tienen un efecto sobre el resultado?

18. El Williamsburg es un puente colgante que cruza el río East y conecta a Brooklyn y Manhattan en Nueva York. Una evaluación de las fuerzas de sus cables se reporta en el artículo “Estimating Strength of the Williamsburg Bridge Cables” (R. Perry, en *The American Statistician*, 2002:211-217). Cada cable de suspensión consta de 7 696 alambres. De uno de los cables se tomó una muestra de alambres de 128 puntos. Estos puntos provenían de cuatro posiciones a lo largo de la longitud del cable (I, II, III, IV). En cada ubicación había ocho puntos igualmente espaciados alrededor de la circunferencia del cable (A, B, C, D, E, F, G, H). De cada uno de los ocho puntos se tomaron muestras de alambre de cuatro profundidades: 1) la superficie externa del cable, 2) dos pulgadas de profundidad, 3) cuatro pulgadas de profundidad y 4) siete pulgadas de profundidad (el cable tiene un radio de 9.625 pulgadas). Bajo los supuestos hechos en el artículo, es adecuado considerar esto último como un experimento de dos factores con la posición en la circunferencia y la profundidad como los factores, y con la posición proporcionando cuatro réplicas respecto cada combinación de estos factores. La fuerza de ruptura mínima (en lbf) de cada uno de los 128 puntos se presenta en la tabla siguiente.

Circunferencia	Profundidad	Posición			
		I	II	III	IV
A	1	6 250	5 910	5 980	5 800
A	2	6 650	6 690	6 780	5 540
A	3	5 390	6 080	6 550	5 690
A	4	6 510	6 580	6 700	5 980
B	1	6 200	6 240	6 180	6 740
B	2	6 430	6 590	6 500	6 110
B	3	5 710	6 230	6 450	6 310
B	4	6 510	6 600	6 250	5 660
C	1	5 570	5 700	6 390	6 170
C	2	6 260	6 290	5 630	6 990
C	3	6 050	6 120	6 290	5 800
C	4	6 390	6 540	6 590	6 620
D	1	6 140	6 210	5 710	5 090
D	2	5 090	6 000	6 020	6 480
D	3	5 280	5 650	5 410	5 730
D	4	6 300	6 320	6 650	6 050
E	1	4 890	4 830	5 000	6 490
E	2	5 360	5 640	5 920	6 390
E	3	5 600	5 500	6 250	6 510
E	4	6 640	6 810	5 760	5 200
F	1	5 920	5 300	5 670	6 200
F	2	5 880	5 840	7 270	5 230

Circunferencia	Profundidad	Posición			
		I	II	III	IV
F	3	6 570	6 130	5 800	6 200
F	4	6 120	6 430	6 100	6 370
G	1	6 070	6 980	6 570	6 980
G	2	6 180	6 340	6 830	6 260
G	3	6 360	6 420	6 370	6 550
G	4	6 340	6 380	6 480	7 020
H	1	5 950	5 950	6 450	5 870
H	2	6 180	6 560	5 730	6 550
H	3	6 560	6 560	6 450	6 790
H	4	6 700	6 690	6 670	6 600

- a) Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- b) ¿Puede concluir que hay interacciones entre la posición en la circunferencia y la profundidad? Explique.
- c) ¿Puede concluir que la tensión varía con la posición en la circunferencia? Explique.
- d) ¿Puede concluir que la tensión varía con la profundidad? Explique.

19. En el artículo “Nitrate Contamination of Alluvial Groundwaters in the Nakdong River Basin, Korea” (J. Min, S. Yun y cols., *Geosciences Journal*, 2002:35-46) se midieron varias propiedades químicas de muestras de agua tomadas de pozos de irrigación en tres ubicaciones. La tabla siguiente presenta las medias, las desviaciones estándar y los tamaños de muestra de las mediciones de pH.

Ubicación	Media	DE	Tamaño de muestra
Río arriba	6.0	0.2	49
Río en medio	6.2	0.4	31
Río abajo	6.4	0.6	30

¿Los datos prueban concluyentemente que el pH difiere en las diferentes ubicaciones?

20. El artículo citado en el ejercicio 19 proporciona las medidas de la conductividad eléctrica (en $\mu\text{S}/\text{cm}$). Los resultados se presentan en la tabla siguiente.

Ubicación	Media	DE	Tamaño de muestra
Río arriba	463	208	49
Río en medio	363	98	31
Río abajo	647	878	30

¿Puede utilizarse un análisis de varianza de un sentido para determinar si la conductividad varía con la ubicación? ¿O uno de los supuestos necesariamente no se cumple? Explique.

21. El artículo “Factorial Experiments in the Optimization of Alkaline Wastewater Pretreatment” (M. Prisciandaro, A. Del Borghi y F. Veglio, en *Industrial Engineering and Chemistry Research*, 2002:5034-5041) presenta los resultados de un experimento para investigar los efectos de las concentraciones de ácido sulfúrico (H_2SO_4) y cloruro de calcio (CaCl_2) en la cantidad de barro negro que se precipita en el tratamiento de agua de residuos alcalinos. Había tres niveles de cada concentración y se reali-

zaron dos réplicas del experimento en cada combinación de niveles. Los resultados se presentan en la tabla siguiente (todas las mediciones están en unidades de kg/m^3).

H_2SO_4	CaCl_2	Precipitado	
110	15	100.2	98.2
110	30	175.8	176.2
110	45	216.5	206.0
123	15	110.5	105.5
123	30	184.7	189.0
123	45	234.0	222.0
136	15	106.5	107.0
136	30	181.7	189.0
136	45	211.8	201.3

- Construya una tabla ANOVA. Puede dar los rangos para los P -valores.
- ¿Es creíble el modelo aditivo? Explique.
- ¿Puede concluir que la concentración de H_2SO_4 afecta la cantidad de precipitado? Explique.
- ¿Puede concluir que la concentración de CaCl_2 afecta la cantidad de precipitado? Explique.

22. Las inclusiones de un fluido son volúmenes microscópicos de fluido atrapados en la roca durante la formación de ésta. El artículo “Fluid Inclusion Study of Metamorphic Gold-Quartz Veins in Northwestern Nevada, U.S.A.: Characteristics of Tectonically Induced Fluid” (S. Cheong, en *Geosciences Journal*, 2002:103-115) describe las propiedades geoquímicas de las inclusiones inestables en algunos filones diferentes en el noroeste de Nevada. La tabla siguiente presenta los datos acerca de la salinidad máxima (% de NaCl por peso) de las inclusiones en algunas muestras de roca de algunas áreas.

Área	Salinidad				
Humboldt Range	9.2	10.0	11.2	8.8	
Santa Rosa Range	5.2	6.1	8.3		
Ten Mile	7.9	6.7	9.5	7.3	10.4
Antelope Range	6.7	8.4	9.9		
Pine Forest Range	10.5	16.7	17.5	15.3	20.0

¿Puede concluir que la salinidad difiere entre las áreas?

23. El artículo “Effect of Microstructure and Weathering on the Strength Anisotropy of Porous Rhyolite” (Y. Matsukura, K. Hashizume, y C. Oguchi, en *Engineering Geology*, 2002:39-47) investiga la relación entre el ángulo entre la división y la estructura de flujo y la fuerza de la riolita porosa. Se midieron las fuerzas (en MPa) para varios elementos cortados en diferentes ángulos. La media y la desviación estándar de las fuerzas para cada ángulo se presentan en la tabla siguiente.

Ángulo	Media	Desviación estándar	Tamaño de muestra
0°	22.9	2.98	12
15°	22.9	1.16	6
30°	19.7	3.00	4
45°	14.9	2.99	5
60°	13.5	2.33	7
75°	11.9	2.10	6
90°	14.3	3.95	6

¿Puede concluir que la fuerza varía con el ángulo?

24. El artículo “Influence of Supplemental Acetate on Bioremediation for Dissolved Polycyclic Aromatic Hydrocarbons” (T. Ebihara y P. Bishop, en *Journal of Environmental Engineering*, 2002:505-513) describe experimentos en los que agua con hidrocarburos de poliaromáticos disueltos (PAH) era introducida en columnas de arena. Se midieron las concentraciones de PAH en diferentes profundidades después de 25, 45 y 90 días. Suponga que se realizaron tres mediciones independientes en cada profundidad cada vez. Los datos que se presentan en la tabla siguiente son las concentraciones de naftaleno (en mg/L) que son consistentes con las medias y desviaciones estándar descritas en el artículo.

Profundidad	25 días			45 días			90 días		
0	11.15	11.39	11.36	9.28	8.15	8.59	7.68	7.59	7.41
5	14.40	11.78	11.92	9.44	9.34	9.33	7.53	7.92	7.12
15	11.51	11.01	11.09	9.34	9.11	8.94	7.43	7.47	7.53
30	12.77	12.18	11.65	9.37	9.27	9.05	7.60	7.48	7.84
50	11.71	11.29	11.20	9.25	8.97	9.29	7.76	7.84	7.68
75	11.18	11.45	11.27	9.09	8.86	8.78	7.72	7.61	7.74

- a) Construya una tabla ANOVA. Puede dar rangos para los P -valores.
- b) Realice una prueba para determinar si el modelo aditivo es creíble. Proporcione el valor del estadístico de prueba y el P -valor.

Capítulo 10

Control estadístico de calidad

Introducción

A medida que el mercado industrial de bienes se hace más global, los fabricantes se han dado cuenta que para ser competitivos la calidad y la confiabilidad de sus productos deben ser lo más altas posible. Hoy se reconoce, por lo general, que la forma más eficiente en costos para mantener una alta calidad se logra mediante un seguimiento constante del proceso de producción. A menudo dicho seguimiento se realiza al extraer una muestra de las unidades de la producción y medir cierta característica de calidad. Debido a que las unidades se extraen de determinada población grande, estos métodos tienen una naturaleza inherentemente estadística.

Uno de los pioneros en el área del control de calidad estadístico fue el doctor Walter A. Shewart, de los laboratorios de la compañía telefónica Bell. En 1924 desarrolló la gráfica del control moderna, la cual sigue siendo hasta hoy una de las herramientas más utilizadas en el control de calidad. Después de la Segunda Guerra Mundial, los descubrimientos de W. Edwards Deming fueron cruciales para despertar el interés en el control de calidad, primero en Japón y después en Estados Unidos y otros países. El científico japonés Genichi Taguchi tuvo también un papel destacado, al desarrollar métodos de diseño experimental con una visión de mejoramiento de la calidad. En este capítulo se destacarán los diagramas de control de Shewart y los diagramas de suma acumulada (CUSUM, por sus siglas en inglés), puesto que éstas son las herramientas más poderosas entre las utilizadas frecuentemente en el control estadístico de calidad.

10.1 Ideas básicas

El principio básico de los diagramas de control consiste en que en cualquier proceso hay siempre una variación en el resultado. Parte de esta variación se debe a causas inherentes al proceso y es muy difícil o imposible especificarlas. A estas causas se les llama **causas comunes** o **probables**. Cuando éstas son las únicas causas de la variación, se dice que el proceso está **en un estado de control estadístico**, o de forma sencilla, **en control**.

Algunas veces se presentan factores especiales que producen mayor variabilidad. Entre los factores más comunes se encuentran las máquinas que no funcionan adecuadamente, errores del operador, fluctuaciones en las condiciones ambientales y variaciones en las propiedades de las materias primas. A éstas se les denomina **causas especiales** o **asignables**. Por lo general, estas últimas producen mayor nivel de variabilidad que las causas comunes; a esta variabilidad se le considera como inaceptable. Cuando un proceso opera cuando se presenta una o más causas especiales, se dice que está **fuera de control estadístico**.

Los diagramas de control le permiten al ingeniero responsable del mismo decidir si un proceso parece estar en control, o si están presente una o más causas especiales. Si se determina que el proceso está fuera de control, debe definirse y corregirse la naturaleza de la causa especial, de tal forma que el proceso vuelva al estado de control estadístico. Existen diversos tipos de diagramas de control; el uso de determinado diagrama dependerá de si la característica de calidad que se midió es una **variable continua**, **variable binaria**, o **variable discreta**. Por ejemplo, cuando se hace el seguimiento de un proceso que fabrica latas de aluminio para bebidas, debe medirse la altura de cada lata de la muestra. La altura es una variable continua. En ciertas ocasiones, basta simplemente con determinar si la altura se encuentra dentro de ciertos límites requeridos. En este caso, la medición de la calidad toma sólo uno de los dos valores: adecuado (dentro de los límites) o no adecuado (fuera de los límites). Esta medición es una variable binaria, ya que sólo tiene dos valores. Finalmente, podría tenerse interés en contar el número de imperfecciones en la superficie de la lata. Ésta es una variable discreta.

A los diagramas de control utilizados para variables continuas se les denominan **diagramas de control de variables**. Entre los ejemplos están el diagrama $\bar{X}$, el diagrama R y el diagrama S . A los diagramas de control utilizados para variables binarias o discretas se les llama **diagramas de control de atributos**. El diagrama p es el diagrama más comúnmente utilizado para variables binarias, mientras que el diagrama c se utiliza para variables discretas.

Recolección de datos: subgrupos racionales

Los datos utilizados en la construcción de un diagrama de control se recolectan mediante cierto número de muestras tomadas durante determinado periodo. A estas muestras se les conoce como **subgrupos racionales**. Existen muchas estrategias diferentes para escoger subgrupos racionales. El principio básico que debía seguirse es que toda la variabilidad dentro de las unidades de un subgrupo racional debe ser resultado de causas comunes, y ninguna debe ser resultado de causas especiales. En general, una buena forma de elegir subgrupos racionales consiste en definir qué causas especiales son las más importantes de detectar, y enseguida elegir a los subgrupos racionales en lo que es más probable detectarlas. Los dos métodos comúnmente más utilizados son:

- Muestreo en intervalos regulares, en el que todos los elementos de cada muestra son producidos casi en el mismo tiempo en que se realiza el muestreo.
- Muestreo en intervalos regulares, en el que los elementos de cada muestra se extraen a partir de todas las unidades producidas desde que se tomó la última muestra.

Para datos de variables la mayoría de las veces el número de unidades en cada muestra es pequeño, con frecuencia se encuentra entre tres y ocho. El número de muestras debe ser al menos veinte. En general, es mejor tomar con cierta frecuencia muchas muestras pequeñas que tomar pocas muestras con poca frecuencia. Para los datos binarios o discretos, en general, las muestras deben ser más grandes.

Control contra capacidad

Es importante entender la diferencia entre el *control* del proceso y la *capacidad* del proceso. Un proceso está en control si no opera ninguna causa especial. La característica distintiva de un proceso que está en control consiste en que los valores de la característica de control varían sin seguir alguna tendencia o patrón, ya que las causas comunes no cambian a lo largo del tiempo. Sin embargo, es muy posible que un proceso esté en control y aún produzca unidades que no satisfagan cierta especificación. Por ejemplo, suponga que un proceso produce barras de acero cuyas longitudes varían de forma aleatoria entre 19.9 y 20.1 cm, sin ningún patrón aparente de fluctuación. Este proceso se encuentra en un estado de control. No obstante, si el diseño requiere una longitud entre 21 y 21.2 cm, muy poca de la producción cumpliría con la especificación. A la **capacidad** de un proceso para producir elementos que satisfagan con una especificación dada se le llama **capacidad** del proceso. En la sección 10.5 se analizará la medición de este último.

El control del proceso debe realizarse continuamente

Existen tres fases básicas en el uso de los diagramas de control. Primero, se reúnen los datos; luego se grafican para determinar si el proceso está en control y, por último, una vez que el proceso ya está en control, debe estimarse su capacidad. Por supuesto, un proceso que en determinado momento está en control y tiene capacidad para estar fuera de control posteriormente, conforme se vuelvan a presentar las causas especiales. Por esta razón debe darse seguimiento a los procesos continuamente.

Ejercicios para la sección 10.1

1. Indique si cada una de las siguientes características de control es una variable continua, binaria o discreta.
 - a) La longitud de una barra de acero.
 - b) El número de imperfecciones que hay en una sección de una lámina de metal.
 - c) Si un tipo de concreto cumple con la resistencia requerida.
 - d) El tiempo necesario para realizar una inspección final a un producto acabado.
 - e) Si un proceso está en un estado de control estadístico, entonces casi no habrá variación en la producción.
2. Verdadero o falso:
 - a) Los diagramas de control se utilizan para determinar si están en curso causas especiales.
 - b) Si no está en curso ninguna causa especial, entonces la mayoría de la producción cumplirá con la especificación.
 - c) La variabilidad debida a causas comunes no aumenta ni disminuye demasiado en periodos cortos.
 - d) La variabilidad dentro de los elementos de un subgrupo racional es resultado de causas especiales.
3. Complete los enunciados. Las opciones son: *está en control; tiene una capacidad alta*.
 - a) Si la variabilidad en un proceso es aproximadamente constante en el tiempo, el proceso _____.
 - b) Si la mayoría de las unidades producidas cumplen la especificación, el proceso _____.
4. Complete el enunciado: Una vez que un proceso está en un estado de control estadístico, _____.
 - i) El seguimiento debe seguirse realizando constantemente.
 - ii) El seguimiento debe detenerse durante cierto tiempo, ya que es poco probable que el proceso se salga de control inmediatamente.
 - iii) No debe darse seguimiento nuevamente al proceso, a menos que éste sea rediseñado.

5. Verdadero o falso:
- a) Cuando un proceso se encuentra en un estado de control estadístico, entonces la mayoría de la producción cumplirá con las especificaciones.
 - b) Cuando un proceso está fuera de control, una proporción inaceptablemente grande de la producción no cumplirá con las especificaciones.
 - c) Cuando un proceso está en un estado de control estadístico, toda la variación en el proceso se debe a causas inherentes al propio proceso.
 - d) Cuando un proceso está fuera de control, parte de la variación en el proceso se debe a causas externas al proceso.
6. Complete el enunciado: Cuando se extraen muestras de unidades para subgrupos racionales, _____.
- a) Es más importante elegir muestras grandes que realizar el muestreo frecuentemente, debido a que las muestras grandes dan una información más precisa sobre el proceso.
 - b) Es más importante realizar el muestreo frecuentemente que elegir muestras grandes, de tal manera que las causas especiales puedan ser detectarse con mayor rapidez.

10.2 Diagramas de control para variables

Cuando se realiza una medición de calidad en una escala de variables, a los datos se les conoce como **datos de variables**. Para estos datos, primero se utiliza un diagrama R o un diagrama S con el fin de controlar la variabilidad del proceso, y después se utiliza un diagrama $\bar{X}$ para controlar la media del proceso. Los métodos descritos en esta sección suponen que todas las mediciones siguen una distribución aproximadamente normal.

Lo anterior se ilustra con un ejemplo. El ingeniero de control de calidad a cargo de un proceso de empaquetado de sal está preocupado por el contenido de humedad en los paquetes de sal. Para determinar si el proceso está en control estadístico, primero se necesitan definir los subgrupos racionales, y después recolectar ciertos datos. Suponga que para el proceso de empaquetado de sal la preocupación principal es que la variación en la humedad del ambiente de la planta pueda variar el contenido de humedad promedio de los paquetes a lo largo del tiempo. Recuerde que los subgrupos racionales deben elegirse de tal forma que la variación dentro de cada muestra se deba sólo a causas comunes, no a causas especiales. Por consiguiente, en este caso una buena opción para los subgrupos racionales es extraer muestras de diversos paquetes, cada una en intervalos regulares. Los paquetes en cada muestra serán producidos lo más próximos en tiempo entre sí. De este modo, la humedad del ambiente será casi la misma para cada paquete de la muestra, de forma que la variación dentro del grupo no se verá afectada por esta causa especial. Suponga que se extraen cinco paquetes de sal cada 15 minutos durante ocho horas, y que el contenido de humedad en cada paquete se mide como un porcentaje del peso total. Los datos se presentan en la tabla 10.1.

Como consecuencia de que la muestra se mide en una escala continua, éstos son datos de variables. Cada renglón de la tabla 10.1 presenta las cinco mediciones de humedad en una muestra dada, junto con su media muestral $\bar{X}$, su desviación estándar muestral s y su rango muestral R (la diferencia entre el valor mayor y el menor). El último renglón de la tabla contiene la media de las medias muestrales ($\bar{\bar{X}}$), la media de los rangos muestrales ($\bar{R}$) y la media de las desviaciones estándar muestrales ($\bar{s}$).

Se supone que cada una de las 32 muestras en la tabla 10.1 constituye una muestra de una población normal con media μ y desviación estándar σ . A la cantidad μ se le llama **media del proceso**, y a σ , **desviación estándar del proceso**. La idea detrás de los diagramas de

TABLA 10.1 Contenido de humedad en los paquetes de sal, como porcentaje del peso total

Muestra	Valores muestrales					Media ($\bar{X}$)	Rango (R)	SD (s)
1	2.53	2.66	1.88	2.21	2.26	2.308	0.780	0.303
2	2.69	2.38	2.34	2.47	2.61	2.498	0.350	0.149
3	2.67	2.23	2.10	2.43	2.54	2.394	0.570	0.230
4	2.10	2.26	2.51	2.58	2.28	2.346	0.480	0.196
5	2.64	2.42	2.56	2.51	2.36	2.498	0.280	0.111
6	2.64	1.63	2.95	2.12	2.67	2.402	1.320	0.525
7	2.58	2.69	3.01	3.01	2.23	2.704	0.780	0.327
8	2.31	2.39	2.60	2.40	2.46	2.432	0.290	0.108
9	3.03	2.68	2.27	2.54	2.63	2.630	0.760	0.274
10	2.86	3.22	2.72	3.09	2.48	2.874	0.740	0.294
11	2.71	2.80	3.09	2.60	3.39	2.918	0.790	0.320
12	2.95	3.54	2.59	3.31	2.87	3.052	0.950	0.375
13	3.14	2.84	3.77	2.80	3.22	3.154	0.970	0.390
14	2.85	3.29	3.25	3.35	3.59	3.266	0.740	0.267
15	2.82	3.71	3.36	2.95	3.37	3.242	0.890	0.358
16	3.17	3.07	3.14	3.63	3.70	3.342	0.630	0.298
17	2.81	3.21	2.95	3.04	2.85	2.972	0.400	0.160
18	2.99	2.65	2.79	2.80	2.95	2.836	0.340	0.137
19	3.11	2.74	2.59	3.01	3.03	2.896	0.520	0.221
20	2.83	2.74	3.03	2.68	2.49	2.754	0.540	0.198
21	2.76	2.85	2.59	2.23	2.87	2.660	0.640	0.265
22	2.54	2.63	2.32	2.48	2.93	2.580	0.610	0.226
23	2.27	2.54	2.82	2.11	2.69	2.486	0.710	0.293
24	2.40	2.62	2.84	2.50	2.51	2.574	0.440	0.168
25	2.41	2.72	2.29	2.35	2.63	2.480	0.430	0.186
26	2.40	2.33	2.40	2.02	2.43	2.316	0.410	0.169
27	2.56	2.47	2.11	2.43	2.85	2.484	0.740	0.266
28	2.21	2.61	2.59	2.24	2.34	2.398	0.400	0.191
29	2.56	2.26	1.95	2.26	2.40	2.286	0.610	0.225
30	2.42	2.37	2.13	2.09	2.41	2.284	0.330	0.161
31	2.62	2.11	2.47	2.27	2.49	2.392	0.510	0.201
32	2.21	2.15	2.18	2.59	2.61	2.348	0.460	0.231
						$\bar{\bar{X}} = 2.6502$	$\bar{R} = 0.6066$	$\bar{s} = 0.2445$

control es que cada valor de $\bar{X}$ se aproxima a la media del proceso durante el tiempo en que se tomó la muestra, mientras que los valores de R y s se pueden utilizar para aproximar la desviación estándar muestral. Si el proceso está en control, entonces la media y la desviación estándar del proceso son iguales en cada muestra. Si el proceso está fuera de control, la media del proceso μ o la desviación estándar σ , o ambas, diferirán de muestra en muestra. Por tanto, los valores de $\bar{X}$, R , y s variarán menos cuando el proceso esté en control que cuando el proceso esté fuera de control. Si fuera el primer caso, los valores de $\bar{X}$, R , y s se mantendrán casi siempre dentro de límites calculables, denominados **límites de control**. Si el proceso está fuera de control, es más probable que los valores de $\bar{X}$, R , o s excedan estos límites. Un diagrama de control grafica los valores de $\bar{X}$, R , o s junto con los límites de control, de forma que puede verse muy fácilmente si la variación es lo bastante grande para concluir que el proceso está fuera de control.

Ahora se verá cómo determinar si el proceso de empaquetado de sal se encuentra en un estado de control estadístico con respecto al contenido de humedad. Puesto que se supone que la variación dentro de cada muestra se debe solamente a causas comunes, esta variación no debe ser muy diferente de una muestra a otra. Por consiguiente, lo primero es comprobar para tener la certeza que la cantidad de la variación dentro de cada muestra, ya sea medida por el rango muestral o desviación estándar muestral, no varía demasiado de una muestra a otra. Con este propósito se utiliza el diagrama R para estimar la variación en el rango muestral, o puede usarse el diagrama S para calcular la variación en la desviación estándar muestral. Primero se estudiará el diagrama R , ya que es el más tradicional. Al final de esta sección se analizará el diagrama S .

La figura 10.1 presenta el diagrama R para los datos de humedad. El eje horizontal representa las muestras, numeradas del 1 al 32. Los rangos muestrales están graficados en el eje vertical. Las tres líneas horizontales son las más importantes. La recta en el centro del diagrama es el valor $\bar{R}$ y se le llama **recta central**. Las rectas superior e inferior indican los **límites de control** 3σ superior e inferior (UCL y LCL, respectivamente). Los límites de control están dibujados de tal forma que cuando el proceso está en control, casi todos los puntos estarán dentro de los límites. Un punto graficado fuera de los límites de control es una evidencia de que el proceso no está en control.

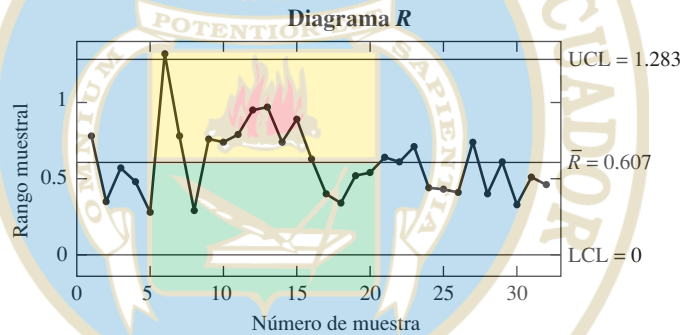


FIGURA 10.1 Diagrama R para los datos de humedad.

Para entender dónde se ubican en el diagrama los límites de control, suponga que los 32 rangos muestrales provienen de una población con media μ_R y desviación estándar σ_R . Los valores de μ_R y σ_R no se conocerán exactamente, pero se sabe que en la mayoría de las poblaciones es poco usual observar un valor que sea diferente a la media por más de tres desviaciones estándar. Por esta razón se tiene la costumbre de graficar los límites de control en puntos que se aproximan a los valores $\mu_R \pm 3\sigma_R$. Se puede demostrar mediante métodos avanzados que las cantidades $\mu_R \pm 3\sigma_R$ pueden estimarse con múltiplos de $\bar{R}$; dichos múltiplos se representan con D_3 y D_4 . La cantidad $\mu_R - 3\sigma_R$ se estima con $D_3\bar{R}$, y la cantidad $\mu_R + 3\sigma_R$ se estima con $D_4\bar{R}$. Las cantidades D_3 y D_4 son constantes cuyos valores dependen del tamaño de la muestra n . A continuación se presenta una tabla de valores de D_3 y D_4 . En la tabla 4.9 (Apéndice A) se muestra una tabulación más extensa. Observe que para tamaños de muestra de 6 o menos, el valor de D_3 es 0. Para estos tamaños de muestra pequeños, la cantidad $\mu_R - 3\sigma_R$ es negativa. En estos casos se establece que el límite de control inferior es 0, debido a que es imposible que el rango sea negativo.

n	2	3	4	5	6	7	8
D_3	0	0	0	0	0	0.076	0.136
D_4	3.267	2.575	2.282	2.114	2.004	1.924	1.864

Ejemplo

10.1

Calcule los límites de control 3σ superior e inferior del diagrama R para los datos de humedad de la tabla 10.1

Solución

El valor de $\bar{R}$ es 0.6066 (tabla 10.1). El tamaño de la muestra es $n = 5$. A partir de la tabla se tiene que $D_3 = 0$ y $D_4 = 2.114$. En consecuencia, el límite de control superior es $(2.114)(0.6066) = 1.283$, y el límite de control inferior es $(0)(0.6066) = 0$.

Resumen

En un diagrama R , la recta central y los límites de control 3σ superior e inferior están dados por

$$\text{Límite } 3\sigma \text{ superior} = D_4 \bar{R}$$

$$\text{Recta central} = \bar{R}$$

$$\text{Límite } 3\sigma \text{ inferior} = D_3 \bar{R}$$

Los valores D_3 y D_4 dependen del tamaño de la muestra. Los valores están tabulados en la tabla A.9.

Una vez que se han calculado los límites de control y se han graficado los puntos, puede utilizarse el diagrama R para determinar si el proceso está en control con respecto a la variación. La figura 10.1 muestra que el rango para la muestra 6 excede el límite de control superior, por lo que evidencia que estaba operando una causa especial y que la variación del proceso no está en control. La acción apropiada es determinar la naturaleza de la causa especial, y después eliminar la muestra que no está en control y calcular de nuevo los límites. Suponga que se descubre que un técnico tuvo la negligencia de cerrar una ventana, lo cual provocó una variación mayor a la usual en el contenido de humedad durante el tiempo en el que se eligió la muestra. Volver a entrenar al técnico corregirá esa causa especial. Se elimina la muestra 6 de los datos y se vuelve a calcular el diagrama R . Los resultados se muestran en la figura 10.2. Ahora, la variación del proceso está en control.

Ahora que se ha puesto en control la variación del proceso, se puede determinar si la media del proceso está bajo control al dibujar el diagrama $\bar{X}$. La figura 10.3 (página 730) presenta el diagrama $\bar{X}$. Las medias muestrales están graficadas en el eje vertical. Observe que no se ha utilizado en este diagrama la muestra 6, ya que es preciso eliminarla para que el proceso de variación esté en control. Como en todas los diagramas de control, el diagrama $\bar{X}$ tiene una recta central y límites de control superior e inferior.

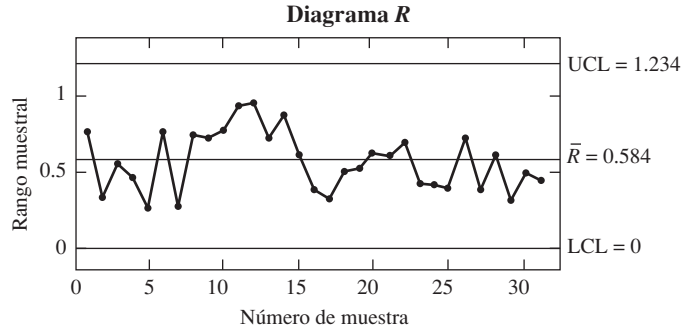


FIGURA 10.2 Diagrama R para los datos de humedad, después de eliminar la muestra fuera de control.

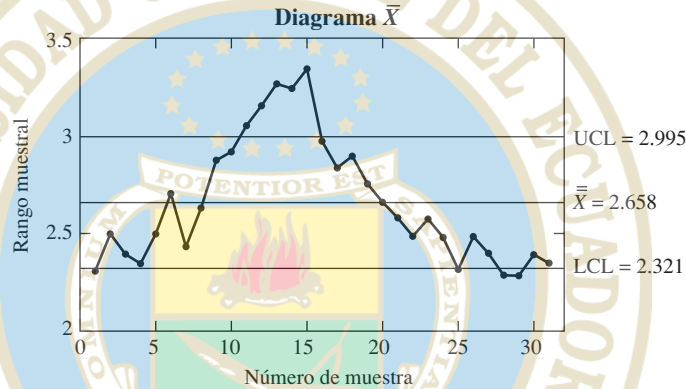


FIGURA 10.3 Diagrama $\bar{X}$ para los datos de humedad. La muestra 6 se ha eliminado con el fin de que el proceso de variación esté en control. Sin embargo, el diagrama $\bar{X}$ muestra que la media muestral está fuera de control.

Para calcular la recta central y los límites de control, se puede suponer que la desviación estándar del proceso es igual en todas las muestras, ya que se ha utilizado al diagrama R para que la variación del proceso esté en control. Si la media del proceso μ también está en control, entonces también es igual en todas las muestras. En ese caso, se extraen 32 medias muestrales de una población normal con media $\mu_{\bar{X}} = \mu$ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, donde n es el tamaño de la muestra, que en este caso es igual a 5. Idealmente, se desearía trazar la recta central en μ y los límites de control 3σ en $\mu \pm 3\sigma_{\bar{X}}$. Sin embargo, con frecuencia no se conocen los valores de μ y $\sigma_{\bar{X}}$ y se deben estimar a partir de los datos. Se estima μ con $\bar{\bar{X}}$, el promedio de las medias muestrales. Por consiguiente, la recta central se ubica en $\bar{\bar{X}}$. La cantidad $\sigma_{\bar{X}}$ se puede estimar ya sea mediante el rango promedio $\bar{R}$ o usando las desviaciones estándar muestrales. Aquí se usará $\bar{R}$ y al final de la sección se analizarán los métodos basados en la desviación estándar, junto con el análisis de los diagramas S . Se puede demostrar por medio de métodos avanzados que la cantidad $3\sigma_{\bar{X}}$ se estima con $A_2\bar{R}$, donde A_2 es una constante cuyo valor depende del tamaño de la muestra. A continuación se presenta una tabla de valores de A_2 . En la tabla A.9 se muestra una tabulación más extensa.

n	2	3	4	5	6	7	8
A_2	1.880	1.023	0.729	0.577	0.483	0.419	0.373

Resumen

En un diagrama $\bar{X}$, cuando se utiliza $\bar{R}$ para estimar a $\sigma_{\bar{X}}$, la recta central y los límites de control 3σ superior e inferior están dados por

$$\text{Límite } 3\sigma \text{ superior} = \bar{\bar{X}} + A_2\bar{R}$$

$$\text{Recta central} = \bar{\bar{X}}$$

$$\text{Límite } 3\sigma \text{ inferior} = \bar{\bar{X}} - A_2\bar{R}$$

El valor A_2 depende del tamaño de la muestra. Los valores están tabulados en la tabla A.9.

Ejemplo 10.2

Calcule los límites de control 3σ superior e inferior del diagrama $\bar{X}$ para los datos de humedad de la tabla 10.1.

Solución

Con la muestra 6 eliminada, el valor de $\bar{\bar{X}}$ es 2.658, y el valor de $\bar{R}$ es 0.5836. El tamaño de la muestra es $n = 5$. A partir de la tabla se tiene que $A_2 = 0.577$. Por tanto, el límite de control superior es $2.658 + (0.577)(0.5836) = 2.995$, y el límite de control inferior es $2.658 - (0.577)(0.5836) = 2.321$.

El diagrama $\bar{X}$ muestra claramente que la media del proceso no está en control, ya que hay diversos puntos ubicados fuera de los límites de control. El jefe de producción instala un higrómetro para darle seguimiento a la humedad del ambiente, y determina que las fluctuaciones en el contenido de humedad se deben a las fluctuaciones en la humedad del ambiente. Se instala un deshumidificador para estabilizar la humedad del ambiente. Después de haber solucionado esta causa especial se recolectan más datos y se construyen un nuevo diagrama R y un diagrama $\bar{X}$. La figura 10.4 (página 732) presenta los resultados. Ahora el proceso se encuentra en un estado de control estadístico. Obviamente, el proceso debe vigilarse en forma constante, debido a que las causas especiales tienen tendencia a cambiar en forma inesperada cada vez y es preciso detectarlas y corregirlas.

Observe que mientras los diagramas de control pueden detectar la presencia de una causa especial, éstos no pueden determinar su naturaleza ni cómo corregirlas. Es necesario que el ingeniero del proceso entienda bien éste para que puedan diagnosticarse y corregirse las causas especiales, detectadas por los diagramas de control.

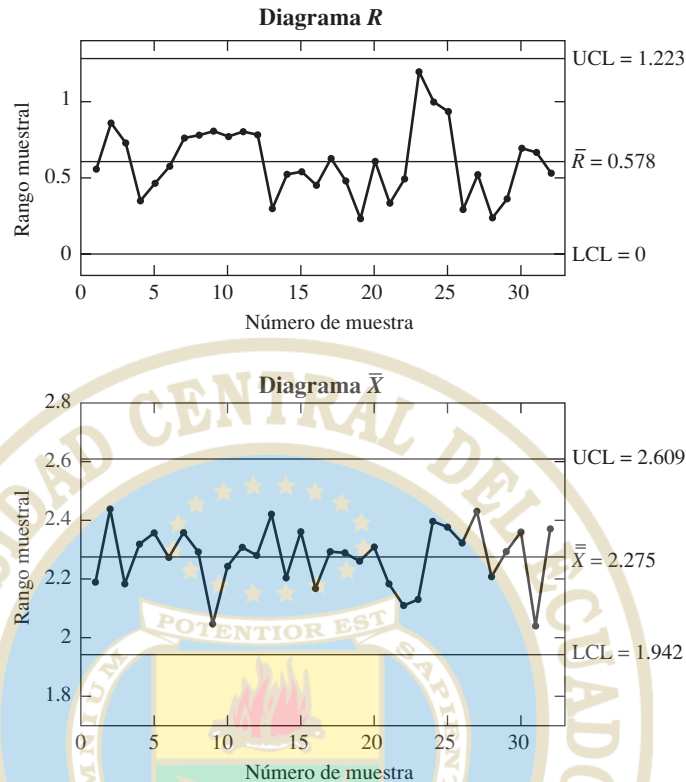


FIGURA 10.4 El diagrama R y el diagrama $\bar{X}$ después de que se ha corregido una causa especial. Ahora el proceso se encuentra en un estado de control estadístico.

Resumen

Los pasos que deben seguirse cuando se utilizan el diagrama R y el diagrama $\bar{X}$ son:

1. Elija subgrupos racionales.
2. Calcule el diagrama R .
3. Determine las causas especiales para cualesquiera de los puntos que no estén en control.
4. Vuelva a construir el diagrama R ; omita las muestras que resultaron en puntos fuera de control.
5. Una vez que el diagrama R indique que se está en un estado de control, calcule el diagrama $\bar{X}$, omitiendo las muestras que resultaron en puntos fuera de control en el diagrama R .
6. Si el diagrama $\bar{X}$ indica que el proceso no está en control, identifique y corrija algunas de las causas especiales.
7. Continúe dando seguimiento a $\bar{X}$ y R .

Funcionamiento del diagrama de control

Existe una estrecha conexión entre los diagramas de control y las pruebas de hipótesis. La hipótesis nula es que los procesos están en un estado de control. Un punto graficado fuera de los límites de control 3σ representa una evidencia en contra de la hipótesis nula. Como en cualquier prueba de hipótesis, se puede cometer un error. Por ejemplo, ocasionalmente algún punto se encontrará fuera de los límites 3σ incluso cuando el proceso está en control. A esto se le llama **falsa alarma**. También puede suceder que un proceso no está en control o no tenga algún punto fuera de los límites de control, en especial si no se observan durante un periodo suficientemente largo. A esto se le denomina **falla de detección**.

Lo deseable es que estos errores ocurran lo menos frecuente posible. Se describe a la frecuencia con la cual estos errores se presentan con una cantidad llamada **longitud de corrida promedio (ARL, por sus siglas en inglés)**. La ARL es el número de muestras que deben observarse, en promedio, antes de que un punto sea graficado fuera de los límites de control. Se desea que la ARL sea grande cuando el proceso está en control, y pequeña cuando está fuera de control. Se puede calcular la ARL de un diagrama $\bar{X}$ si se supone que se conocen la media del proceso μ y la desviación estándar del proceso σ . Entonces la recta central se localizará en la media del proceso μ y los límites centrales estarán en $\mu \pm 3\sigma_{\bar{X}}$. También debe suponerse, como siempre ha sido el caso con el diagrama $\bar{X}$, que la cantidad que se medirá sigue una distribución aproximadamente normal. Los ejemplos 10.3 al 10.6 muestran cómo calcular la ARL.

Ejemplo 10.3

Para un diagrama $\bar{X}$ con límites de control en $\mu \pm 3\sigma_{\bar{X}}$, calcule la ARL de un proceso que está en control.

Solución

Sea $\bar{X}$ la media de una muestra. Entonces $\bar{X} \sim N(\mu, \sigma_{\bar{X}}^2)$. La probabilidad de que un punto esté graficado fuera de los límites de control es igual a $P(\bar{X} < \mu - 3\sigma_{\bar{X}}) + P(\bar{X} > \mu + 3\sigma_{\bar{X}})$. Esta probabilidad es igual a $0.00135 + 0.00135 = 0.0027$ (véase la figura 10.5). Por tanto, en promedio, 27 de cada 10 000 puntos estarán graficados fuera de los límites de control. Esto es equivalente a 1 de cada $10\,000/27 = 370.4$ puntos. Por consiguiente, la longitud de corrida promedio es igual a 370.4.

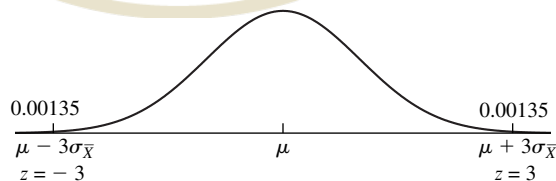


FIGURA 10.5 La probabilidad de que un punto esté graficado fuera de los límites de control 3σ , cuando el proceso está en control, es 0.0027 ($0.00135 + 0.00135$).

El resultado del ejemplo 10.3 puede interpretarse de la siguiente manera: si un proceso está en control, se espera observar alrededor de 370 muestras, en promedio, antes de encontrar

trar una cuyos diagramas estén fuera de los límites de control, provocando una falsa alarma. Observe también que la ARL en el ejemplo 10.3 fue de $10\,000/27$, que es igual a $1/0.0027$, donde 0.0027 es la probabilidad de cualquiera de los diagramas de la muestra estén fuera de los límites de control. En general, esto es cierto.

Resumen

La longitud de corrida promedio (ARL) es el número de muestras que se observarán, en promedio, antes de que algún punto del diagrama esté fuera de los límites de control. Si p es la probabilidad de que cualquier punto dado del diagrama esté fuera de los límites de control, entonces

$$ARL = \frac{1}{p} \quad (10.1)$$

Si un proceso está fuera de control, entonces la ARL será menor a 370.4. El ejemplo 10.4 muestra cómo calcular la ARL para una situación en la que el proceso cambia a una situación fuera de control.

Ejemplo 10.4

Un proceso tiene una media $\mu = 3$ y desviación estándar $\sigma = 1$. Se toman muestras de tamaño $n = 4$. Si una causa especial cambia a la media del proceso a un valor de 3.5, encuentre la ARL.

Solución

Primero se calcula la probabilidad p de que un punto del diagrama esté fuera de los límites de control. Entonces $ARL = 1/p$. Los límites de control se grafican con base en un proceso que está en control. Por consiguiente, éstos se encuentran en $\mu \pm 3\sigma_{\bar{X}}$, donde $\mu = 3$ y $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 1/\sqrt{4} = 0.5$. Por tanto, el límite de control inferior está en 1.5 y el de control superior está en 4.5. Si $\bar{X}$ es la media de una muestra tomada después de que cambió la media del proceso, entonces $\bar{X} \sim N(3.5, 0.5^2)$. La probabilidad de que el diagrama $\bar{X}$ se encuentre fuera de los límites de control es igual a $P(\bar{X} < 1.5) + P(\bar{X} > 4.5)$. Esta probabilidad es 0.0228 (véase la figura 10.6). En consecuencia, la ARL es igual a $1/0.0228 = 43.9$. Se tienen que observar alrededor de 44 muestras, en promedio, antes de detectar este cambio.

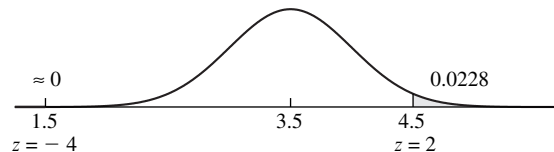


FIGURA 10.6 La media del proceso ha cambiado de $\mu = 3$ a $\mu = 3.5$. Ahora el límite de control superior se encuentra a sólo $2\sigma_{\bar{X}}$ de la media, que es indicado por el hecho de que $z = 2$. Ahora el límite de control inferior está a $4\sigma_{\bar{X}}$ por debajo de la media. La probabilidad de que los puntos del diagrama estén fuera de los límites de control es 0.0228 ($0 + 0.0228$).

Ejemplo 10.5

Con referencia al ejemplo 10.4, ¿a qué valor puede detectarse un cambio positivo con una ARL de 20?

Solución

M es la nueva media a la que ha cambiado el proceso. Dado que se ha especificado un cambio positivo, $m > 3$. En el ejemplo 10.4 se calculó que los límites de control eran 1.5 y 4.5. Si $\bar{X}$ es la media de la muestra tomada después de que la media del proceso cambió, entonces $\bar{X} \sim N(m, 0.5^2)$. La probabilidad de que el diagrama $\bar{X}$ esté fuera de los límites de control es igual a $P(\bar{X} < 1.5) + P(\bar{X} > 4.5)$ (véase la figura 10.7). Esta probabilidad es igual a $1/ARL = 1/20 = 0.05$. Puesto que $m > 3$, m se aproxima más a 4.5 que a 1.5. Se comenzará suponiendo que el área a la izquierda de 1.5 es insignificante y que el área a la derecha de 4.5 es igual a 0.05. El puntaje z de 4.5 es, entonces, 1.645, por lo que $(4.5 - m)/0.5 = 1.645$. Al despejar a m , se tiene que $m = 3.68$. Se termina por comprobar la suposición de que el área a la izquierda de 1.5 es insignificante. Con $m = 3.68$, el puntaje z para 1.5 es $(1.5 - 3.68)/0.5 = -4.36$. El área a la izquierda de 1.5 es en realidad insignificante.

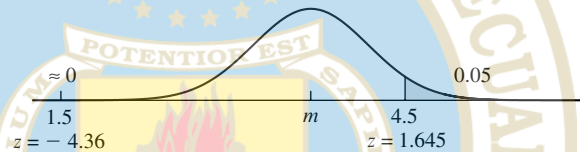


FIGURA 10.7 Solución al ejemplo 10.5.

Ejemplo 10.6

Con referencia al ejemplo 10.4, si el tamaño de la muestra sigue siendo $n = 4$, ¿cuál debe ser el valor de la desviación estándar σ para tener una ARL de 10 cuando la media del proceso cambia a 3.5?

Solución

Sea σ la nueva desviación estándar del proceso. Los nuevos límites de control son $3 \pm 3\sigma/\sqrt{n}$, o $3 \pm 3\sigma/2$. Si la media del proceso cambia a 3.5, entonces $\bar{X} \sim N(3.5, \sigma^2/4)$. La probabilidad de que el diagrama $\bar{X}$ esté fuera de los límites de control es igual a $P(\bar{X} < 3 - 3\sigma/2) + P(\bar{X} > 3 + 3\sigma/2)$. Esta probabilidad es igual a $1/ARL = 1/10 = 0.10$ (véase la figura 10.8, página 736). La media del proceso, 3.5, se aproxima más a $3 + 3\sigma/2$ que a $3 - 3\sigma/2$. Se supondrá que el área a la izquierda de $3 - 3\sigma/2$ es insignificante y que el área a la izquierda de $3 + 3\sigma/2$ es igual a 0.10. El puntaje z para $3 + 3\sigma/2$ es, entonces, 1.28, por lo que

$$\frac{(3 + 3\sigma/2) - 3.5}{\sigma/2} = 1.28$$

Al despejar σ se obtiene que $\sigma = 0.58$. Se concluye que el área a la izquierda de $3 - 3\sigma/2$ es insignificante. Al sustituir $\sigma = 0.58$, se obtiene que $3 - 3\sigma/2 = 2.13$. El puntaje z es $(2.13 - 3.5)/(0.58/2) = -4.72$. El área a la izquierda de $3 - 3\sigma/2$ es verdaderamente insignificante.

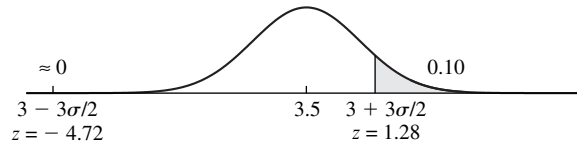


FIGURA 10.8 Solución al ejemplo 10.6.

Los ejemplos 10.4 a 10.6 muestran que usualmente los diagramas $\bar{X}$ no detectan los cambios pequeños de forma rápida. En otras palabras, la ARL es alta cuando los cambios en la media del proceso son pequeños. En principio, se puede reducir la ARL al mover los límites de control cerca de la recta central. Esto último reduciría el tamaño del cambio que se necesita para detectar una condición fuera de control, por lo que los cambios en la media del proceso se detectarían más rápidamente. No obstante, hay un intercambio. La tasa de falsa alarma también puede incrementarse, debido a que es más probable que ocurran cambios fuera de los límites de control. La situación se parece mucho a las pruebas de hipótesis de nivel fijo. La hipótesis nula consiste en que el proceso esté en control. El diagrama de control realiza una prueba de hipótesis para cada muestra. Cuando un punto del diagrama está fuera de los límites de control, se rechaza la hipótesis nula. Con los límites de control en $\pm 3\sigma_{\bar{X}}$, un error del tipo I (rechazo de una hipótesis nula verdadera) ocurrirá aproximadamente una vez por cada 370 muestras. El precio que debe pagarse por esta baja tasa de falsa alarma es la falta de capacidad de rechazar una hipótesis nula cuando ésta es falsa. La respuesta no es mover los límites de control para que se aproximen entre sí. Aunque incrementará la capacidad, también aumentará la razón de falsa alarma.

Dos de los medios por los cuales los practicantes han intentado mejorar su capacidad para detectar pequeños cambios de forma rápida son el uso de las **reglas de Western Electric** para interpretar el diagrama de control y el uso de los diagramas CUSUM. Enseguida se describirán las reglas de Western Electric. En la sección 10.4 se presentarán los diagramas CUSUM.

Las reglas de Western Electric

La figura 10.9 presenta un diagrama $\bar{X}$. A pesar de que ninguno de los puntos está fuera de los límites de control 3σ , claramente el proceso no se encuentra en un estado de control, ya que las medias muestrales siguen un patrón no aleatorio. Al darse cuenta de que un proceso puede estar fuera de control aun cuando ninguno de los puntos del diagrama se encuentre fuera de los límites de control, en 1956 los ingenieros de la compañía Western Electric propusieron una lista de condiciones, de las cuales cualquiera puede utilizarse para demostrar que un proceso está fuera de control. La idea detrás de estas condiciones es que si en el diagrama de control persiste una tendencia o patrón durante un gran periodo, puede ser una señal de ausencia de control, incluso si ninguno de los puntos del diagrama está fuera de los límites de control 3σ .

La aplicación de las reglas de Western Electric requiere calcular los límites de control 1σ y 2σ . Los límites de control están dados por $\bar{\bar{X}} \pm A_2\bar{R}/3$, y los límites de control 2σ están dados por $\bar{\bar{X}} \pm 2A_2\bar{R}/3$.

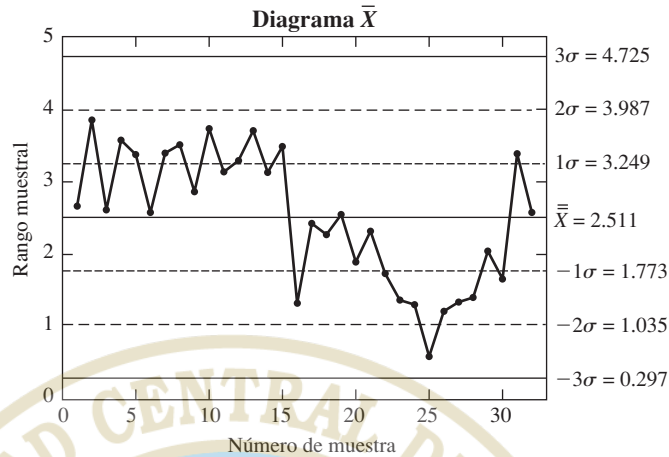


FIGURA 10.9 Este diagrama $\bar{X}$ exhibe patrones no aleatorios, lo cual es muestra de una falta de control estadístico, aun cuando no se encuentren puntos fuera de los límites de control 3σ . En este diagrama se muestran los límites de control 1σ y 2σ , de tal forma que puedan aplicarse las reglas de Western Electric.

Las reglas de Western Electric

Cualquiera de las siguientes condiciones es prueba de que un proceso está fuera de control:

1. Cualquier punto del diagrama fuera de los límites 3σ .
2. Dos de tres puntos consecutivos del diagrama por arriba del límite superior 2σ , o dos de tres puntos consecutivos del diagrama por abajo del límite inferior 2σ .
3. Cuatro de cinco puntos consecutivos del diagrama por arriba del límite superior 1σ , o cuatro de cinco puntos consecutivos del diagrama por debajo del límite inferior 1σ .
4. Ocho puntos consecutivos del diagrama que estén del mismo lado que la recta central.

En la figura 10.9 las reglas de Western Electric indican que un proceso está fuera de control en el número de muestra 8, lugar en el cual cuatro de cinco puntos consecutivos se encuentran por arriba del límite de control superior 1σ . Para mayor información sobre el uso de las reglas de Western Electric para interpretar los diagramas de control, véase Montgomery (2001b).

El diagrama S

El diagrama S es una alternativa al diagrama R . Ambos son útiles para controlar la variabilidad en un proceso. Mientras que el diagrama R evalúa la variabilidad con el rango muestral, el S utiliza la desviación estándar muestral. La figura 10.10 presenta el diagrama de los datos de humedad de la tabla 10.1.

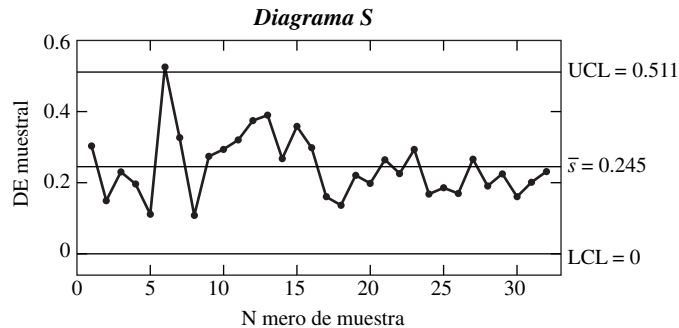


FIGURA 10.10 Diagrama S para los datos de humedad. Compárese con la figura 10.1.

Observe que el diagrama S para los datos de humedad es similar en apariencia al diagrama R (figura 10.1) para los mismos datos. Al igual que el diagrama R, el S indica que la variación estaba fuera de control en la muestra 6.

Para entender dónde deben graficarse los límites de control, suponga que 32 desviaciones estándar provienen de una población con media μ y desviación estándar σ . Idealmente se desearía que en el diagrama la recta central estuviera en μ_s y los límites de control en $\mu_s \pm 3\sigma_s$. Generalmente no se conocen estas cantidades. Se aproxima μ_s con $\bar{s}$, el promedio de las desviaciones estándar. Por tanto, se grafica la recta central en $\bar{s}$. Puede demostrarse mediante métodos avanzados que las cantidades $\mu_s \pm 3\sigma_s$ pueden estimarse con múltiplos de $\bar{s}$; estos múltiplos se representan mediante B_3 y B_4 . La cantidad $\mu_s - 3\sigma_s$ se estima con $B_3\bar{s}$, mientras que la cantidad $\mu_s + 3\sigma_s$ se estima con $B_4\bar{s}$. Las cantidades B_3 y B_4 son constantes cuyos valores dependen del tamaño de la muestra n . A continuación se presenta una pequeña tabla de valores de B_3 y B_4 . La tabla A.9 (Apéndice A) muestra una tabulación más extensa. Observe que para muestras de tamaño 5 o menos, el valor de B_3 es 0. Para muestras así de pequeñas, el valor de $\mu_s - 3\sigma_s$ es negativo. En estos casos se establece que el límite de control es 0, debido a que es imposible que una desviación estándar sea negativa.

n	2	3	4	5	6	7	8
B_3	0	0	0	0	0.030	0.118	0.185
B_4	3.267	2.568	2.266	2.089	1.970	1.882	1.815

Ejemplo
10.7

Calcule la recta central y los límites de control 3σ superior e inferior del diagrama para los datos de humedad de la tabla 10.1.

Solución

El valor de $\bar{s}$ es 0.2445 (tabla 10.1). El tamaño de la muestra es $n = 5$. De la tabla anterior se tiene que $B_3 = 0$ y $B_4 = 2.089$. Por consiguiente, el límite de control superior es $(2.089)(0.2445) = 0.5108$, y el límite de control inferior es $(0)(0.2445) = 0$.

Resumen

En un diagrama S , la recta central y los límites de control 3σ superior e inferior están dados por

$$\begin{aligned}\text{Límite superior } 3\sigma &= B_4\bar{s} \\ \text{Recta central} &= \bar{s} \\ \text{Límite inferior } 3\sigma &= B_3\bar{s}\end{aligned}$$

Los valores B_3 y B_4 dependen del tamaño de la muestra. La tabla A.9 tabula estos valores.

El diagrama S de la figura 10.10 muestra que la variación del proceso se encuentra fuera de control en la muestra 6. Se elimina esta muestra y se vuelve a calcular el diagrama S . La figura 10.11 presenta los resultados. Ahora la variación está bajo control. Observe que este diagrama tiene una apariencia similar al diagrama R de la figura 10.2.

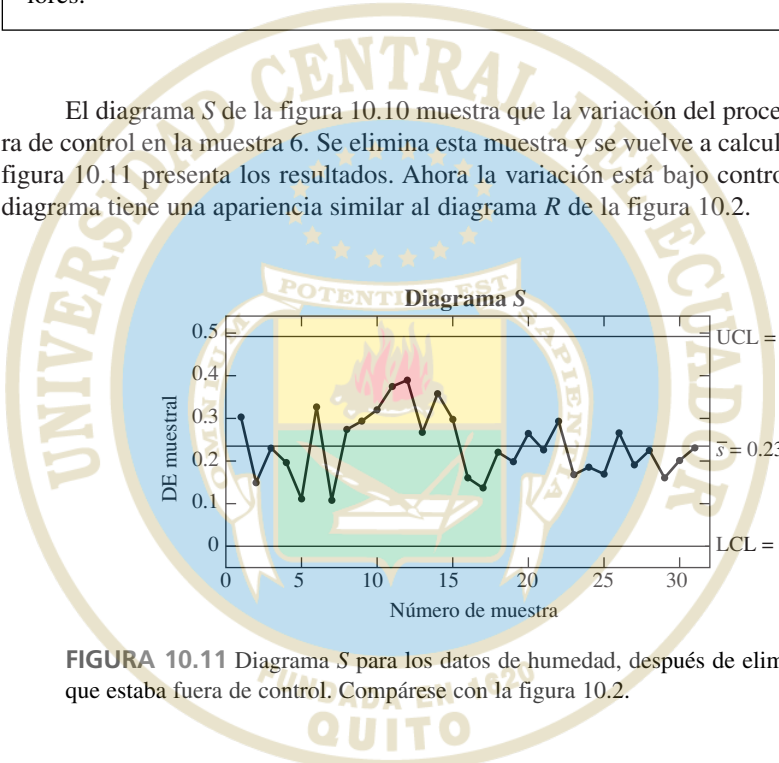


FIGURA 10.11 Diagrama S para los datos de humedad, después de eliminar la muestra que estaba fuera de control. Compárese con la figura 10.2.

Una vez que la variación está en control, se calcula el diagrama $\bar{X}$ para evaluar la media muestral. Recuerde que para el diagrama $\bar{X}$ la recta central se encuentra en $\bar{\bar{X}}$, y que lo ideal sería que los límites de control superior e inferior se localizaran a una distancia $3\sigma_{\bar{X}}$ por debajo y por arriba de la recta central. Puesto que se empleó el diagrama S para evaluar la variación del proceso, se estimará la cantidad $3\sigma_{\bar{X}}$ con un múltiplo de $\bar{s}$. Específicamente, se estima a $3\sigma_{\bar{X}}$ con $A_3\bar{s}$, donde A_3 es una constante cuyo valor depende del tamaño de muestra n . A continuación se muestra una pequeña tabla de valores de A_3 . La tabla A.9 muestra una tabulación más extensa.

n	2	3	4	5	6	7	8
A_3	2.659	1.954	1.628	1.427	1.287	1.182	1.099

Resumen

En un diagrama $\bar{X}$, cuando se utiliza a $\bar{s}$ para estimar $\sigma_{\bar{X}}$, la recta central y los límites 3σ superior e inferior están dados por

$$\text{Límite superior } 3\sigma = \bar{\bar{X}} + A_3\bar{s}$$

$$\text{Recta central} = \bar{\bar{X}}$$

$$\text{Límite inferior } 3\sigma = \bar{\bar{X}} - A_3\bar{s}$$

El valor A_3 depende del tamaño de la muestra. Los valores están tabulados en la tabla A.9.

Si se utilizan las reglas de Western Electric, deben calcularse los límites de control 1σ y 2σ . Los límites 1σ son $\bar{\bar{X}} \pm A_3\bar{s}/3$; los límites 2σ son $\bar{\bar{X}} \pm 2A_3\bar{s}/3$.

Ejemplo 10.8

Calcule los límites de control 3σ superior e inferior del diagrama $\bar{X}$ para los datos de humedad de la tabla 10.1.

Solución

Con la muestra 6 eliminada, el valor de $\bar{\bar{X}}$ es 2.658, y el valor de $\bar{s}$ es 0.2354. El tamaño de la muestra es $n = 5$. De la tabla se tiene que $A_3 = 1.427$. Por tanto, el límite de control superior es $2.658 + (1.427)(0.2354) = 2.994$, y el límite de control inferior es $2.658 - (1.427)(0.2354) = 2.322$.

La figura 10.12 muestra el diagrama $\bar{X}$ de los datos de humedad con la muestra 6 eliminada. Los límites de control se parecen mucho a los calculados a partir de los rangos mues-

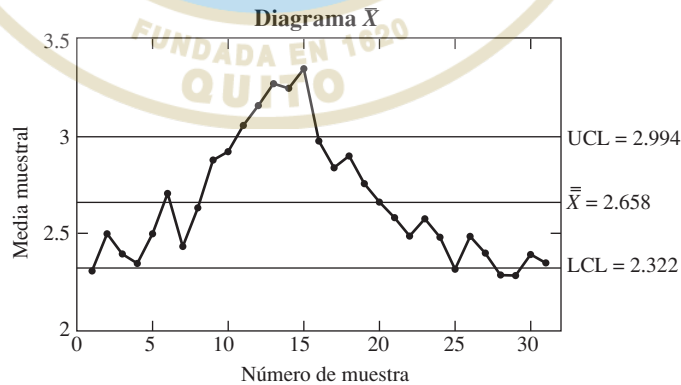


FIGURA 10.12 Diagrama $\bar{X}$ para los datos de humedad. Los límites de control están basados en las desviaciones estándar en vez de los rangos muestrales. Compárese con la figura 10.3.

trales, como lo muestra la figura 10.3. La figura 10.12 indica que el proceso no está en control. Después de corregirse, se construyen nuevos diagramas S y $\bar{X}$. La figura 10.13 presenta los resultados. Ahora, el proceso está en un estado de control estadístico.

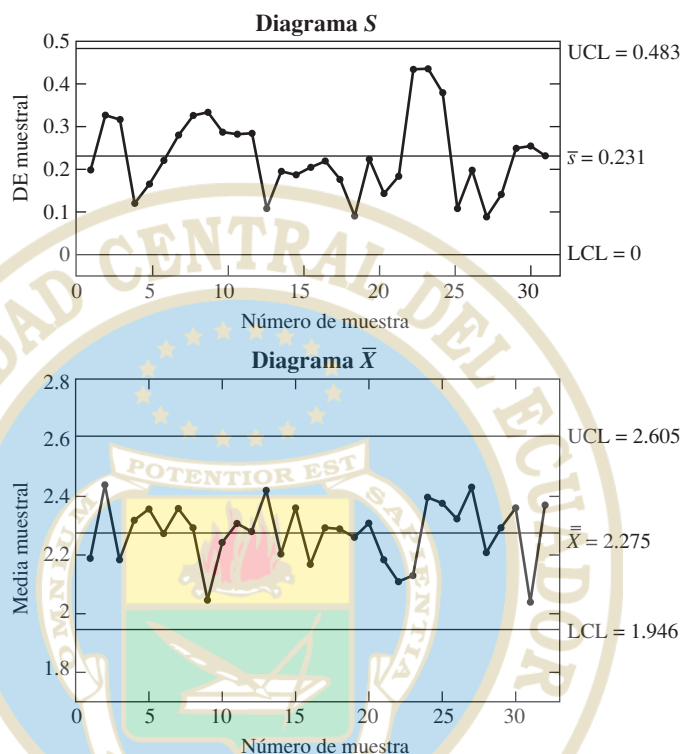


FIGURA 10.12 Diagramas S y $\bar{X}$ después de que se corrigió la causa especial. Ahora el proceso se encuentra en un estado de control estadístico. Compárese con la figura 10.4.

En resumen, el diagrama S es una alternativa al diagrama R que puede utilizarse junto con el diagrama $\bar{X}$. Para los datos de humedad, los dos diagramas produjeron resultados muy similares. Esto es cierto en muchos casos, pero algunas veces los resultados son diferentes.

¿Cuál es mejor, el diagrama S o el diagrama R ?

Ambos diagramas tienen el mismo objetivo: estimar la desviación estándar del proceso y determinar si está en control. Parece ser más natural estimar la desviación estándar del proceso con la desviación estándar muestral s que con el rango R . De hecho, cuando la población es normal, s es una estimación más precisa de la desviación estándar del proceso que R , debido a una incertidumbre menor. Para ver esto de forma intuitiva, note que el cálculo de s abarca

todas las mediciones de cada muestra, mientras que el cálculo de R sólo toma en cuenta dos mediciones (la mayor y la menor). Resulta que el mejoramiento de la precisión obtenido con s en contraposición con R aumenta a medida que el tamaño de muestra es mayor. De ahí que el diagrama S sea una mejor opción, especialmente para tamaños de muestras más grandes (mayores a 5 y en adelante). El diagrama R se sigue utilizando aún, en gran parte por tradición. En determinada época, el diagrama R tenía la ventaja de que podía calcularse el rango muestral con menos aritmética que la desviación estándar muestral. Hoy en que la mayoría de los cálculos se hacen electrónicamente, esta ventaja ya no es válida. Por tanto, el diagrama S es en general la mejor opción.

Muestras de tamaño 1

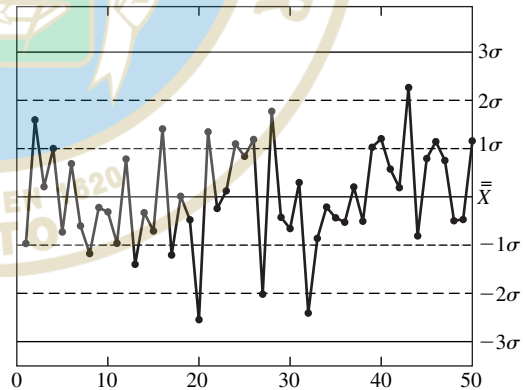
En algunas ocasiones se necesitan definir los grupos racionales de tal manera que cada muestra sólo contenga un valor. Por ejemplo, si la tasa de producción es muy lenta, no sería conveniente esperar a acumular muestras mayores a $n = 1$. Es imposible calcular un rango muestral o una desviación estándar muestral para una muestra de tamaño 1, por lo que no pueden utilizarse los diagramas R y S . Se puede disponer de muchos otros métodos. Uno de ellos es el diagrama CUSUM, que se analiza en la sección 10.4.

Ejercicios para la sección 10.2

1. El plan de control de calidad para cierto proceso de producción consiste en tomar muestras de tamaño 4. Los resultados de las últimas 30 muestras pueden resumirse de la siguiente manera:

$$\sum_{i=1}^{30} \bar{X}_i = 712.5 \quad \sum_{i=1}^{30} R_i = 143.7 \quad \sum_{i=1}^{30} s_i = 62.5$$

- a) Calcule los límites de control 3σ para el diagrama R .
 - b) Calcule los límites de control 3σ para el diagrama S .
 - c) Utilizando los rangos muestrales, calcule los límites de control 3σ para el diagrama $\bar{X}$.
 - d) Utilizando las desviaciones estándar muestrales, calcule los límites de control 3σ para el diagrama $\bar{X}$.
2. El siguiente diagrama $\bar{X}$ presenta las últimas 50 muestras tomadas de la producción de un proceso. Utilizando las reglas de Western Electric, ¿podría detectarse en cualquier momento que el proceso no está en control? Si es así, especifique en qué muestra se detectaría primero que el proceso está fuera de control y qué regla no se cumple.



3. Se mide la distancia (en mm) entre los electrodos centrales y laterales de las bujías de motores de combustión interna en muestras de tamaño 5. La siguiente tabla presenta las medias, los rangos y las desviaciones estándar de 20 muestras consecutivas.

Muestra	$\bar{X}$	R	s
1	1.11	0.17	0.07
2	1.09	0.13	0.06
3	1.10	0.13	0.06
4	1.09	0.22	0.09
5	1.11	0.15	0.06
6	1.07	0.13	0.06
7	1.05	0.11	0.05
8	1.04	0.12	0.06
9	1.05	0.12	0.05
10	1.07	0.05	0.02
11	1.11	0.14	0.07
12	1.06	0.10	0.04
13	1.10	0.14	0.07
14	1.14	0.14	0.05
15	1.10	0.19	0.08
16	1.13	0.09	0.04
17	1.19	0.18	0.07
18	1.20	0.06	0.03
19	1.21	0.11	0.05
20	1.18	0.14	0.06

Las medias son $\bar{\bar{X}} = 1.110$, $\bar{R} = 0.131$ y $\bar{s} = 0.057$.

- Calcule los límites de control 3σ para el diagrama R . ¿La varianza está bajo control? Si no es así, elimine las muestras que no lo están y vuelva a calcular a $\bar{\bar{X}}$ y $\bar{R}$.
 - Con base en el rango muestral R , calcule los límites de control 3σ para el diagrama $\bar{X}$. Con base en los límites 3σ , ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no está en control?
 - Con base en las reglas de Western Electric, ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no está en control?
- Repita el ejercicio 3 utilizando el diagrama S en lugar del R .
 - Un proceso tiene una media de 10 y desviación estándar de 2. Se da seguimiento al proceso tomando muestras de tamaño 4 en intervalos regulares. Se concluye que el proceso no está en control si es que un punto del diagrama $\bar{X}$ se encuentra fuera de los límites de control 3σ .
 - Si la media del proceso pasa a ser 11, ¿cuál es el número de muestras promedio que se extraerán antes de que se detecte el cambio en el diagrama $\bar{X}$?
 - ¿A qué valor necesita darse el cambio positivo para detectarlo con una ARL de 6?
 - Si el tamaño de la muestra sigue siendo de 4, ¿a qué valor debe reducirse la desviación estándar para que produzca una ARL de 6 cuando la media del proceso cambia a 11?
 - Un proceso tiene una media de 7.2 y desviación estándar de 1.3. Se da seguimiento al proceso tomando muestras de tamaño 6 en intervalos regulares. Se concluye que el proceso no está en control si es que un punto del diagrama $\bar{X}$ se encuentra fuera de los límites de control 3σ .
 - Si la media del proceso cambia a 6.7, ¿cuál es el número de muestras promedio que se extraerán antes de que se detecte el cambio en el diagrama $\bar{X}$?
 - ¿A qué valor necesita darse el cambio positivo para detectarlo con una ARL de 15?
 - Si el tamaño de la muestra sigue siendo de 6, ¿a qué valor debe reducirse la desviación estándar para que produzca una ARL de 15 cuando la media del proceso cambia a 6.7?
 - Si la desviación estándar sigue siendo de 1.3, ¿qué tamaño de muestra debe utilizarse para que se produzca una ARL no mayor a 6 cuando la media del proceso cambia a 6.7?
 - Se da seguimiento a un proceso tomando muestras en intervalos regulares y se concluye que está fuera de control si un punto del diagrama está fuera de los límites de control 3σ . Suponga que el proceso está en control.
 - ¿Cuál es la probabilidad de que ocurra una falsa alarma en las próximas 50 muestras?
 - ¿Cuál es la probabilidad de que ocurra una falsa alarma en las próximas 100 muestras?
 - ¿Cuál es la probabilidad de que no haya una falsa alarma en las próximas 200 muestras?
 - Complete el enunciado: La probabilidad es de 0.5 de que haya una falsa alarma en las próximas _____ muestras.
 - Se toman periódicamente muestras de seis cojinetes de bola y se miden sus diámetros (en mm). La siguiente tabla presenta las medias, los rangos y las desviaciones estándar de 25 muestras consecutivas.

Muestra	$\bar{X}$	R	s
1	199.5	7.2	2.38
2	200.4	7.9	2.70
3	202.0	3.6	1.35
4	198.9	5.3	1.95
5	199.1	3.4	1.53
6	200.2	4.8	2.19
7	199.5	2.1	0.87
8	198.1	5.5	2.22
9	200.0	5.5	2.09
10	199.0	9.2	3.17
11	199.3	7.8	2.64
12	199.3	5.8	1.98
13	200.5	3.6	1.36
14	200.1	5.7	2.34
15	200.1	3.0	1.11
16	200.4	5.2	1.99
17	200.9	4.1	1.54
18	200.8	5.5	2.01
19	199.3	8.3	2.92
20	199.8	3.8	1.78
21	199.5	3.6	1.60
22	198.9	7.2	2.60
23	199.6	3.8	1.69
24	200.3	3.1	1.15
25	199.9	4.1	1.86

Las medias son $\bar{\bar{X}} = 199.816$, $\bar{R} = 5.164$ y $\bar{s} = 1.961$.

- Calcule los límites de control 3σ para el diagrama R . ¿La varianza está bajo control? Si no es así, elimine las muestras que no están en control y vuelva a calcular a $\bar{\bar{X}}$ y $\bar{R}$.
- Con base en el rango muestral R , calcule los límites de control 3σ para el diagrama $\bar{X}$. Con base en los límites 3σ , ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no lo está?
- Con base en las reglas de Western Electric, ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no lo está?

9. Repita el ejercicio 8 utilizando el diagrama S en lugar del R .

10. Cierta tipo de circuito integrado está conectado a su marco mediante cuatro alambres. Se toman treinta muestras de cinco unidades cada una, y se mide la resistencia de empuje (en gramos) de un alambre de cada unidad. La tabla E10 presenta los datos. Las medias son $\bar{\bar{X}} = 9.81$, $\bar{R} = 1.14$ y $\bar{s} = 0.4647$.

- Calcule los límites de control 3σ para el diagrama R . ¿La varianza está fuera control en algún punto? Si es así, elimine las muestras que no están en control y vuelva a calcular a $\bar{\bar{X}}$ y a $\bar{R}$.
- Calcule los límites de control 3σ para el diagrama $\bar{X}$. Con base en los límites 3σ , ¿la media del proceso está bajo control? Si no es así, ¿en qué punto se detecta por primera vez que no lo está?
- Con base en las reglas de Western Electric, ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no lo está?

11. Repita el ejercicio 10 utilizando el diagrama S en lugar del R .

12. Los alambres de cobre están revestidos con una fina capa de plástico. Cada hora se toman muestras de cuatro alambres y se mide el espesor del revestimiento (en milésimas). La tabla E12 de la página 746 presenta los datos de las últimas 30 muestras. Las medias son $\bar{\bar{X}} = 150.075$, $\bar{R} = 6.97$ y $\bar{s} = 3.082$.

- Calcule los límites de control 3σ para el diagrama R . ¿La varianza está fuera control en algún punto? Si es así, elimine las muestras que no lo están y vuelva a calcular a $\bar{\bar{X}}$ y a $\bar{R}$.
- Calcule los límites de control 3σ para el diagrama $\bar{X}$. Con base en los límites 3σ , ¿la media del proceso está bajo control? Si no es así, ¿en qué punto se detecta por primera vez que no lo está?
- Con base en las reglas de Western Electric, ¿la media del proceso está bajo control? Si no es así, ¿cuándo se detecta por primera vez que no lo está?

13. Repita el ejercicio 12 utilizando el diagrama S en lugar del R .

TABLA E10 Datos para el ejercicio 10

Muestra	Valores muestrales					$\bar{X}$	R	s
1	10.3	9.8	9.7	9.9	10.2	9.98	0.6	0.26
2	9.9	9.4	10.0	9.4	10.2	9.78	0.8	0.36
3	9.0	9.9	9.6	9.2	10.6	9.66	1.6	0.63
4	10.1	10.6	10.3	9.6	9.7	10.06	1.0	0.42
5	10.8	9.4	9.9	10.1	10.1	10.06	1.4	0.50
6	10.3	10.1	10.0	9.5	9.8	9.94	0.8	0.30
7	8.8	9.3	9.9	8.9	9.3	9.24	1.1	0.43
8	9.4	9.7	9.4	9.9	10.5	9.78	1.1	0.45
9	9.1	8.9	9.8	9.0	9.3	9.22	0.9	0.36
10	8.9	9.4	10.6	9.4	8.7	9.40	1.9	0.74
11	9.0	8.6	9.9	9.6	10.5	9.52	1.9	0.75
12	9.5	9.2	9.4	9.3	9.6	9.40	0.4	0.16
13	9.0	9.4	9.7	9.4	8.6	9.22	1.1	0.43
14	9.4	9.2	9.4	9.3	9.7	9.40	0.5	0.19
15	9.4	10.2	9.0	8.8	10.2	9.52	1.4	0.66
16	9.6	9.5	10.0	9.3	9.4	9.56	0.7	0.27
17	10.2	8.8	10.0	10.1	10.1	9.84	1.4	0.59
18	10.4	9.4	9.9	9.4	9.9	9.80	1.0	0.42
19	11.1	10.5	10.6	9.8	9.4	10.28	1.7	0.68
20	9.3	9.9	10.9	9.5	10.6	10.04	1.6	0.69
21	9.5	10.2	9.7	9.4	10.0	9.76	0.8	0.34
22	10.5	10.5	10.1	9.5	10.3	10.18	1.0	0.41
23	9.8	8.9	9.6	9.8	9.6	9.54	0.9	0.37
24	9.3	9.7	10.3	10.1	9.7	9.82	1.0	0.39
25	10.2	9.6	8.8	9.9	10.2	9.74	1.4	0.58
26	10.8	9.5	10.5	10.5	10.1	10.28	1.3	0.50
27	10.4	9.9	10.1	9.9	10.9	10.24	1.0	0.42
28	11.0	10.8	10.1	9.2	9.9	10.20	1.8	0.72
29	10.3	10.0	10.6	10.0	11.1	10.40	1.1	0.46
30	10.9	10.6	9.9	10.0	10.8	10.44	1.0	0.46

TABLA E12 Datos para el ejercicio 12

Muestra	Valores muestrales				$\bar{X}$	R	s
1	146.0	147.4	151.9	155.2	150.125	9.2	4.22
2	147.1	147.5	151.4	149.4	148.850	4.3	1.97
3	148.7	148.4	149.6	154.1	150.200	5.7	2.65
4	151.3	150.0	152.4	148.2	150.475	4.2	1.81
5	146.4	147.5	152.9	150.3	149.275	6.5	2.92
6	150.2	142.9	152.5	155.5	150.275	12.6	5.37
7	147.8	148.3	145.7	149.7	147.875	4.0	1.66
8	137.1	156.6	147.2	148.9	147.450	19.5	8.02
9	151.1	148.1	145.6	147.6	148.100	5.5	2.27
10	151.3	151.3	142.5	146.2	147.825	8.8	4.29
11	151.3	153.5	150.2	148.7	150.925	4.8	2.02
12	151.9	152.2	149.3	154.2	151.900	4.9	2.01
13	152.8	149.1	148.5	146.9	149.325	5.9	2.50
14	152.9	149.9	151.9	150.4	151.275	3.0	1.38
15	149.0	149.9	153.1	152.8	151.200	4.1	2.06
16	153.9	150.8	153.9	145.0	150.900	8.9	4.20
17	150.4	151.8	151.3	153.0	151.625	2.6	1.08
18	157.2	152.6	148.4	152.6	152.700	8.8	3.59
19	152.7	156.2	146.8	148.7	151.100	9.4	4.20
20	150.2	148.2	149.8	142.1	147.575	8.1	3.75
21	151.0	151.7	148.5	147.0	149.550	4.7	2.19
22	143.8	154.5	154.8	151.6	151.175	11.0	5.12
23	143.0	156.4	149.2	152.2	150.200	13.4	5.64
24	148.8	147.7	147.1	148.2	147.950	1.7	0.72
25	153.8	145.4	149.5	153.4	150.525	8.4	3.93
26	151.6	149.3	155.0	149.0	151.225	6.0	2.77
27	149.4	151.4	154.6	150.0	151.350	5.2	2.32
28	149.8	149.0	146.8	145.7	147.825	4.1	1.90
29	155.8	152.4	150.2	154.8	153.300	5.6	2.51
30	153.9	145.7	150.7	150.4	150.175	8.2	3.38

10.3 Diagramas de control para atributos

El diagrama p

El diagrama p se utiliza cuando la característica de calidad que se medirá en cada unidad sólo toma dos valores; por lo general, “defectuoso” y “no defectuoso”. En cada muestra se calcula la proporción de unidades defectuosas; después se grafican estas proporciones muestrales. Ahora se describirá cómo calcular la recta central y los límites de control.

Sea p la probabilidad de que una unidad dada esté defectuosa. Si el proceso está en control, esta probabilidad es constante a lo largo del tiempo. Sea k el número de muestras. Se supondrá que todas las muestras tienen el mismo tamaño, y éste se representará por n . Sea X_i el número de unidades defectuosas en la i -ésima muestra, y $\hat{p}_i = X_i/n$ la proporción de unidades defectuosas en la i -ésima muestra. Ahora, $X_i \sim \text{Bin}(n, p)$, y si $np > 10$, es aproximadamente cierto que $\hat{p}_i \sim N(p, p(1 - p)/n)$ (véase p. 274). Dado que $\hat{p}_i$ tiene una media $\mu = p$ y una

desviación estándar $\sigma = \sqrt{p(1 - p)/n}$, se tiene que la recta central debe ubicarse en p , y que los límites de control 3σ deben estar en $p \pm 3\sqrt{p(1 - p)/n}$. Usualmente no se conoce a p y se estima con $\bar{p} = \sum_{i=1}^k \hat{p}_i/k$, el promedio de las proporciones muestrales $\hat{p}_i$.

Resumen

En un diagrama p , en el que el número de unidades en cada muestra es n , la recta central y los límites de control 3σ superior e inferior están dados por

Límite superior $3\sigma = \bar{p} + 3\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$

Recta central $= \bar{p}$

Límite inferior $3\sigma = \bar{p} - 3\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$

Estos límites de control serán válidos si $n\bar{p} > 10$.

Estas ideas se ilustran en el ejemplo 10.9.

Ejemplo

10.9

En la producción de placas de silicio se extraen 30 muestras de tamaño 500, y se calcula la proporción de placas defectuosas en cada muestra. La tabla 10.2 muestra los resultados. Calcule la recta central y los límites de control del diagrama p . Dibuje el diagrama. ¿Parece que el proceso esté en control?

TABLA 10.2 Número y proporción de unidades defectuosas, para el ejemplo 10.9

Proporción de unidades defectuosas ($\bar{p}$)			Proporción de unidades defectuosas ($\bar{p}$)		
Muestra	Número de unidades defectuosas		Muestra	Número de unidades defectuosas	
1	17	0.034	16	26	0.052
2	26	0.052	17	19	0.038
3	31	0.062	18	31	0.062
4	25	0.050	19	27	0.054
5	26	0.052	20	24	0.048
6	29	0.058	21	22	0.044
7	36	0.072	22	24	0.048
8	26	0.052	23	30	0.060
9	25	0.050	24	25	0.050
10	21	0.042	25	26	0.052
11	18	0.036	26	28	0.056
12	33	0.066	27	22	0.044
13	29	0.058	28	31	0.062
14	17	0.034	29	18	0.036
15	28	0.056	30	23	0.046

Solución

El promedio de las 30 proporciones muestrales es $\bar{p} = 0.050867$. En consecuencia, la recta central se ubica en 0.050867. Los límites de control se grafican en $0.050867 \pm 3\sqrt{(0.050867)(0.949133)/500}$. Por tanto, el límite de control superior es 0.0803, y el límite de control inferior es 0.0214. La figura 10.4 muestra el diagrama p . Parece que el proceso sí está en control.

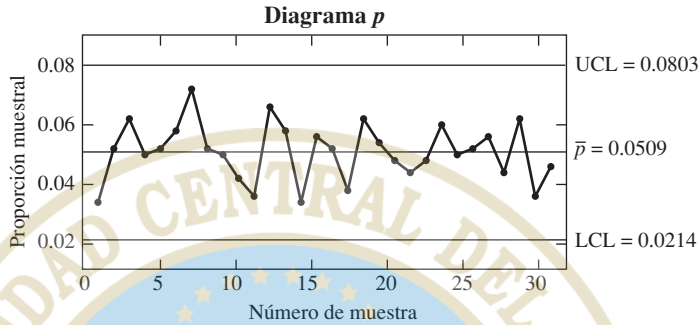


FIGURA 10.4 Diagrama p para los datos de la tabla 10.2.

El tamaño de muestra necesario para construir un diagrama p usualmente es mayor que el requerido para un diagrama $\bar{X}$. La razón es que el tamaño de la muestra debe ser lo suficientemente grande para que haya varias unidades defectuosas en la mayoría de las muestras. Si las unidades defectuosas no se presentan frecuentemente, el tamaño de las muestras debe ser muy grande.

Interpretación de las señales de fuera de control en diagramas de atributos

Cuando se utiliza un diagrama de control de atributos para darle seguimiento a la frecuencia de las unidades defectuosas, un punto del diagrama que está por arriba del límite de control superior requiere de una respuesta muy diferente a un punto del diagrama que está por debajo del límite de control inferior. Ambas situaciones indican que una causa especial ha cambiado la proporción de unidades defectuosas. Un punto del diagrama que está por arriba del límite de control superior señala que la proporción de unidades defectuosas ha aumentado, por lo que debe emprenderse alguna acción para identificar y eliminar la causa especial. Sin embargo, un punto del diagrama que está por debajo del límite de control inferior indica que la causa especial ha reducido la proporción de unidades defectuosas. Aun así debe identificarse la causa especial; en este caso, debe hacerse algo para que ésta siga operando, de tal forma que la proporción de unidades defectuosas pueda disminuir permanentemente.

El diagrama c

El diagrama c se utiliza cuando la medida de calidad es el conteo del número de defectos, o imperfecciones, en una unidad dada. Una *unidad* puede ser un solo elemento o un grupo de elementos lo suficientemente grande para que el número esperado de imperfecciones sea lo bastante grande. El uso del diagrama c requiere que el número de defectos siga una distribución Poisson. Suponga que se extraen k unidades, y que c_i denota el número total de defectos en la i -ésima unidad. Sea que λ denote la media del número total de fallas por unidad. Enton-

ces $c_i \sim \text{Poisson}(\lambda)$. Si el proceso está en control, el valor de λ es constante en el tiempo. Ahora, si λ es razonablemente grande, por ejemplo $\lambda > 10$, entonces aproximadamente $c_i \sim N(\lambda, \lambda)$ (véase p. 278). Observe que el valor de λ puede hacerse grande al escoger un número de elementos por unidad lo suficientemente grande. El diagrama c se construye al graficar los valores de c_i . Puesto que c_i tiene una media λ y una desviación estándar igual $\sqrt{\lambda}$, la recta central debe graficarse en λ y los límites de control deben graficarse en $\lambda \pm 3\sqrt{\lambda}$. Usualmente no se conoce el valor de λ y tiene que estimarse a partir de los datos. La estimación apropiada es $\bar{c} = \sum_{i=1}^k c_i / k$, el número de defectos por unidad.

Resumen

En un diagrama c , la recta central y los límites de control 3σ superior e inferior están dados por

$$\text{Límite superior } 3\sigma = \bar{c} + 3\sqrt{\bar{c}}$$

$$\text{Recta central} = \bar{c}$$

$$\text{Límite inferior } 3\sigma = \bar{c} - 3\sqrt{\bar{c}}$$

Estos límites de control serán válidos si $\bar{c} > 10$.

El ejemplo 10.10 muestra estas ideas.

Ejemplo 10.10

Se examinan las grietas en la superficie de los rollos de láminas de aluminio utilizados para fabricar latas. La tabla 10.3 presenta el número de grietas que hay en 40 muestras de 100 m² cada una. Calcule la recta central y los límites de control 3σ para el diagrama c . Dibuje el diagrama. ¿El proceso parece estar en control?

Solución

El promedio de los 40 conteos es $\bar{c} = 12.275$. Por tanto, la recta central se localiza en 12.275. Los límites de control 3σ se grafican en $12.275 \pm 3\sqrt{12.275}$. En consecuencia, el límite de control superior es 22.7857, y el de control inferior es 1.7643. La figura 10.15 en la página 750 presenta el diagrama c . El proceso parece estar en control.

TABLA 10.3 Número de grietas, para el ejemplo 10.10

Muestra	Número de grietas (c)	Muestra	Número de grietas (c)	Muestra	Número de grietas (c)	Muestra	Número de grietas (c)
1	16	11	14	21	11	31	10
2	12	12	11	22	16	32	10
3	9	13	10	23	16	33	10
4	13	14	9	24	13	34	12
5	15	15	9	25	12	35	14
6	5	16	14	26	17	36	10
7	13	17	10	27	15	37	15
8	11	18	12	28	13	38	12
9	15	19	8	29	15	39	11
10	12	20	14	30	13	40	14

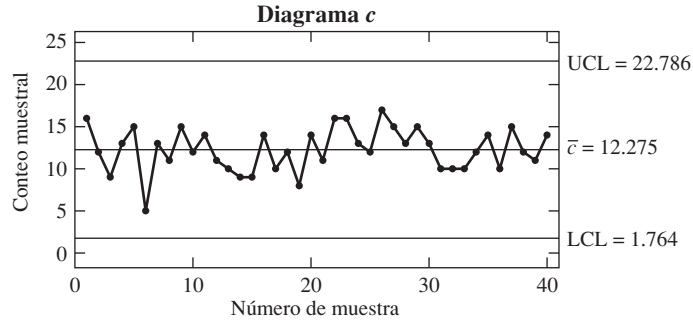


FIGURA 10.15 Diagrama c para los datos de la tabla 10.3.

Ejercicios para la sección 10.3

- Se da seguimiento a un proceso para determinar las unidades defectuosas tomando diariamente una muestra de 200 unidades y calculando la proporción que está defectuosa. Sea p_i la proporción de unidades defectuosas en la i -ésima muestra. Para las últimas 30 muestras, la suma de las proporciones es $\sum_{i=1}^{30} p_i = 1.64$. Calcule la recta central y los límites de control 3σ superior e inferior para un diagrama p .
- Se requiere que el diámetro de cierta cabeza remachadora mida entre 13.3-13.5 mm. Diariamente se toma una muestra de 400 remachadoras, y se cuenta el número de éstas que no cumplen con el requisito. Los números de remachadoras defectuosas para cada uno de los últimos 20 días son los siguientes:

12	25	23	15	20	11	16	20	18	25
22	16	22	18	37	35	40	40	36	30

 - Calcule los límites 3σ superior e inferior para un diagrama p .
 - ¿El proceso está en control? Si no es así, ¿cuándo es la primera vez que se detectó que no lo está?
- Se da seguimiento a un proceso para determinar las unidades defectuosas tomando periódicamente una muestra de 100 unidades y calculando el número de defectuosas. En las últimas 50 muestras hubo un total de 622 unidades defectuosas. ¿Se tiene la suficiente información para calcular los límites de control 3σ para un diagrama p ? Si es así, calcule los límites. Si no, diga qué información adicional se requiere.
- Con referencia al ejercicio 3, en las últimas 50 muestras hubo 622 unidades defectuosas. El número máximo de éstas en las muestras fue de 24, mientras que el mínimo fue de 6.
 - ¿Esta información es suficiente para determinar si el proceso se encontraba fuera de control en cualquier momento durante las últimas 50 muestras? Si es así, diga si el proceso estaba o no fuera de control. Si no, diga qué información adicional se requiere para hacer la determinación.
- Un programa de control de calidad apenas diseñado para cierto proceso consiste en tomar muestras de 20 unidades al día y contar el número de unidades defectuosas. El número éstas en las primeras 10 muestras son 0, 0, 1, 0, 1, 0, 0, 0, 1, 0. Un miembro del grupo de control de calidad pide consejo dada su preocupación de que el número de unidades defectuosas sea muy pequeño para construir un diagrama p preciso. ¿Cuál de los siguientes consejos es el mejor?
 - No necesita cambiarse nada. Un diagrama p preciso puede construirse cuando el número de unidades defectuosas es así de pequeño.
 - Dado que la proporción de unidades que están defectuosas es tan pequeña, no se necesita construir un diagrama p para este proceso.
 - Aumente el valor de p para incrementar el número de unidades defectuosas por muestra.
 - Aumente el tamaño de las muestras para incrementar el número de unidades defectuosas por muestra.
- Se da seguimiento a un proceso que produce latas de aluminio al tomar muestras de mil latas y contar el número total de grietas visuales que hay en todas las latas de la muestra. Sea c_i el número total de grietas en las latas en la i -ésima muestra. Para las últimas 50 muestras se ha calculado la cantidad $\sum_{i=1}^{50} c_i = 1476$. Calcule la recta central y los límites de control 3σ superior e inferior para un diagrama c .

7. Con referencia al ejercicio 6, el número de grietas en la vigésima muestra fue 48. ¿Se puede determinar si el proceso estaba en control en dicho momento? Si es así, establezca si el proceso estaba o no en control. Si no, diga qué información adicional se necesita para hacer dicha determinación.
8. Cada hora se inspecciona una sección 10 m^2 de un tejido para determinar sus grietas. Los números de grietas observadas durante las últimas 20 horas son:
- | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|
| 38 | 35 | 35 | 49 | 33 | 48 | 40 | 47 | 45 | 46 |
| 41 | 53 | 36 | 41 | 51 | 63 | 35 | 58 | 55 | 57 |
- a) Calcule los límites 3σ superior e inferior para un diagrama p .
- b) ¿El proceso está en control? Si no es así, ¿cuándo es la primera vez que se detectó que no lo está?

10.4 El diagrama CUSUM

Uno de los objetivos de un diagrama $\bar{X}$ es detectar un cambio en la media del proceso. No obstante, a menos que un cambio sea bastante grande, puede transcurrir cierto tiempo antes de que un punto del diagrama se ubique fuera de los límites de control 3σ . El ejemplo 10.4 (en la sección 10.2) demostró que cuando la media de un proceso cambia por una cantidad igual a $\sigma_{\bar{X}}$, la longitud de corrida promedio (ARL) es aproximadamente de 44, lo cual significa que en promedio deben observarse 44 muestras antes de que se determine que el proceso está fuera de control. Las reglas de Western Electric (sección 10.2) son un método para reducir la ARL; el diagrama CUSUM (*cumulative sums*) es otro.

Una forma en que se manifiestan los cambios por sí mismos es con una corrida de puntos por arriba o debajo de la recta central. Las reglas de Western Electric están diseñadas para responder a las corridas. Otra forma de detectar pequeños cambios es con **sumas acumuladas**. Imagine que la media de un proceso tiene un ligero cambio positivo. Entonces los puntos tendrán cierta tendencia a estar por arriba de la recta central. Si se suman las desviaciones desde la recta central en adelante, y se grafican las sumas acumuladas, los puntos se desplazarán hacia arriba y excederán un límite control mucho antes de lo que lo harían en un diagrama $\bar{X}$.

Ahora se describe cómo graficar los puntos en un diagrama CUSUM. Se supone que se tienen m muestras de tamaño n , con medias muestrales $\bar{X}_1, \dots, \bar{X}_m$. Para comenzar, debe especificarse un valor objetivo μ para la media del proceso. Con frecuencia, el valor de μ se toma igual al valor $\bar{\bar{X}}$. Por lo que se necesita una estimación de $\sigma_{\bar{X}}$, la desviación estándar de las medias muestrales. Ésta puede obtenerse ya sea con rangos muestrales, utilizando la estimación $\sigma_{\bar{X}} \approx A_2 \bar{R}/3$, o con desviaciones estándar muestrales, empleando la estimación $\sigma_{\bar{X}} \approx A_3 \bar{s}/3$. Si sólo hay una unidad por muestra ($n = 1$), entonces se necesita de una estimación externa. Aunque una estimación aproximada puede producir buenos resultados, el procedimiento CUSUM también puede ser útil cuando $n = 1$. Finalmente, deben especificarse dos constantes, de manera usual denominadas k y h . Los valores grandes de estas constantes resultan en longitudes de corridas promedio más largas y, por tanto, en menos falsas alarmas, pero también resultan en mayores tiempos de espera para descubrir que un proceso está fuera de control. A menudo se utilizan los valores $k = 0.5$ y $h = 4$ o 5 , ya que resultan en una ARL razonablemente larga cuando el proceso está en control, aún así siguen teniendo buen poder de detectar un cambio de magnitud $1\sigma_{\bar{X}}$ o mayor en la media del proceso.

Para cada muestra, la cantidad $\bar{X}_i - \mu$ es la desviación del valor objetivo. Se definen dos sumas acumuladas, SH y SL. La suma SH siempre es positiva o cero e indica que la media del proceso se ha hecho mayor que el valor objetivo. La suma SL siempre es negativa o cero e indica que la media del proceso se ha hecho menor al valor objetivo. Ambas sumas se

calculan de forma recursiva; es decir, se utiliza el valor actual en la secuencia para calcular el siguiente valor. Los valores iniciales de SH y SL son

$$SH_0 = 0 \quad SL_0 = 0 \quad (10.2)$$

Para $i \geq 1$, los valores son

$$SH_i = \max[0, \bar{X}_i - \mu - k\sigma_{\bar{X}} + SH_{i-1}] \quad (10.3)$$

$$SL_i = \min[0, \bar{X}_i - \mu + k\sigma_{\bar{X}} + SL_{i-1}] \quad (10.4)$$

Si $SH_i > h\sigma_{\bar{X}}$ para alguna i , se concluye que la media del proceso es mayor que el valor objetivo. Si $SL_i < -h\sigma_{\bar{X}}$ para alguna i , se concluye que la media del proceso es menor que el valor objetivo.

La figura 10.16 presenta un diagrama CUSUM para los datos de la figura 10.9 (sección 10.2). Se utilizaron los valores $k = 0.5$ y $h = 4$. El valor 2.962 es la cantidad $h\sigma_{\bar{X}} = 4(0.738)$. El diagrama CUSUM muestra que en la décima muestra hay una situación fuera de control. Para estos datos, el diagrama CUSUM funciona tan bien como las reglas de Western Electric, las cuales determinaron que el proceso se encontraba fuera de control en la octava muestra (véase la figura 10.9).

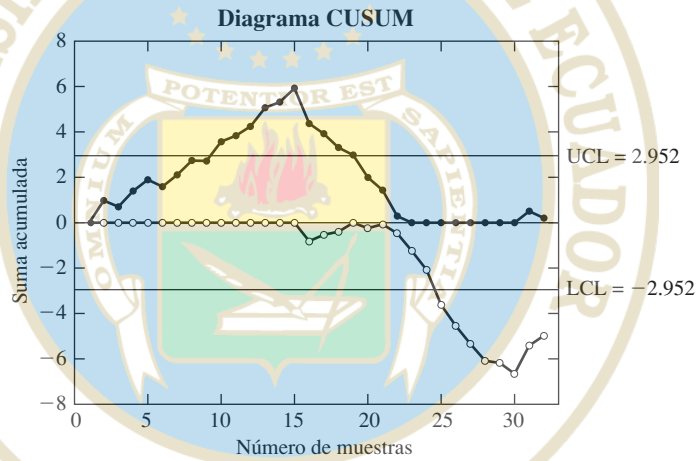


FIGURA 10.16 Diagrama CUSUM para los datos de la figura 10.9.

Resumen

En un diagrama CUSUM se grafican dos sumas acumuladas, SH y SL.

Los valores iniciales son $SH_0 = SL_0 = 0$. Para $i \geq 1$,

$$SH_i = \max[0, \bar{X}_i - \mu - k\sigma_{\bar{X}} + SH_{i-1}]$$

$$SL_i = \min[0, \bar{X}_i - \mu + k\sigma_{\bar{X}} + SL_{i-1}]$$

Deben especificarse las constantes k y h . Con frecuencia se obtienen buenos resultados con los valores $k = 0.5$ y $h = 4$ o 5.

Si para cualquier i , $SH_i > h\sigma_{\bar{X}}$ o $SL_i < -h\sigma_{\bar{X}}$, se dice que el proceso está fuera de control.

Existen algunos otros métodos para construir diagramas CUSUM, los cuales son equivalentes, o casi equivalentes, al método presentado aquí. Alguna gente define que las desviaciones son los puntajes $z z_i = (\bar{X}_i - \mu)/\sigma_{\bar{X}}$, y después se utiliza z_i en lugar de $X_i - \mu$, y k en lugar de $k\sigma_{\bar{X}}$ en las fórmulas de SH y SL. Con esta definición, los límites de control se grafican en $\pm h$ en vez de en $\pm h\sigma_{\bar{X}}$.

También existen otros métodos para dibujar el diagrama CUSUM. La alternativa más común es el enfoque “máscara V”. Para mayor información, puede consultarse un libro sobre control estadístico de calidad, como Montgomery (2000).

Ejercicios para la sección 10.4

1. Considérese el ejercicio 3 de la sección 10.2.

- Elimine todas las muestras que sean necesarias para que la variación del proceso esté bajo control. (Usted ya realizó esto último si resolvió el ejercicio 3 en la sección 10.2.)
- Utilice $\bar{R}$ para estimar $\sigma_{\bar{X}}$ ($\sigma_{\bar{X}}$ es la diferencia entre $\bar{\bar{X}}$ y el límite de control 1σ en un diagrama $\bar{X}$).
- Construya un diagrama CUSUM utilizando $\bar{\bar{X}}$ para la media objetivo, μ , y la estimación de $\sigma_{\bar{X}}$ encontrada en el inciso b) para la desviación estándar. Use los valores $k = 0.5$ y $h = 4$.
- ¿La media del proceso está en control? Si no es así, ¿cuándo se detecta por primera que no lo está?
- Construya un diagrama $\bar{X}$ y emplee las reglas de Western Electric para determinar si la media del proceso está en control. (Usted ya lo hizo en el ejercicio 3 de la sección 10.2.) ¿Las reglas de Western Electric dan los mismos resultados que el diagrama CUSUM? Si no, ¿en qué difieren?

2. Considere el ejercicio 8 de la sección 10.2.

- Elimine todas las muestras que sean necesarias para que la variación del proceso esté bajo control. (Usted ya realizó esto último si resolvió el ejercicio 8 en la sección 10.2.)
- Utilice $\bar{R}$ para estimar $\sigma_{\bar{X}}$ ($\sigma_{\bar{X}}$ es la diferencia entre $\bar{\bar{X}}$ y el límite de control 1σ en un diagrama $\bar{X}$).
- Construya un diagrama CUSUM utilizando $\bar{\bar{X}}$ para la media objetivo, μ , y la estimación de $\sigma_{\bar{X}}$ encontrada en el inciso b) para la desviación estándar. Use los valores $k = 0.5$ y $h = 4$.
- ¿La media del proceso está en control? Si no es así, ¿cuándo se detecta por primera que no lo está?

- Construya un diagrama $\bar{X}$ y emplee las reglas de Western Electric para determinar si la media del proceso está en control. (Usted ya hizo esto si resolvió el ejercicio 8 de la sección 10.2.) ¿Las reglas de Western Electric dan los mismos resultados que el diagrama CUSUM? Si no es así, ¿en qué difieren?

3. Considere el ejercicio 10 de la sección 10.2.

- Elimine todas las muestras que sean necesarias para que la variación del proceso esté bajo control. (Usted ya realizó esto si resolvió el ejercicio 8 en la sección 10.2.)
- Utilice $\bar{R}$ para estimar $\sigma_{\bar{X}}$ ($\sigma_{\bar{X}}$ es la diferencia entre $\bar{\bar{X}}$ y el límite de control 1σ en un diagrama $\bar{X}$).
- Construya un diagrama CUSUM utilizando $\bar{\bar{X}}$ para la media objetivo, μ , y la estimación de $\sigma_{\bar{X}}$ encontrada en el inciso b) para la desviación estándar. Use los valores $k = 0.5$ y $h = 4$.
- ¿La media del proceso está en control? Si no es así, ¿cuándo se detecta por primera que no lo está?
- Construya un diagrama $\bar{X}$ y emplee las reglas de Western Electric para determinar si la media del proceso está en control. (Usted ya hizo esto si resolvió el ejercicio 10 de la sección 10.2.) ¿Las reglas de Western Electric producen los mismos resultados que el diagrama CUSUM? Si no es así, ¿en qué difieren?

4. Considere el ejercicio 3 de la sección 10.2.

- Elimine todas las muestras que sean necesarias para que la variación del proceso esté bajo control. (Usted ya realizó esto si resolvió el ejercicio 12 en la sección 10.2.)
- Utilice $\bar{R}$ para estimar $\sigma_{\bar{X}}$ ($\sigma_{\bar{X}}$ es la diferencia entre $\bar{\bar{X}}$ y el límite de control 1σ en un diagrama $\bar{X}$).
- Construya un diagrama CUSUM utilizando $\bar{\bar{X}}$ para la media objetivo, μ , y la estimación de $\sigma_{\bar{X}}$ encontrada en

el inciso b) para la desviación estándar. Use los valores $k = 0.5$ y $h = 4$.

- d) ¿La media del proceso está en control? Si no es así, ¿cuándo se detecta por primera que no lo está?
- e) Construya un diagrama $\bar{X}$ y emplee las reglas de Western Electric para determinar si la media del proceso está en control. (Usted ya hizo esto si resolvió el ejercicio 12 de la sección 10.2.) ¿Las reglas de Western Electric dan los mismos resultados que el diagrama CUSUM? Si no es así, ¿en qué difieren?
5. Se supone que los bloques de concreto utilizados en cierta aplicación tienen una fuerza de compresión promedio de 1 500 MPa. Se usan muestras de tamaño 1 para el control de calidad. Las fuerzas de compresión de las últimas 40 muestras están dadas en la siguiente tabla.

Muestras	Fuerza	Muestras	Fuerza
1	1487	21	1507
2	1463	22	1474
3	1499	23	1515
4	1502	24	1533
5	1473	25	1487
6	1520	26	1518
7	1520	27	1526
8	1495	28	1469
9	1503	29	1472
10	1499	30	1512
11	1497	31	1483
12	1516	32	1505
13	1489	33	1507
14	1545	34	1505
15	1498	35	1517
16	1503	36	1504
17	1522	37	1515
18	1502	38	1467
19	1499	39	1491
20	1484	40	1488

Resultados anteriores sugieren que un valor de $\sigma = 15$ es razonable para este proceso.

- a) Construya un diagrama CUSUM utilizando el valor de 1500 para la media objetivo μ , y los valores $k = 0.5$ y $h = 4$.
- b) ¿La media del proceso está en control? Si no es así, ¿cuándo se detecta por primera vez que no lo está?
6. Un aprendiz de control de calidad está construyendo un diagrama CUSUM. Los valores calculados para SL y SH se presentan en la siguiente tabla. Tres de los valores no se calcularon correctamente. ¿Cuáles son éstos?

Muestra	SL	SH
1	0	0
2	0	0
3	0	0
4	-1.3280	0
5	-1.4364	0
6	-2.0464	0
7	-1.6370	0
8	-0.8234	0.2767
9	-0.4528	0.1106
10	0	0.7836
11	0.2371	0.0097
12	0.7104	0
13	0	0.2775
14	0	0.5842
15	0	0.3750
16	0	0.4658
17	0	0.1866
18	0	0.3277
19	-0.2036	0
20	0	-0.7345

10.5 Capacidad del proceso

Una vez que un proceso se encuentra en un estado de control estadístico es importante evaluar su habilidad para producir elementos que se apeguen a las especificaciones del diseño. Se toman en cuenta los datos de variables y se supone que la característica relevante de calidad sigue una distribución normal.

El primer paso al evaluar la capacidad del proceso es estimar la media y su desviación estándar. Estas estimaciones se denotan mediante $\hat{\mu}$ y $\hat{\sigma}$, respectivamente. De manera usual, los datos utilizados para calcular $\hat{\mu}$ y $\hat{\sigma}$ se extraen de los diagramas de control justo cuando el proceso está en un estado de control. La media del proceso se estima con $\hat{\mu} = \bar{\bar{X}}$. La desviación estándar del proceso se puede estimar mediante el rango muestral promedio $\bar{R}$ o la desviación estándar muestral promedio $\bar{s}$. Específicamente, se ha demostrado que $\hat{\sigma}$ puede calcularse al dividir $\bar{R}$ entre una constante llamada d_2 , o al dividir $\bar{s}$ entre una constante llamada c_4 . Los valores de las constantes d_2 y c_4 dependen del tamaño de la muestra. Los valores están tabulados en la tabla A.9 (en el Apéndice A).

Resumen

Si una característica de calidad de un proceso que se encuentra en un estado de control sigue una distribución normal, entonces la media $\hat{\mu}$ y la desviación estándar del proceso $\hat{\sigma}$ pueden estimarse a partir de los datos del diagrama de control, de la siguiente manera:

$$\hat{\mu} = \bar{\bar{X}}$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \quad \text{o} \quad \hat{\sigma} = \frac{\bar{s}}{c_4}$$

Los valores de las constantes d_2 y c_4 dependen del tamaño de la muestra. Los valores están tabulados en la tabla A.9.

Observe que la desviación estándar del proceso σ no es la misma cantidad que se utilizó para calcular los límites de control 3σ en el diagrama $\bar{X}$. Los límites de control son $\mu \pm 3\sigma_{\bar{X}}$, donde $\sigma_{\bar{X}}$ es la desviación estándar de la media muestral. La desviación estándar del proceso σ es la desviación de la característica de calidad de las unidades individuales. Éstas se relacionan entre sí por medio de $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, donde n es el tamaño de la muestra.

Para que pueda utilizarse, una característica de calidad debe encontrarse entre un límite de especificación menor (LSL, por sus siglas en inglés) y un límite de especificación superior (USL, por sus siglas en inglés). En algunas ocasiones sólo hay un límite; esta situación se analizará al final de esta sección. Los límites de especificación los determinan los requerimientos del proceso. En los diagramas de control *no* hay límites de control. Se supondrá que la media del proceso está entre LSL y USL.

Se analizarán dos índices de la capacidad del proceso, C_{pk} y C_p . El índice C_{pk} describe la capacidad del proceso tal como ésta es, mientras que C_p describe la capacidad potencial del proceso. Observe que el índice de capacidad del proceso C_p no tiene relación alguna con la cantidad denominada C_p de Mallows que se utiliza para la selección de modelos lineales (véase el capítulo 8). Es una coincidencia que las dos cantidades tengan el mismo nombre.

Se define al índice C_{pk} como la distancia desde $\hat{\mu}$ hasta el límite de especificación más cercano, dividida entre $3\hat{\sigma}$. La figura 10.17 presenta una ilustración en donde $\hat{\mu}$ se acerca más al límite de especificación superior.

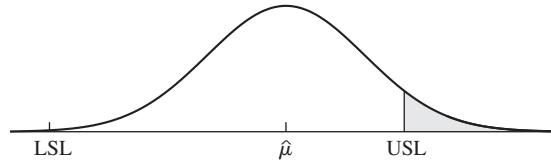


FIGURA 10.17 La curva normal representa la población de unidades producidas por un proceso. La media del proceso se acerca más al límite de especificación superior (USL) que al límite de especificación inferior (LSL). Por tanto, el índice C_{pk} es igual a $(USL - \hat{\mu})/3\hat{\sigma}$.

Definición

El índice C_{pk} es igual a

$$\frac{\hat{\mu} - LSL}{3\hat{\sigma}} \quad \text{o} \quad \frac{USL - \hat{\mu}}{3\hat{\sigma}}$$

cualquiera que sea menor.

Convencionalmente, el valor mínimo aceptable para C_{pk} es 1. Es decir, se considera que un proceso es mínimamente capaz si la media del proceso se encuentra a tres desviaciones estándar del límite de especificación más cercano. En general, se considera que un valor C_{pk} de 1.33, el cual indica que la media del proceso está a cuatro desviaciones estándar del límite de especificación más cercano, es bueno.

Ejemplo

10.11

Las especificaciones de diseño de una barra de un pistón utilizada en una transmisión automática requieren que la longitud de la barra esté entre 71.4 y 72.8 mm. El proceso se vigila con un diagrama $\bar{X}$ y un diagrama S , que utilizan muestras de tamaño $n = 5$. Éstas demuestran que el proceso está en control. Los valores de $\bar{X}$ y $\bar{s}$ son $\bar{X} = 71.8$ mm y $\bar{s} = 0.20$ mm. Calcule el valor de C_{pk} . ¿La capacidad del proceso es aceptable?

Solución

Se estima que $\hat{\mu} = \bar{X} = 71.8$. Para calcular $\hat{\sigma}$, se tiene, a partir de la tabla A.9, que $c_4 = 0.9400$ cuando el tamaño de la muestra es 5. Por consiguiente, $\hat{\sigma} = \bar{s}/c_4 = 0.20/0.9400 = 0.2128$. Los límites de especificación son $LSL = 71.4$ mm y $USL = 72.8$ mm. El valor $\hat{\mu}$ es más cercano al LSL que al USL. En consecuencia,

$$\begin{aligned} C_{pk} &= \frac{\hat{\mu} - LSL}{3\hat{\sigma}} = \frac{71.8 - 71.4}{(3)(0.2128)} \\ &= 0.6266 \end{aligned}$$

Dado que $C_{pk} < 1$, la capacidad del proceso no es aceptable.

Ejemplo

10.12

Con referencia al ejemplo 10.11, suponga que se puede ajustar la media del proceso a cualquier valor deseado. ¿Qué valor debe establecerse para maximizar el valor C_{pk} ? ¿Cuál será el valor de C_{pk} ?

Solución

Los límites de especificación son $LSL = 71.4$ y $USL = 72.8$. El valor de C_{pk} se maximizará si se ajusta a la media del proceso en el punto medio entre los límites de especificación; esto es, si $\mu = 72.1$. La desviación estándar del proceso se estima con $\hat{\sigma} = 0.2128$. Por tanto, el valor máximo de C_{pk} es $(72.1 - 71.4)/(3)(0.2128) = 1.0965$. La capacidad del proceso sería aceptable.

A la capacidad que puede alcanzarse potencialmente al cambiar la media del proceso al punto medio entre los límites de especificación superior e inferior se le llama **índice de capacidad del proceso**, que se denota por C_p . Si la media del proceso se encuentra en el punto medio entre LSL y USL , entonces la distancia desde de la media hasta cualquier límite de especificación es igual a la mitad de la distancia entre los límites de especificación, esto es $\mu - LSL = USL - \mu = (USL - LSL)/2$ (véase la figura 10.18). De ahí que

$$C_p = \frac{USL - LSL}{6\hat{\sigma}} \quad (10.5)$$

El índice de capacidad del proceso C_p mide la capacidad potencial del proceso; es decir, la capacidad máxima que puede alcanzar el proceso sin reducir la desviación estándar de éste.

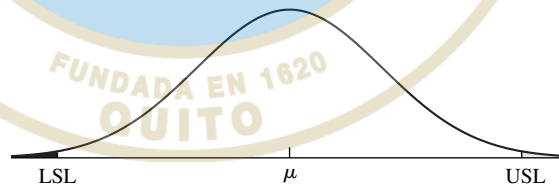


FIGURA 10.18 Un proceso alcanza su capacidad máxima cuando la media del proceso se encuentra en el punto medio entre los límites de especificación. En este caso $\mu - LSL = USL - \mu = (USL - LSL)/2$.

Ejemplo

10.13

Cierto circuito eléctrico requiere que el voltaje de salida esté entre 48 y 52 V. El proceso está en control con $\hat{\sigma} = 0.482$ V. Calcule el índice de capacidad del proceso C_p .

Solución

$$C_p = \frac{USL - LSL}{6\hat{\sigma}} = \frac{52 - 48}{(6)(0.482)} = 1.38$$

La capacidad del proceso es potencialmente buena.

Estimación de la proporción de unidades que no se ajustan a las especificaciones a partir de la capacidad del proceso

Mucha gente utiliza el valor de C_p para tratar de estimar la proporción de unidades que no siguen las especificaciones. Por ejemplo, si $C_p = 1$, entonces los límites de especificación son iguales a $\hat{\mu} - 3\hat{\sigma}$ y $\hat{\mu} + 3\hat{\sigma}$, respectivamente. Por tanto, una unidad no seguirá las especificaciones sólo si se encuentra a más de tres desviaciones estándar de la media del proceso. Ahora, para una población normal la proporción de unidades que se encuentran a más de tres desviaciones estándar de la media es igual a 0.0027. Por consiguiente, con frecuencia se afirma que un proceso con $C_p = 1$ producirá 27 partes que no seguirán las especificaciones de 10 000.

El problema con esto es que el supuesto de normalidad sólo es aproximado para los procesos reales. La aproximación puede ser muy buena cerca de la mitad de la curva, pero con frecuencia no es buena en las colas. Por tanto, la proporción verdadera de las partes que no siguen las especificaciones puede ser muy diferente de la predicha a partir de la curva normal, especialmente cuando la proporción es muy pequeña. En general, las estimaciones de las probabilidades pequeñas que están basadas en una aproximación normal son en el mejor de los casos extremadamente burdas.

Calidad six-sigma

La frase “calidad six-sigma” se ha convertido en un término muy común en los estudios de calidad de control de los últimos años. Se dice que un proceso tiene una calidad six-sigma si el índice de la capacidad del proceso C_p tiene un valor mayor o igual a 2.0. De forma equivalente, un proceso tiene una calidad six-sigma si la diferencia $USL - LSL$ es de al menos 12σ . Cuando un proceso tiene una calidad six-sigma, entonces la media del proceso se ajusta óptimamente y se encuentra a seis desviaciones estándar de cada límite de especificación. En este caso, la proporción de unidades que no siguen las especificaciones es virtualmente igual a cero.

Una característica importante de un proceso six-sigma es que puede resistir cambios moderados en la media del proceso sin que haya un deterioro importante en la capacidad. Por ejemplo, aunque la media del proceso cambiara en 3σ en una dirección u otra, aún se encontraría a 3σ del límite de especificación más cercano, por lo que el índice de capacidad seguiría siendo aceptable.

Ejemplo 10.14

Considere el ejemplo 10.13, ¿a qué valor debe reducirse la desviación estándar del proceso para que el proceso alcance una calidad six-sigma?

Solución

Para alcanzar una calidad six-sigma, el valor de C_p debe ser al menos de 2.0. El valor de σ en el que esto último ocurre se encuentra al establecer que $C_p = 2.0$ y despejando σ . Se obtiene

$$2.0 = \frac{52 - 48}{6\sigma}$$

a partir del cual $\sigma = 0.33$.

Tolerancias unilaterales

Algunas características tienen sólo un límite de especificación. Por ejemplo, los esfuerzos tienen usualmente un límite de especificación inferior pero no un límite superior, ya que la mayoría de las aplicaciones de una parte no puede ser demasiado fuerte. El equivalente de C_{pk} cuando sólo hay un límite inferior es el **índice de capacidad inferior** C_{pl} ; cuando sólo hay un límite superior se refiere al **índice de capacidad superior** C_{pu} . Cada una de estas cantidades es la diferencia entre la media del proceso $\hat{\mu}$ y el límite de especificación, dividida entre $3\hat{\sigma}$.

Resumen

Si un proceso sólo tiene un límite de especificación inferior (LSL), entonces el índice de capacidad inferior es

$$C_{pl} = \frac{\hat{\mu} - \text{LSL}}{3\hat{\sigma}}$$

Si un proceso sólo tiene un límite de especificación superior (USL), entonces el índice de capacidad superior es

$$C_{pu} = \frac{\text{USL} - \hat{\mu}}{3\hat{\sigma}}$$

No hay ningún equivalente de C_p para los procesos que tienen sólo un límite de especificación.

Ejercicios para la sección 10.5

- Las especificaciones de diseño para la válvula de entrada de un motor de combustión interna requieren que la apertura de la válvula mida entre 0.18 y 0.22 mm. Con base en muestras de tamaño 4, los datos de un diagrama $\bar{X}$, que demuestran que el proceso está en control, producen valores de $\bar{\bar{X}} = 0.205$ mm y $\bar{s} = 0.002$ mm.
 - Calcule el valor de C_{pk} para este proceso.
 - ¿La capacidad del proceso es aceptable? Explique.
- Las especificaciones para el volumen de llenado de latas de bebidas se encuentran entre 11.95 y 12.10 onzas. Con base en muestras de tamaño 5, los datos de un diagrama $\bar{X}$, que demuestran que el proceso está en control, producen valores de $\bar{\bar{X}} = 12.01$ oz y $\bar{R} = 0.124$ oz.
 - Calcule el valor de C_{pk} para este proceso.
 - ¿La capacidad del proceso es aceptable? Explique.

3. Considérese el ejercicio 1.

- ¿A qué valor debe establecerse la media del proceso para que la capacidad del proceso sea máxima?
- ¿Cuál sería entonces la capacidad del proceso?

4. Considérese el ejercicio 2.

- ¿A qué valor debe establecerse la media del proceso para que la capacidad del proceso sea máxima?
- ¿Es posible convertir en aceptable la capacidad del proceso simplemente gastando la media del mismo? Explique.
- Cuando la media del proceso se encuentra en su valor óptimo, ¿qué valor debe tener la desviación estándar del proceso para que la capacidad del proceso sea aceptable?
- Cuando la media del proceso se encuentra en su valor óptimo, ¿qué valor debe tener la desviación estándar del

proceso para que la capacidad del proceso tenga una calidad six-sigma?

5. Un proceso tiene un índice de capacidad del proceso de $C_p = 1.2$.

- Suponga que la media del proceso está en su valor óptimo. Expresé los límites de especificación superior e inferior en términos de la media y la desviación estándar del proceso.
- Utilizando la curva normal, estime la proporción de unidades que no seguirán la especificación.
- ¿Existe o no la posibilidad de que la verdadera proporción de unidades que no siguen la especificación sea muy diferente a la estimación del inciso b)? Explique.

Ejercicios adicionales para el capítulo 10

1. Se da seguimiento a un proceso para encontrar el número de unidades defectuosas tomando diariamente una muestra de 250 unidades y calculando la proporción que está defectuosa. Sea p_i la proporción de unidades defectuosas en la i -ésima muestra. Para las últimas 50 muestras, la suma de las proporciones es $\sum_{i=1}^{50} p_i = 2.98$. Calcule la recta central y los límites de control 3σ superior e inferior para un diagrama p .

2. Alguien construye un diagrama $\bar{X}$ en que los límites de control se encuentran en $\pm 2.5\sigma_{\bar{X}}$ en vez de $\pm 3\sigma_{\bar{X}}$.

- Si el proceso está en control, ¿cuál es la ARL de este diagrama?
- Si la media del proceso cambia a $1\sigma_{\bar{X}}$, ¿cuál es la ARL de este diagrama?
- En unidades de $\sigma_{\bar{X}}$, ¿qué tan grande puede ser un cambio positivo para que sea detectado con una ARL de 10?

3. Se toman periódicamente muestras de tres resistores y se miden sus resistencias en ohms. La siguiente tabla presenta las medias, los rangos y las desviaciones estándar de 30 muestras consecutivas.

Muestra	$\bar{X}$	R	s
1	5.114	0.146	0.077
2	5.144	0.158	0.085
3	5.220	0.057	0.031
4	5.196	0.158	0.081
5	5.176	0.172	0.099
6	5.222	0.030	0.017
7	5.209	0.118	0.059
8	5.212	0.099	0.053
9	5.238	0.157	0.085
10	5.152	0.104	0.054
11	5.163	0.051	0.026
12	5.221	0.105	0.055
13	5.144	0.132	0.071
14	5.098	0.123	0.062
15	5.070	0.083	0.042
16	5.029	0.073	0.038
17	5.045	0.161	0.087
18	5.008	0.138	0.071
19	5.029	0.082	0.042
20	5.038	0.109	0.055
21	4.962	0.066	0.034
22	5.033	0.078	0.041
23	4.993	0.085	0.044
24	4.961	0.126	0.066
25	4.976	0.094	0.047
26	5.005	0.135	0.068
27	5.022	0.120	0.062
28	5.077	0.140	0.074
29	5.033	0.049	0.026
30	5.068	0.146	0.076

Las medias son $\bar{\bar{X}} = 5.095$, $\bar{R} = 0.110$ y $\bar{s} = 0.058$.

- a) Calcule los límites 3σ para el diagrama R . ¿La varianza está fuera de control en algún punto? Si es así, elimine las muestras que están fuera de control y vuelva a calcular $\bar{\bar{X}}$ y $\bar{R}$.
 - b) Calcule los límites 3σ para el diagrama $\bar{X}$. Con base en los límites 3σ , ¿la media del proceso está en control? Si no es así, ¿en qué punto se detecta por primera vez que está fuera de control?
 - c) Con base en las reglas de Western Electric, ¿la media del proceso está en control? Si no es así, ¿en qué punto se detecta por primera vez que no lo está?
4. Repita el ejercicio 3, utilizando el diagrama S en lugar del R .
5. Considérese el ejercicio 3.
- a) Elimine todas las muestras que sean necesarias para que la variación del proceso esté bajo control. (Usted ya hizo esto si resolvió el ejercicio 3.)
 - b) Utilice $\bar{R}$ para estimar $\sigma_{\bar{X}}$ ($\sigma_{\bar{X}}$ es la diferencia entre $\bar{\bar{X}}$ y el límite de control 1σ en un diagrama $\bar{X}$).
 - c) Construya un diagrama CUSUM utilizando $\bar{\bar{X}}$ para la media objetivo μ y la estimación de $\sigma_{\bar{X}}$ encontrada en el inciso b) para la desviación estándar. Utilice los valores $k = 0.5$ y $h = 0.4$.
 - d) ¿La media del proceso está en control? Si no es así, ¿en qué punto se detecta por primera vez que no lo está?
 - e) Construya un diagrama $\bar{X}$ y utilice las reglas de Western Electric para determinar si la media del proceso está en control. (Usted ya hizo esto si resolvió el ejercicio 3.)
- ¿Las reglas de Western Electric producen los mismos resultados que el diagrama CUSUM? Si no es así, ¿en qué son diferentes?
6. Se da seguimiento a un proceso para determinar las grietas tomando cada hora una muestra de tamaño 50 y contando el número total de grietas en las unidades extraídas. El número total de grietas en las últimas 30 muestras fue de 658.
- a) Calcule la recta central y los límites de control 3σ superior e inferior.
 - b) La décima muestra tiene tres grietas. ¿El proceso estaba fuera de control en dicho momento? Explique.
7. Para construir un diagrama p que vigile un proceso que produce chips computacionales, diariamente se toman muestras de 500 chips y se cuenta el número de éstos defectuosos en cada muestra. Los números de chips defectuosos para cada uno de los últimos 25 días son los siguientes:
- | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|---|----|----|
| 25 | 22 | 14 | 24 | 18 | 16 | 20 | 27 | 19 | 20 | 22 | 7 | 24 | 26 |
| 11 | 14 | 18 | 29 | 21 | 32 | 29 | 34 | 34 | 30 | 24 | | | |
- a) Calcule los límites 3σ superior e inferior para un diagrama p .
 - b) ¿En qué muestra se detecta por primera vez que el proceso está fuera de control?
 - c) Suponga que se determina la causa especial que da como resultado la condición fuera de control. ¿Debería remediarse esta causa? Explique.



Tablas

- Tabla A.1: Distribución binomial acumulativa
- Tabla A.2: Distribución normal acumulativa
- Tabla A.3: Puntos porcentuales superiores para la distribución t de Student
- Tabla A.4: Puntos críticos para la prueba de rango con signo de Wilcoxon
- Tabla A.5: Puntos críticos para la prueba de rango-suma de Wilcoxon
- Tabla A.6: Puntos porcentuales superiores para la distribución χ^2
- Tabla A.7: Puntos porcentuales superiores para la distribución F
- Tabla A.8: Puntos porcentuales superiores para la distribución de rango studentizado q_{v_1, v_2}
- Tabla A.9: Constantes de diagramas de control

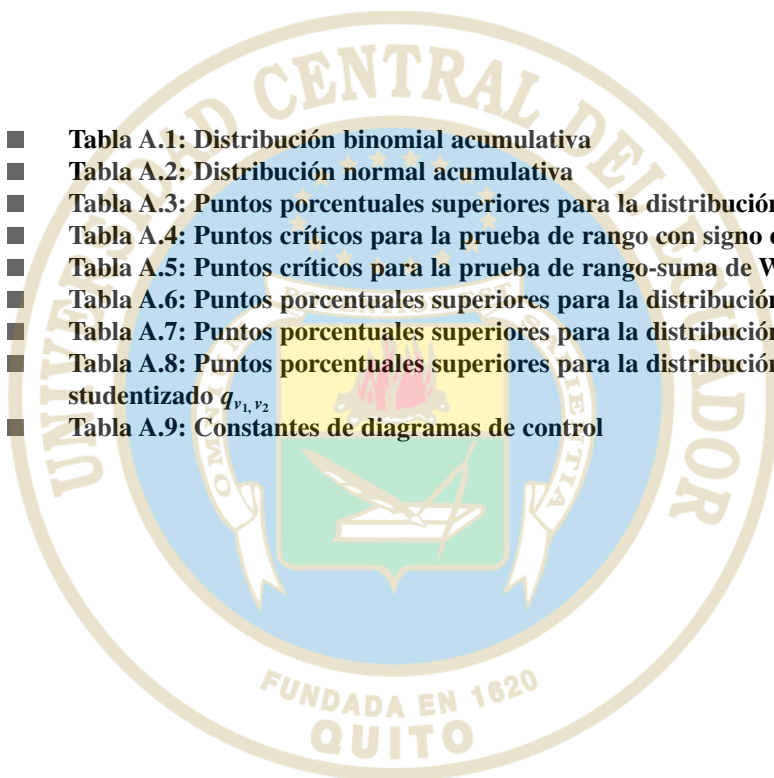


TABLA A.1 Distribución binomial acumulativa

$$F(x) = P(X \leq x) = \sum_{k=0}^x \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$$

		p												
n	x	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
2	0	0.902	0.810	0.640	0.562	0.490	0.360	0.250	0.160	0.090	0.062	0.040	0.010	0.003
	1	0.997	0.990	0.960	0.938	0.910	0.840	0.750	0.640	0.510	0.438	0.360	0.190	0.098
	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0	0.857	0.729	0.512	0.422	0.343	0.216	0.125	0.064	0.027	0.016	0.008	0.001	0.000
	1	0.993	0.972	0.896	0.844	0.784	0.648	0.500	0.352	0.216	0.156	0.104	0.028	0.007
	2	1.000	0.999	0.992	0.984	0.973	0.936	0.875	0.784	0.657	0.578	0.488	0.271	0.143
	3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0	0.815	0.656	0.410	0.316	0.240	0.130	0.062	0.026	0.008	0.004	0.002	0.000	0.000
	1	0.986	0.948	0.819	0.738	0.652	0.475	0.313	0.179	0.084	0.051	0.027	0.004	0.000
	2	1.000	0.996	0.973	0.949	0.916	0.821	0.688	0.525	0.348	0.262	0.181	0.052	0.014
	3	1.000	1.000	0.998	0.996	0.992	0.974	0.938	0.870	0.760	0.684	0.590	0.344	0.185
	4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0	0.774	0.590	0.328	0.237	0.168	0.078	0.031	0.010	0.002	0.001	0.000	0.000	0.000
	1	0.977	0.919	0.737	0.633	0.528	0.337	0.187	0.087	0.031	0.016	0.007	0.000	0.000
	2	0.999	0.991	0.942	0.896	0.837	0.683	0.500	0.317	0.163	0.104	0.058	0.009	0.001
	3	1.000	1.000	0.993	0.984	0.969	0.913	0.812	0.663	0.472	0.367	0.263	0.081	0.023
	4	1.000	1.000	1.000	0.999	0.998	0.990	0.969	0.922	0.832	0.763	0.672	0.410	0.226
	5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
6	0	0.735	0.531	0.262	0.178	0.118	0.047	0.016	0.004	0.001	0.000	0.000	0.000	0.000
	1	0.967	0.886	0.655	0.534	0.420	0.233	0.109	0.041	0.011	0.005	0.002	0.000	0.000
	2	0.998	0.984	0.901	0.831	0.744	0.544	0.344	0.179	0.070	0.038	0.017	0.001	0.000
	3	1.000	0.999	0.983	0.962	0.930	0.821	0.656	0.456	0.256	0.169	0.099	0.016	0.002
	4	1.000	1.000	0.998	0.995	0.989	0.959	0.891	0.767	0.580	0.466	0.345	0.114	0.033
	5	1.000	1.000	1.000	1.000	0.999	0.996	0.984	0.953	0.882	0.822	0.738	0.469	0.265
	6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0	0.698	0.478	0.210	0.133	0.082	0.028	0.008	0.002	0.000	0.000	0.000	0.000	0.000
	1	0.956	0.850	0.577	0.445	0.329	0.159	0.063	0.019	0.004	0.001	0.000	0.000	0.000
	2	0.996	0.974	0.852	0.756	0.647	0.420	0.227	0.096	0.029	0.013	0.005	0.000	0.000
	3	1.000	0.997	0.967	0.929	0.874	0.710	0.500	0.290	0.126	0.071	0.033	0.003	0.000
	4	1.000	1.000	0.995	0.987	0.971	0.904	0.773	0.580	0.353	0.244	0.148	0.026	0.004
	5	1.000	1.000	1.000	0.999	0.996	0.981	0.938	0.841	0.671	0.555	0.423	0.150	0.044
	6	1.000	1.000	1.000	1.000	1.000	0.998	0.992	0.972	0.918	0.867	0.790	0.522	0.302
	7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Continúa en la página 765

TABLA A.1 Distribución binomial acumulativa (continuación)

		p												
n	x	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
8	0	0.663	0.430	0.168	0.100	0.058	0.017	0.004	0.001	0.000	0.000	0.000	0.000	0.000
	1	0.943	0.813	0.503	0.367	0.255	0.106	0.035	0.009	0.001	0.000	0.000	0.000	0.000
	2	0.994	0.962	0.797	0.679	0.552	0.315	0.145	0.050	0.011	0.004	0.001	0.000	0.000
	3	1.000	0.995	0.944	0.886	0.806	0.594	0.363	0.174	0.058	0.027	0.010	0.000	0.000
	4	1.000	1.000	0.990	0.973	0.942	0.826	0.637	0.406	0.194	0.114	0.056	0.005	0.000
	5	1.000	1.000	0.999	0.996	0.989	0.950	0.855	0.685	0.448	0.321	0.203	0.038	0.006
	6	1.000	1.000	1.000	1.000	0.999	0.991	0.965	0.894	0.745	0.633	0.497	0.187	0.057
	7	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.983	0.942	0.900	0.832	0.570	0.337
9	0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0	0.630	0.387	0.134	0.075	0.040	0.010	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.929	0.775	0.436	0.300	0.196	0.071	0.020	0.004	0.000	0.000	0.000	0.000	0.000
	2	0.992	0.947	0.738	0.601	0.463	0.232	0.090	0.025	0.004	0.001	0.000	0.000	0.000
	3	0.999	0.992	0.914	0.834	0.730	0.483	0.254	0.099	0.025	0.010	0.003	0.000	0.000
	4	1.000	0.999	0.980	0.951	0.901	0.733	0.500	0.267	0.099	0.049	0.020	0.001	0.000
	5	1.000	1.000	0.997	0.990	0.975	0.901	0.746	0.517	0.270	0.166	0.086	0.008	0.001
	6	1.000	1.000	1.000	0.999	0.996	0.975	0.910	0.768	0.537	0.399	0.262	0.053	0.008
10	7	1.000	1.000	1.000	1.000	1.000	0.996	0.980	0.929	0.804	0.700	0.564	0.225	0.071
	8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.990	0.960	0.925	0.866	0.613	0.370
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0	0.599	0.349	0.107	0.056	0.028	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.914	0.736	0.376	0.244	0.149	0.046	0.011	0.002	0.000	0.000	0.000	0.000	0.000
	2	0.988	0.930	0.678	0.526	0.383	0.167	0.055	0.012	0.002	0.000	0.000	0.000	0.000
	3	0.999	0.987	0.879	0.776	0.650	0.382	0.172	0.055	0.011	0.004	0.001	0.000	0.000
	4	1.000	0.998	0.967	0.922	0.850	0.633	0.377	0.166	0.047	0.020	0.006	0.000	0.000
11	5	1.000	1.000	0.994	0.980	0.953	0.834	0.623	0.367	0.150	0.078	0.033	0.002	0.000
	6	1.000	1.000	0.999	0.996	0.989	0.945	0.828	0.618	0.350	0.224	0.121	0.013	0.001
	7	1.000	1.000	1.000	1.000	0.998	0.988	0.945	0.833	0.617	0.474	0.322	0.070	0.012
	8	1.000	1.000	1.000	1.000	1.000	0.998	0.989	0.954	0.851	0.756	0.624	0.264	0.086
	9	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.972	0.944	0.893	0.651	0.401
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0	0.569	0.314	0.086	0.042	0.020	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.898	0.697	0.322	0.197	0.113	0.030	0.006	0.001	0.000	0.000	0.000	0.000	0.000
	2	0.985	0.910	0.617	0.455	0.313	0.119	0.033	0.006	0.001	0.000	0.000	0.000	0.000
	3	0.998	0.981	0.839	0.713	0.570	0.296	0.113	0.029	0.004	0.001	0.000	0.000	0.000
	4	1.000	0.997	0.950	0.885	0.790	0.533	0.274	0.099	0.022	0.008	0.002	0.000	0.000
	5	1.000	1.000	0.988	0.966	0.922	0.753	0.500	0.247	0.078	0.034	0.012	0.000	0.000
	6	1.000	1.000	0.998	0.992	0.978	0.901	0.726	0.467	0.210	0.115	0.050	0.003	0.000
	7	1.000	1.000	1.000	0.999	0.996	0.971	0.887	0.704	0.430	0.287	0.161	0.019	0.002
	8	1.000	1.000	1.000	1.000	0.999	0.994	0.967	0.881	0.687	0.545	0.383	0.090	0.015

Continúa en la página 766

TABLA A.1 Distribución binomial acumulativa (continuación)

<i>n</i>	<i>x</i>	<i>p</i>												
		0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
11	9	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.970	0.887	0.803	0.678	0.303	0.102
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.980	0.958	0.914	0.686	0.431
	11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
12	0	0.540	0.282	0.069	0.032	0.014	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.882	0.659	0.275	0.158	0.085	0.020	0.003	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.980	0.889	0.558	0.391	0.253	0.083	0.019	0.003	0.000	0.000	0.000	0.000	0.000
	3	0.998	0.974	0.795	0.649	0.493	0.225	0.073	0.015	0.002	0.000	0.000	0.000	0.000
	4	1.000	0.996	0.927	0.842	0.724	0.438	0.194	0.057	0.009	0.003	0.001	0.000	0.000
	5	1.000	0.999	0.981	0.946	0.882	0.665	0.387	0.158	0.039	0.014	0.004	0.000	0.000
	6	1.000	1.000	0.996	0.986	0.961	0.842	0.613	0.335	0.118	0.054	0.019	0.001	0.000
	7	1.000	1.000	0.999	0.997	0.991	0.943	0.806	0.562	0.276	0.158	0.073	0.004	0.000
	8	1.000	1.000	1.000	1.000	0.998	0.985	0.927	0.775	0.507	0.351	0.205	0.026	0.002
	9	1.000	1.000	1.000	1.000	1.000	0.997	0.981	0.917	0.747	0.609	0.442	0.111	0.020
	10	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.980	0.915	0.842	0.725	0.341	0.118
	11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.986	0.968	0.931	0.718	0.460
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
13	0	0.513	0.254	0.055	0.024	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.865	0.621	0.234	0.127	0.064	0.013	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.975	0.866	0.502	0.333	0.202	0.058	0.011	0.001	0.000	0.000	0.000	0.000	0.000
	3	0.997	0.966	0.747	0.584	0.421	0.169	0.046	0.008	0.001	0.000	0.000	0.000	0.000
	4	1.000	0.994	0.901	0.794	0.654	0.353	0.133	0.032	0.004	0.001	0.000	0.000	0.000
	5	1.000	0.999	0.970	0.920	0.835	0.574	0.291	0.098	0.018	0.006	0.001	0.000	0.000
	6	1.000	1.000	0.993	0.976	0.938	0.771	0.500	0.229	0.062	0.024	0.007	0.000	0.000
	7	1.000	1.000	0.999	0.994	0.982	0.902	0.709	0.426	0.165	0.080	0.030	0.001	0.000
	8	1.000	1.000	1.000	0.999	0.996	0.968	0.867	0.647	0.346	0.206	0.099	0.006	0.000
	9	1.000	1.000	1.000	1.000	0.999	0.992	0.954	0.831	0.579	0.416	0.253	0.034	0.003
	10	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.942	0.798	0.667	0.498	0.134	0.025
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.987	0.936	0.873	0.766	0.379	0.135
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.976	0.945	0.746	0.487
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
14	0	0.488	0.229	0.044	0.018	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.847	0.585	0.198	0.101	0.047	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.970	0.842	0.448	0.281	0.161	0.040	0.006	0.001	0.000	0.000	0.000	0.000	0.000
	3	0.996	0.956	0.698	0.521	0.355	0.124	0.029	0.004	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.991	0.870	0.742	0.584	0.279	0.090	0.018	0.002	0.000	0.000	0.000	0.000
	5	1.000	0.999	0.956	0.888	0.781	0.486	0.212	0.058	0.008	0.002	0.000	0.000	0.000
	6	1.000	1.000	0.988	0.962	0.907	0.692	0.395	0.150	0.031	0.010	0.002	0.000	0.000
	7	1.000	1.000	0.998	0.990	0.969	0.850	0.605	0.308	0.093	0.038	0.012	0.000	0.000

Continúa en la página 767

TABLA A.1 Distribución binomial acumulativa (continuación)

		p												
n	x	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
14	8	1.000	1.000	1.000	0.998	0.992	0.942	0.788	0.514	0.219	0.112	0.044	0.001	0.000
	9	1.000	1.000	1.000	1.000	0.998	0.982	0.910	0.721	0.416	0.258	0.130	0.009	0.000
	10	1.000	1.000	1.000	1.000	1.000	0.996	0.971	0.876	0.645	0.479	0.302	0.044	0.004
	11	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.960	0.839	0.719	0.552	0.158	0.030
	12	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.953	0.899	0.802	0.415	0.153
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.993	0.982	0.956	0.771	0.512
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
15	0	0.463	0.206	0.035	0.013	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.829	0.549	0.167	0.080	0.035	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.964	0.816	0.398	0.236	0.127	0.027	0.004	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.995	0.944	0.648	0.461	0.297	0.091	0.018	0.002	0.000	0.000	0.000	0.000	0.000
	4	0.999	0.987	0.836	0.686	0.515	0.217	0.059	0.009	0.001	0.000	0.000	0.000	0.000
	5	1.000	0.998	0.939	0.852	0.722	0.403	0.151	0.034	0.004	0.001	0.000	0.000	0.000
	6	1.000	1.000	0.982	0.943	0.869	0.610	0.304	0.095	0.015	0.004	0.001	0.000	0.000
	7	1.000	1.000	0.996	0.983	0.950	0.787	0.500	0.213	0.050	0.017	0.004	0.000	0.000
	8	1.000	1.000	0.999	0.996	0.985	0.905	0.696	0.390	0.131	0.057	0.018	0.000	0.000
	9	1.000	1.000	1.000	0.999	0.996	0.966	0.849	0.597	0.278	0.148	0.061	0.002	0.000
	10	1.000	1.000	1.000	1.000	0.999	0.991	0.941	0.783	0.485	0.314	0.164	0.013	0.001
	11	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.909	0.703	0.539	0.352	0.056	0.005
	12	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.973	0.873	0.764	0.602	0.184	0.036
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.965	0.920	0.833	0.451	0.171
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.987	0.965	0.794	0.537
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
16	0	0.440	0.185	0.028	0.010	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.811	0.515	0.141	0.063	0.026	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.957	0.789	0.352	0.197	0.099	0.018	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.993	0.932	0.598	0.405	0.246	0.065	0.011	0.001	0.000	0.000	0.000	0.000	0.000
	4	0.999	0.983	0.798	0.630	0.450	0.167	0.038	0.005	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.997	0.918	0.810	0.660	0.329	0.105	0.019	0.002	0.000	0.000	0.000	0.000
	6	1.000	0.999	0.973	0.920	0.825	0.527	0.227	0.058	0.007	0.002	0.000	0.000	0.000
	7	1.000	1.000	0.993	0.973	0.926	0.716	0.402	0.142	0.026	0.007	0.001	0.000	0.000
	8	1.000	1.000	0.999	0.993	0.974	0.858	0.598	0.284	0.074	0.027	0.007	0.000	0.000
	9	1.000	1.000	1.000	0.998	0.993	0.942	0.773	0.473	0.175	0.080	0.027	0.001	0.000
	10	1.000	1.000	1.000	1.000	0.998	0.981	0.895	0.671	0.340	0.190	0.082	0.003	0.000
	11	1.000	1.000	1.000	1.000	1.000	0.995	0.962	0.833	0.550	0.370	0.202	0.017	0.001
	12	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.935	0.754	0.595	0.402	0.068	0.007
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.901	0.803	0.648	0.211	0.043
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.974	0.937	0.859	0.485	0.189
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.990	0.972	0.815	0.560
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

Continúa en la página 768

TABLA A.1 Distribución binomial acumulativa (continuación)

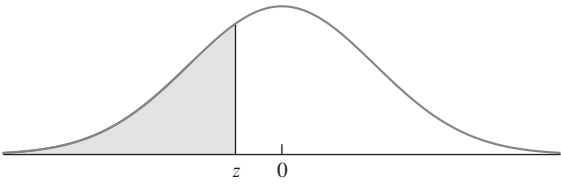
<i>n</i>	<i>x</i>	<i>p</i>												
		0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
17	0	0.418	0.167	0.023	0.008	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.792	0.482	0.118	0.050	0.019	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.950	0.762	0.310	0.164	0.077	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.991	0.917	0.549	0.353	0.202	0.046	0.006	0.000	0.000	0.000	0.000	0.000	0.000
	4	0.999	0.978	0.758	0.574	0.389	0.126	0.025	0.003	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.995	0.894	0.765	0.597	0.264	0.072	0.011	0.001	0.000	0.000	0.000	0.000
	6	1.000	0.999	0.962	0.893	0.775	0.448	0.166	0.035	0.003	0.001	0.000	0.000	0.000
	7	1.000	1.000	0.989	0.960	0.895	0.641	0.315	0.092	0.013	0.003	0.000	0.000	0.000
	8	1.000	1.000	0.997	0.988	0.960	0.801	0.500	0.199	0.040	0.012	0.003	0.000	0.000
	9	1.000	1.000	1.000	0.997	0.987	0.908	0.685	0.359	0.105	0.040	0.011	0.000	0.000
	10	1.000	1.000	1.000	0.999	0.997	0.965	0.834	0.552	0.225	0.107	0.038	0.001	0.000
	11	1.000	1.000	1.000	1.000	0.999	0.989	0.928	0.736	0.403	0.235	0.106	0.005	0.000
	12	1.000	1.000	1.000	1.000	1.000	0.997	0.975	0.874	0.611	0.426	0.242	0.022	0.001
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.954	0.798	0.647	0.451	0.083	0.009
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.923	0.836	0.690	0.238	0.050
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.981	0.950	0.882	0.518	0.208
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.992	0.977	0.833	0.582
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
18	0	0.397	0.150	0.018	0.006	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.774	0.450	0.099	0.039	0.014	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.942	0.734	0.271	0.135	0.060	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.989	0.902	0.501	0.306	0.165	0.033	0.004	0.000	0.000	0.000	0.000	0.000	0.000
	4	0.998	0.972	0.716	0.519	0.333	0.094	0.015	0.001	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.994	0.867	0.717	0.534	0.209	0.048	0.006	0.000	0.000	0.000	0.000	0.000
	6	1.000	0.999	0.949	0.861	0.722	0.374	0.119	0.020	0.001	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.984	0.943	0.859	0.563	0.240	0.058	0.006	0.001	0.000	0.000	0.000
	8	1.000	1.000	0.996	0.981	0.940	0.737	0.407	0.135	0.021	0.005	0.001	0.000	0.000
	9	1.000	1.000	0.999	0.995	0.979	0.865	0.593	0.263	0.060	0.019	0.004	0.000	0.000
	10	1.000	1.000	1.000	0.999	0.994	0.942	0.760	0.437	0.141	0.057	0.016	0.000	0.000
	11	1.000	1.000	1.000	1.000	0.999	0.980	0.881	0.626	0.278	0.139	0.051	0.001	0.000
	12	1.000	1.000	1.000	1.000	1.000	0.994	0.952	0.791	0.466	0.283	0.133	0.006	0.000
	13	1.000	1.000	1.000	1.000	1.000	0.999	0.985	0.906	0.667	0.481	0.284	0.028	0.002
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.967	0.835	0.694	0.499	0.098	0.011
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.940	0.865	0.729	0.266	0.058
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.986	0.961	0.901	0.550	0.226
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.994	0.982	0.850	0.603
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
19	0	0.377	0.135	0.014	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.755	0.420	0.083	0.031	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Continúa en la página 769

TABLA A.1 Distribución binomial acumulativa (continuación)

n	x	p												
		0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
19	2	0.933	0.705	0.237	0.111	0.046	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.987	0.885	0.455	0.263	0.133	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	4	0.998	0.965	0.673	0.465	0.282	0.070	0.010	0.001	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.991	0.837	0.668	0.474	0.163	0.032	0.003	0.000	0.000	0.000	0.000	0.000
	6	1.000	0.998	0.932	0.825	0.666	0.308	0.084	0.012	0.001	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.977	0.923	0.818	0.488	0.180	0.035	0.003	0.000	0.000	0.000	0.000
	8	1.000	1.000	0.993	0.971	0.916	0.667	0.324	0.088	0.011	0.002	0.000	0.000	0.000
	9	1.000	1.000	0.998	0.991	0.967	0.814	0.500	0.186	0.033	0.009	0.002	0.000	0.000
	10	1.000	1.000	1.000	0.998	0.989	0.912	0.676	0.333	0.084	0.029	0.007	0.000	0.000
	11	1.000	1.000	1.000	1.000	0.997	0.965	0.820	0.512	0.182	0.077	0.023	0.000	0.000
	12	1.000	1.000	1.000	1.000	0.999	0.988	0.916	0.692	0.334	0.175	0.068	0.002	0.000
	13	1.000	1.000	1.000	1.000	1.000	0.997	0.968	0.837	0.526	0.332	0.163	0.009	0.000
	14	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.930	0.718	0.535	0.327	0.035	0.002
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.977	0.867	0.737	0.545	0.115	0.013
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.954	0.889	0.763	0.295	0.067
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.969	0.917	0.580	0.245
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.986	0.865	0.623
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	20	0	0.358	0.122	0.012	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1		0.736	0.392	0.069	0.024	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2		0.925	0.677	0.206	0.091	0.035	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3		0.984	0.867	0.411	0.225	0.107	0.016	0.001	0.000	0.000	0.000	0.000	0.000	0.000
4		0.997	0.957	0.630	0.415	0.238	0.051	0.006	0.000	0.000	0.000	0.000	0.000	0.000
5		1.000	0.989	0.804	0.617	0.416	0.126	0.021	0.002	0.000	0.000	0.000	0.000	0.000
6		1.000	0.998	0.913	0.786	0.608	0.250	0.058	0.006	0.000	0.000	0.000	0.000	0.000
7		1.000	1.000	0.968	0.898	0.772	0.416	0.132	0.021	0.001	0.000	0.000	0.000	0.000
8		1.000	1.000	0.990	0.959	0.887	0.596	0.252	0.057	0.005	0.001	0.000	0.000	0.000
9		1.000	1.000	0.997	0.986	0.952	0.755	0.412	0.128	0.017	0.004	0.001	0.000	0.000
10		1.000	1.000	0.999	0.996	0.983	0.872	0.588	0.245	0.048	0.014	0.003	0.000	0.000
11		1.000	1.000	1.000	0.999	0.995	0.943	0.748	0.404	0.113	0.041	0.010	0.000	0.000
12		1.000	1.000	1.000	1.000	0.999	0.979	0.868	0.584	0.228	0.102	0.032	0.000	0.000
13		1.000	1.000	1.000	1.000	1.000	0.994	0.942	0.750	0.392	0.214	0.087	0.002	0.000
14		1.000	1.000	1.000	1.000	1.000	0.998	0.979	0.874	0.584	0.383	0.196	0.011	0.000
15		1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.949	0.762	0.585	0.370	0.043	0.003
16		1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.984	0.893	0.775	0.589	0.133	0.016
17		1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.965	0.909	0.794	0.323	0.075
18		1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.976	0.931	0.608	0.264
19		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	0.988	0.878	0.642
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

TABLA A.2 Distribución normal acumulativa (tabla z)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Continúa en la página 771

TABLA A.2 Distribución normal acumulativa (continuación)

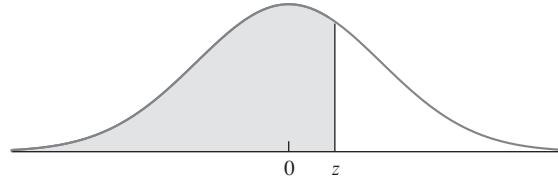
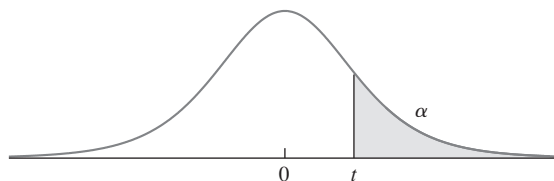
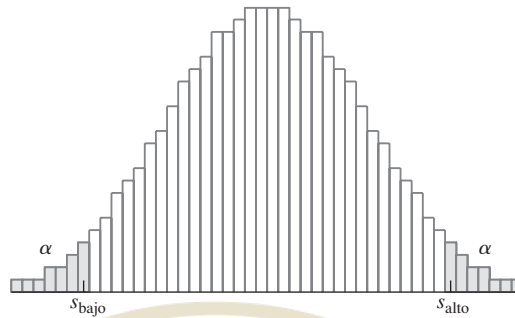
[illegible]

TABLA A.3 Puntos porcentuales superiores para la distribución t de Student



ν	α								
	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
35	0.255	0.682	1.306	1.690	2.030	2.438	2.724	3.340	3.591
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

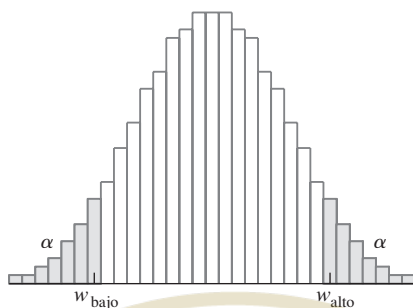
TABLA A.4 Puntos críticos para la prueba de rango con signo de Wilcoxon



n	S_{bajo}	S_{alto}	α	n	S_{bajo}	S_{alto}	α	n	S_{bajo}	S_{alto}	α	n	S_{bajo}	S_{alto}	α
4	1	9	0.1250	10	15	40	0.1162	12	79	0.0085	18	35	118	0.0253	
	0	10	0.0625		14	41	0.0967		10	81		0.0052	34	119	0.0224
5	3	12	0.1562	11	44	0.0527	14	9	82	0.0040	28	125	0.0101		
	2	13	0.0938		10	45		0.0420	32	73		0.1083	27	126	0.0087
	1	14	0.0625		9	46		0.0322	31	74		0.0969	24	129	0.0055
	0	15	0.0312		8	47		0.0244	26	79		0.0520	23	130	0.0047
6	4	17	0.1094	11	6	49	0.0137	15	25	80	0.0453	56	115	0.1061	
	3	18	0.0781		5	50	0.0098		22	83	0.0290		55	116	0.0982
	2	19	0.0469		4	51	0.0068		21	84	0.0247		48	123	0.0542
	1	20	0.0312		3	52	0.0049		16	89	0.0101		47	124	0.0494
	0	21	0.0156		18	48	0.1030		15	90	0.0083		41	130	0.0269
7	6	22	0.1094	12	17	49	0.0874	15	13	92	0.0054	19	40	131	0.0241
	5	23	0.0781		14	52	0.0508		12	93	0.0043		33	138	0.0104
	4	24	0.0547		13	53	0.0415		37	83	0.1039		32	139	0.0091
	3	25	0.0391		11	55	0.0269		36	84	0.0938		28	143	0.0052
	2	26	0.0234		10	56	0.0210		31	89	0.0535		27	144	0.0045
	1	27	0.0156		8	58	0.0122		30	90	0.0473		63	127	0.1051
	0	28	0.0078		7	59	0.0093		26	94	0.0277		62	128	0.0978
	8	9	27		0.1250	12	6		60	0.0068	16		25	95	0.0240
8		28	0.0977	5	61		0.0049	20	100	0.0108		53	137	0.0478	
6		30	0.0547	22	56		0.1018	19	101	0.0090		47	143	0.0273	
5		31	0.0391	21	57		0.0881	16	104	0.0051		46	144	0.0247	
4		32	0.0273	18	60		0.0549	15	105	0.0042		38	152	0.0102	
3		33	0.0195	17	61		0.0461	43	93	0.1057		37	153	0.0090	
2		34	0.0117	14	64		0.0261	42	94	0.0964		33	157	0.0054	
1		35	0.0078	13	65		0.0212	36	100	0.0523		32	158	0.0047	
0		36	0.0039	10	68		0.0105	35	101	0.0467		70	140	0.1012	
9		11	34	0.1016	13		9	69	0.0081	17		30	106	0.0253	24
	10	35	0.0820	8		70	0.0061	29	107		0.0222	61	149	0.0527	
	9	36	0.0645	7		71	0.0046	24	112		0.0107	60	150	0.0487	
	8	37	0.0488	27		64	0.1082	23	113		0.0091	53	157	0.0266	
	6	39	0.0273	26		65	0.0955	20	116		0.0055	52	158	0.0242	
	5	40	0.0195	22		69	0.0549	19	117		0.0046	44	166	0.0107	
	4	41	0.0137	21		70	0.0471	49	104		0.1034	43	167	0.0096	
	3	42	0.0098	18		73	0.0287	48	105		0.0950	38	172	0.0053	
	2	43	0.0059	17		74	0.0239	42	111		0.0544	37	173	0.0047	
	1	44	0.0039	13		78	0.0107	41	112		0.0492				

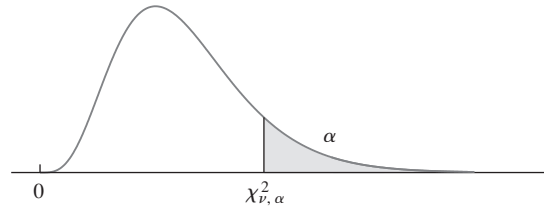
Para $n > 20$, calcule $z = \frac{S_+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$ y utilice la tabla z (tabla A.2).

TABLA A.5 Puntos críticos para la prueba de rango-suma de Wilcoxon



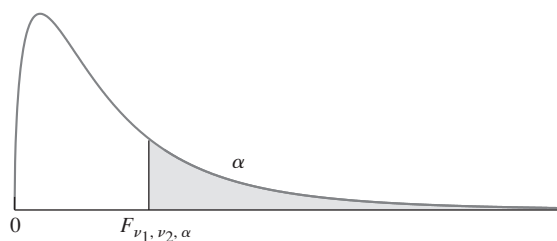
<i>m</i>	<i>n</i>	<i>W</i> _{bajo}	<i>W</i> _{alto}	α	<i>m</i>	<i>n</i>	<i>W</i> _{bajo}	<i>W</i> _{alto}	α	<i>m</i>	<i>n</i>	<i>W</i> _{bajo}	<i>W</i> _{alto}	α	<i>m</i>	<i>n</i>	<i>W</i> _{bajo}	<i>W</i> _{alto}	α
2	5	4	12	0.0952	11	29	0.0159	7	22	43	0.0530	30	60	0.0296					
		3	13	0.0476		10	30	0.0079		21	44	0.0366	29	61	0.0213				
	6	4	14	0.0714	6	14	0.0571	20	45	0.0240		28	62	0.0147					
		3	15	0.0357		13	31	0.0333	19	46	0.0152	27	63	0.0100					
	7	4	16	0.0556	12	32	0.0190	18	47	0.0088		26	64	0.0063					
		3	17	0.0278	11	33	0.0095	17	48	0.0051		25	65	0.0040					
	8	5	17	0.0889	10	34	0.0048	16	49	0.0025	7	7	40	65	0.0641				
		4	18	0.0444	7	15	0.0545	24	46	0.0637		39	66	0.0487					
		3	19	0.0222	14	34	0.0364	23	47	0.0466		37	68	0.0265					
3	4	7	17	0.0571	13	35	0.0212	22	48	0.0326		36	69	0.0189					
		6	18	0.0286	12	36	0.0121	21	49	0.0225		35	70	0.0131					
	5	8	19	0.0714	11	37	0.0061	20	50	0.0148		34	71	0.0087					
		7	20	0.0357	10	38	0.0030	19	51	0.0093		33	72	0.0055					
		6	21	0.0179	8	16	0.0545	18	52	0.0054		32	73	0.0035					
	6	9	21	0.0833	15	37	0.0364	17	53	0.0031	8	42	70	0.0603					
		8	22	0.0476	14	38	0.0242					41	71	0.0469					
		7	23	0.0238	13	39	0.0141	6	6	29	49	0.0660	39	73	0.0270				
	7	9	24	0.0583	12	40	0.0081	28	50	0.0465		38	74	0.0200					
		8	25	0.0333	11	41	0.0040	27	51	0.0325		36	76	0.0103					
		7	26	0.0167				26	52	0.0206		35	77	0.0070					
	6	27	0.0083	5	5	20	35	0.0754	25	53	0.0130	34	78	0.0047					
		10	26	0.0667	19	36	0.0476	24	54	0.0076					8	8	52	84	0.0524
		9	27	0.0424	18	37	0.0278	23	55	0.0043		51	85	0.0415					
		8	28	0.0242	17	38	0.0159	7	30	54	0.0507	50	86	0.0325					
		7	29	0.0121	16	39	0.0079	29	55	0.0367		49	87	0.0249					
		6	30	0.0061	15	40	0.0040	28	56	0.0256		46	90	0.0103					
4	4	12	24	0.0571	6	21	0.0628	27	57	0.0175		45	91	0.0074					
		11	25	0.0286	20	40	0.0411	26	58	0.0111		44	92	0.0052					
		10	26	0.0143	19	41	0.0260	25	59	0.0070		43	93	0.0035					
	5	13	27	0.0556	18	42	0.0152	24	60	0.0041									
		12	28	0.0317	17	43	0.0087	8	32	58	0.0539								
					16	44	0.0043	31	59	0.0406									

Cuando tanto *m* como *n* sean mayores que 8, calcule $z = \frac{W - m(m+n+1)/2}{\sqrt{mn(m+n+1)/12}}$ y utilice la tabla *z* (tabla A.2).

TABLA A.6 Puntos porcentuales superiores para la distribución χ^2 

ν	α									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.191	55.003
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	43.745	47.400	50.725	54.776	57.648
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	46.059	49.802	53.203	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	48.363	52.192	55.668	59.893	62.883
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.996	21.426	23.654	25.695	28.196	50.660	54.572	58.120	62.428	65.476
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766

Para $\nu > 40$, $\chi^2_{\nu, \alpha} \approx 0.5(z_\alpha + \sqrt{2\nu - 1})$.

TABLA A.7 Puntos porcentuales superiores para la distribución F 

ν_2	α	ν_1								
		1	2	3	4	5	6	7	8	9
1	0.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	0.010	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47
	0.001	405284	500012	540382	562501	576405	585938	592874	598144	603040
2	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	0.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
3	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
4	0.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	0.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	0.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	0.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	0.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
6	0.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
7	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33
8	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	0.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11

Continúa en la página 777

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

ν_2	α	ν_1							
		10	12	15	20	25	30	40	60
1	0.100	60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.79
	0.050	241.88	243.91	245.95	248.01	249.26	250.10	251.14	252.20
	0.010	6055.85	6106.32	6157.29	6208.73	6239.83	6260.65	6286.78	6313.03
	0.001	606316	611276	616292	621362	624430	626486	659725	6610390
2	0.100	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47
	0.050	19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48
	0.010	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48
	0.001	999.40	999.42	999.43	999.45	999.46	999.47	999.47	999.48
3	0.100	5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15
	0.050	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.57
	0.010	27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.32
	0.001	129.25	128.32	127.37	126.42	125.84	125.45	124.96	124.47
4	0.100	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79
	0.050	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69
	0.010	14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.65
	0.001	48.05	47.41	46.76	46.10	45.70	45.43	45.09	44.75
5	0.100	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14
	0.050	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.43
	0.010	10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.20
	0.001	26.92	26.42	25.91	25.39	25.08	24.87	24.60	24.33
6	0.100	2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.76
	0.050	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.74
	0.010	7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.06
	0.001	18.41	17.99	17.56	17.12	16.85	16.67	16.44	16.21
7	0.100	2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.51
	0.050	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.30
	0.010	6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.82
	0.001	14.08	13.71	13.32	12.93	12.69	12.53	12.33	12.12
8	0.100	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34
	0.050	3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.01
	0.010	5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.03
	0.001	11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.73
9	0.100	2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.21
	0.050	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.79
	0.010	5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.48
	0.001	9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.19

Continúa en la página 778

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

ν_2	α	ν_1								
		1	2	3	4	5	6	7	8	9
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	0.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48
13	0.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	0.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	0.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	0.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	0.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	0.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	0.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	0.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	0.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	0.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	0.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	0.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	0.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	0.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	0.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	0.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	0.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	0.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	0.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75
18	0.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	0.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	0.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	0.001	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	0.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	0.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	0.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	0.001	15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	0.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	0.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	0.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	0.001	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24

Continúa en la página 779

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

ν_2	α	ν_1								
		10	12	15	20	25	30	40	50	60
10	0.100	2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11
	0.050	2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62
	0.010	4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08
	0.001	8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12
11	0.100	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03
	0.050	2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49
	0.010	4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78
	0.001	7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35
12	0.100	2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96
	0.050	2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38
	0.010	4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54
	0.001	7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76
13	0.100	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.92	1.90
	0.050	2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.31	2.30
	0.010	4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.38	3.34
	0.001	6.80	6.52	6.23	5.93	5.75	5.63	5.47	5.37	5.30
14	0.100	2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.87	1.86
	0.050	2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.24	2.22
	0.010	3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.22	3.18
	0.001	6.40	6.13	5.85	5.56	5.38	5.25	5.10	5.00	4.94
15	0.100	2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.83	1.82
	0.050	2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.18	2.16
	0.010	3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.08	3.05
	0.001	6.08	5.81	5.54	5.25	5.07	4.95	4.80	4.70	4.64
16	0.100	2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.79	1.78
	0.050	2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.12	2.11
	0.010	3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.97	2.93
	0.001	5.81	5.55	5.27	4.99	4.82	4.70	4.54	4.45	4.39
17	0.100	2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.76	1.75
	0.050	2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.08	2.06
	0.010	3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.87	2.83
	0.001	5.58	5.32	5.05	4.78	4.60	4.48	4.33	4.24	4.18
18	0.100	1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.74	1.72
	0.050	2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.04	2.02
	0.010	3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.78	2.75
	0.001	5.39	5.13	4.87	4.59	4.42	4.30	4.15	4.06	4.00
19	0.100	1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.71	1.70
	0.050	2.38	2.31	2.23	2.16	2.11	2.07	2.03	2.00	1.98
	0.010	3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.71	2.67
	0.001	5.22	4.97	4.70	4.43	4.26	4.14	3.99	3.90	3.84
20	0.100	1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.69	1.68
	0.050	2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.97	1.95
	0.010	3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.64	2.61
	0.001	5.08	4.82	4.56	4.29	4.12	4.00	3.86	3.77	3.70

Continúa en la página 780

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

ν_2	α	ν_1								
		1	2	3	4	5	6	7	8	9
21	0.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	0.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	0.001	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	0.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	0.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	0.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	0.001	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	0.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	0.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	0.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	0.001	14.20	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	0.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	0.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	0.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	0.001	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80
25	0.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	0.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	0.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	0.001	13.88	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	0.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	0.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	0.001	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	0.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	0.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	0.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	0.001	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57
28	0.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	0.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	0.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	0.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	0.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	0.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	0.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	0.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	0.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	0.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	0.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	0.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
31	0.100	2.87	2.48	2.27	2.14	2.04	1.97	1.92	1.88	1.84
	0.050	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20
	0.010	7.53	5.36	4.48	3.99	3.67	3.45	3.28	3.15	3.04
	0.001	13.20	8.70	6.99	6.07	5.48	5.07	4.77	4.53	4.34

Continúa en la página 781

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

ν_2	α	ν_1							
		10	12	15	20	25	30	40	60
21	0.100	1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.66
	0.050	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92
	0.010	3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.55
	0.001	4.95	4.70	4.44	4.17	4.00	3.88	3.74	3.58
22	0.100	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64
	0.050	2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.89
	0.010	3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.50
	0.001	4.83	4.58	4.33	4.06	3.89	3.78	3.63	3.48
23	0.100	1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.62
	0.050	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86
	0.010	3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.45
	0.001	4.73	4.48	4.23	3.96	3.79	3.68	3.53	3.38
24	0.100	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61
	0.050	2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.84
	0.010	3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.40
	0.001	4.64	4.39	4.14	3.87	3.71	3.59	3.45	3.29
25	0.100	1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.59
	0.050	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82
	0.010	3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.36
	0.001	4.56	4.31	4.06	3.79	3.63	3.52	3.37	3.22
26	0.100	1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.58
	0.050	2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.80
	0.010	3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.33
	0.001	4.48	4.24	3.99	3.72	3.56	3.44	3.30	3.15
27	0.100	1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.57
	0.050	2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.79
	0.010	3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.29
	0.001	4.41	4.17	3.92	3.66	3.49	3.38	3.23	3.08
28	0.100	1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.56
	0.050	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77
	0.010	3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.26
	0.001	4.35	4.11	3.86	3.60	3.43	3.32	3.18	3.02
29	0.100	1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.55
	0.050	2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.75
	0.010	3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.23
	0.001	4.29	4.05	3.80	3.54	3.38	3.27	3.12	2.97
30	0.100	1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.54
	0.050	2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.74
	0.010	2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.21
	0.001	4.24	4.00	3.75	3.49	3.33	3.22	3.07	2.92
31	0.100	1.81	1.77	1.71	1.66	1.62	1.60	1.56	1.53
	0.050	2.15	2.08	2.00	1.92	1.87	1.83	1.78	1.73
	0.010	2.96	2.82	2.68	2.52	2.43	2.36	2.27	2.18
	0.001	4.19	3.95	3.71	3.45	3.28	3.17	3.03	2.87

Continúa en la página 782

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

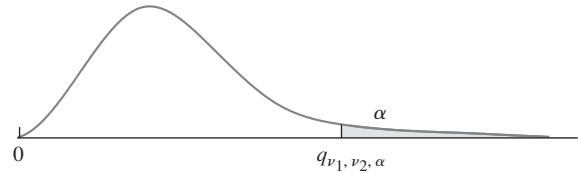
ν_2	α	ν_1								
		1	2	3	4	5	6	7	8	9
32	0.100	2.87	2.48	2.26	2.13	2.04	1.97	1.91	1.87	1.83
	0.050	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19
	0.010	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02
	0.001	13.12	8.64	6.94	6.01	5.43	5.02	4.72	4.48	4.30
33	0.100	2.86	2.47	2.26	2.12	2.03	1.96	1.91	1.86	1.83
	0.050	4.14	3.28	2.89	2.66	2.50	2.39	2.30	2.23	2.18
	0.010	7.47	5.31	4.44	3.95	3.63	3.41	3.24	3.11	3.00
	0.001	13.04	8.58	6.88	5.97	5.38	4.98	4.67	4.44	4.26
34	0.100	2.86	2.47	2.25	2.12	2.02	1.96	1.90	1.86	1.82
	0.050	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17
	0.010	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98
	0.001	12.97	8.52	6.83	5.92	5.34	4.93	4.63	4.40	4.22
35	0.100	2.85	2.46	2.25	2.11	2.02	1.95	1.90	1.85	1.82
	0.050	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16
	0.010	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96
	0.001	12.90	8.47	6.79	5.88	5.30	4.89	4.59	4.36	4.18
36	0.100	2.85	2.46	2.24	2.11	2.01	1.94	1.89	1.85	1.81
	0.050	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15
	0.010	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95
	0.001	12.83	8.42	6.74	5.84	5.26	4.86	4.56	4.33	4.14
37	0.100	2.85	2.45	2.24	2.10	2.01	1.94	1.89	1.84	1.81
	0.050	4.11	3.25	2.86	2.63	2.47	2.36	2.27	2.20	2.14
	0.010	7.37	5.23	4.36	3.87	3.56	3.33	3.17	3.04	2.93
	0.001	12.77	8.37	6.70	5.80	5.22	4.82	4.53	4.30	4.11
38	0.100	2.84	2.45	2.23	2.10	2.01	1.94	1.88	1.84	1.80
	0.050	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14
	0.010	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92
	0.001	12.71	8.33	6.66	5.76	5.19	4.79	4.49	4.26	4.08
39	0.100	2.84	2.44	2.23	2.09	2.00	1.93	1.88	1.83	1.80
	0.050	4.09	3.24	2.85	2.61	2.46	2.34	2.26	2.19	2.13
	0.010	7.33	5.19	4.33	3.84	3.53	3.30	3.14	3.01	2.90
	0.001	12.66	8.29	6.63	5.73	5.16	4.76	4.46	4.23	4.05
40	0.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	0.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	0.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	0.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	0.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	0.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	0.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	0.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	0.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	0.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	0.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	0.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
120	0.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	0.050	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
	0.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	0.001	11.38	7.32	5.78	4.95	4.42	4.04	3.77	3.55	3.38

Continúa en la página 783

TABLA A.7 Puntos porcentuales superiores para la distribución F (continuación)

v_2	α	v_1								
		10	12	15	20	25	30	40	50	60
32	0.100	1.81	1.76	1.71	1.65	1.62	1.59	1.56	1.53	1.52
	0.050	2.14	2.07	1.99	1.91	1.85	1.82	1.77	1.74	1.71
	0.010	2.93	2.80	2.65	2.50	2.41	2.34	2.25	2.20	2.16
	0.001	4.14	3.91	3.66	3.40	3.24	3.13	2.98	2.89	2.83
33	0.100	1.80	1.75	1.70	1.64	1.61	1.58	1.55	1.53	1.51
	0.050	2.13	2.06	1.98	1.90	1.84	1.81	1.76	1.72	1.70
	0.010	2.91	2.78	2.63	2.48	2.39	2.32	2.23	2.18	2.14
	0.001	4.10	3.87	3.62	3.36	3.20	3.09	2.94	2.85	2.79
34	0.100	1.79	1.75	1.69	1.64	1.60	1.58	1.54	1.52	1.50
	0.050	2.12	2.05	1.97	1.89	1.83	1.80	1.75	1.71	1.69
	0.010	2.89	2.76	2.61	2.46	2.37	2.30	2.21	2.16	2.12
	0.001	4.06	3.83	3.58	3.33	3.16	3.05	2.91	2.82	2.75
35	0.100	1.79	1.74	1.69	1.63	1.60	1.57	1.53	1.51	1.50
	0.050	2.11	2.04	1.96	1.88	1.82	1.79	1.74	1.70	1.68
	0.010	2.88	2.74	2.60	2.44	2.35	2.28	2.19	2.14	2.10
	0.001	4.03	3.79	3.55	3.29	3.13	3.02	2.87	2.78	2.72
36	0.100	1.78	1.73	1.68	1.63	1.59	1.56	1.53	1.51	1.49
	0.050	2.11	2.03	1.95	1.87	1.81	1.78	1.73	1.69	1.67
	0.010	2.86	2.72	2.58	2.43	2.33	2.26	2.18	2.12	2.08
	0.001	3.99	3.76	3.51	3.26	3.10	2.98	2.84	2.75	2.69
37	0.100	1.78	1.73	1.68	1.62	1.58	1.56	1.52	1.50	1.48
	0.050	2.10	2.02	1.95	1.86	1.81	1.77	1.72	1.68	1.66
	0.010	2.84	2.71	2.56	2.41	2.31	2.25	2.16	2.10	2.06
	0.001	3.96	3.73	3.48	3.23	3.07	2.95	2.81	2.72	2.66
38	0.100	1.77	1.72	1.67	1.61	1.58	1.55	1.52	1.49	1.48
	0.050	2.09	2.02	1.94	1.85	1.80	1.76	1.71	1.68	1.65
	0.010	2.83	2.69	2.55	2.40	2.30	2.23	2.14	2.09	2.05
	0.001	3.93	3.70	3.45	3.20	3.04	2.92	2.78	2.69	2.63
39	0.100	1.77	1.72	1.67	1.61	1.57	1.55	1.51	1.49	1.47
	0.050	2.08	2.01	1.93	1.85	1.79	1.75	1.70	1.67	1.65
	0.010	2.81	2.68	2.54	2.38	2.29	2.22	2.13	2.07	2.03
	0.001	3.90	3.67	3.43	3.17	3.01	2.90	2.75	2.66	2.60
40	0.100	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.48	1.47
	0.050	2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.66	1.64
	0.010	2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.06	2.02
	0.001	3.87	3.64	3.40	3.14	2.98	2.87	2.73	2.64	2.57
50	0.100	1.73	1.68	1.63	1.57	1.53	1.50	1.46	1.44	1.42
	0.050	2.03	1.95	1.87	1.78	1.73	1.69	1.63	1.60	1.58
	0.010	2.70	2.56	2.42	2.27	2.17	2.10	2.01	1.95	1.91
	0.001	3.67	3.44	3.20	2.95	2.79	2.68	2.53	2.44	2.38
60	0.100	1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.41	1.40
	0.050	1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.56	1.53
	0.010	2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.88	1.84
	0.001	3.54	3.32	3.08	2.83	2.67	2.55	2.41	2.32	2.25
120	0.100	1.65	1.60	1.55	1.48	1.44	1.41	1.37	1.34	1.32
	0.050	1.91	1.83	1.75	1.66	1.60	1.55	1.50	1.46	1.43
	0.010	2.47	2.34	2.19	2.03	1.93	1.86	1.76	1.70	1.66
	0.001	3.24	3.02	2.78	2.53	2.37	2.26	2.11	2.02	1.95

TABLA A.8 Puntos porcentuales superiores para la distribución de rango studentizado q_{v_1, v_2}



		v_1														
v_2	α	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	0.10	8.93	13.44	16.36	18.49	20.15	21.51	22.64	23.62	24.48	25.24	25.92	26.54	27.10	27.62	
	0.05	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	54.33	55.36	
	0.01	90.02	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253.2	260.0	266.2	271.8	277.0	
2	0.10	4.13	5.73	6.77	7.54	8.14	8.63	9.05	9.41	9.72	10.01	10.26	10.49	10.70	10.89	
	0.05	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	
	0.01	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69	32.59	33.40	34.13	34.81	35.43	
3	0.10	3.33	4.47	5.20	5.74	6.16	6.51	6.81	7.06	7.29	7.49	7.67	7.83	7.98	8.12	
	0.05	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.15	10.35	10.52	
	0.01	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13	17.53	17.89	18.22	18.52	
4	0.10	3.01	3.98	4.59	5.04	5.39	5.68	5.93	6.14	6.33	6.49	6.65	6.78	6.91	7.02	
	0.05	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.66	
	0.01	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27	12.57	12.84	13.09	13.32	13.53	
5	0.10	2.85	3.72	4.26	4.66	4.98	5.24	5.46	5.65	5.82	5.97	6.10	6.22	6.34	6.44	
	0.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	
	0.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	
6	0.10	2.75	3.56	4.07	4.44	4.73	4.97	5.17	5.34	5.50	5.64	5.76	5.87	5.98	6.07	
	0.05	3.46	4.34	4.90	5.31	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14	
	0.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	9.95	
7	0.10	2.68	3.45	3.93	4.28	4.55	4.78	4.97	5.14	5.28	5.41	5.53	5.64	5.74	5.83	
	0.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76	
	0.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	
8	0.10	2.63	3.37	3.83	4.17	4.43	4.65	4.83	4.99	5.13	5.25	5.36	5.46	5.56	5.64	
	0.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	
	0.01	4.75	5.64	6.20	6.63	6.96	7.24	7.47	7.68	7.86	8.03	8.18	8.31	8.44	8.55	
9	0.10	2.59	3.32	3.76	4.08	4.34	4.54	4.72	4.87	5.01	5.13	5.23	5.33	5.42	5.51	
	0.05	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98	6.09	6.19	6.28	
	0.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	7.78	7.91	8.03	8.13	
10	0.10	2.56	3.27	3.70	4.02	4.26	4.47	4.64	4.78	4.91	5.03	5.13	5.23	5.32	5.40	
	0.05	3.15	3.88	4.33	4.65	4.91	5.12	5.31	5.46	5.60	5.72	5.83	5.93	6.03	6.11	
	0.01	4.48	5.27	5.77	6.14	6.43	6.67	6.88	7.05	7.21	7.36	7.49	7.60	7.71	7.81	
11	0.10	2.54	3.23	3.66	3.96	4.20	4.40	4.57	4.71	4.84	4.95	5.05	5.15	5.23	5.31	
	0.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.99	
	0.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	
12	0.10	2.52	3.20	3.62	3.92	4.16	4.35	4.51	4.65	4.78	4.89	4.99	5.08	5.16	5.24	
	0.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.62	5.71	5.80	5.88	
	0.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	

Continúa en la página 785

TABLA A.8 Puntos porcentuales superiores para la distribución de rango studentizado q_{v_1, v_2} (continuación)

v_2	α	v_1													
		2	3	4	5	6	7	8	9	10	11	12	13	14	15
13	0.10	2.50	3.18	3.59	3.88	4.12	4.30	4.46	4.60	4.72	4.83	4.93	5.02	5.10	5.18
	0.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79
	0.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19
14	0.10	2.49	3.16	3.56	3.85	4.08	4.27	4.42	4.56	4.68	4.79	4.88	4.97	5.05	5.12
	0.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.72
	0.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05
15	0.10	2.48	3.14	3.54	3.83	4.05	4.23	4.39	4.52	4.64	4.75	4.84	4.93	5.01	5.08
	0.05	3.01	3.67	4.08	4.37	4.60	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.58	5.65
	0.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93
16	0.10	2.47	3.12	3.52	3.80	4.03	4.21	4.36	4.49	4.61	4.71	4.81	4.89	4.97	5.04
	0.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59
	0.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82
17	0.10	2.46	3.11	3.50	3.78	4.00	4.18	4.33	4.46	4.58	4.68	4.77	4.86	4.93	5.01
	0.05	2.98	3.63	4.02	4.30	4.52	4.71	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.55
	0.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73
18	0.10	2.45	3.10	3.49	3.77	3.98	4.16	4.31	4.44	4.55	4.65	4.75	4.83	4.90	4.98
	0.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	5.50
	0.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65
19	0.10	2.44	3.09	3.47	3.75	3.97	4.14	4.29	4.42	4.53	4.63	4.72	4.80	4.88	4.95
	0.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.32	5.39	5.46
	0.01	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58
20	0.10	2.44	3.08	3.46	3.74	3.95	4.12	4.27	4.40	4.51	4.61	4.70	4.78	4.85	4.92
	0.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43
	0.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.52
24	0.10	2.42	3.05	3.42	3.69	3.90	4.07	4.21	4.34	4.45	4.54	4.63	4.71	4.78	4.85
	0.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32
	0.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33
30	0.10	2.40	3.02	3.39	3.65	3.85	4.02	4.16	4.28	4.38	4.47	4.56	4.64	4.71	4.77
	0.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00	5.08	5.15	5.21
	0.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14
40	0.10	2.38	2.99	3.35	3.60	3.80	3.96	4.10	4.21	4.32	4.41	4.49	4.56	4.63	4.69
	0.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74	4.82	4.90	4.98	5.05	5.11
	0.01	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60	5.69	5.76	5.83	5.90	5.96
60	0.10	2.36	2.96	3.31	3.56	3.75	3.91	4.04	4.16	4.25	4.34	4.42	4.49	4.56	4.62
	0.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00
	0.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79
120	0.10	2.34	2.93	3.28	3.52	3.71	3.86	3.99	4.10	4.19	4.28	4.35	4.42	4.48	4.54
	0.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90
	0.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.50	5.56	5.61
∞	0.10	2.33	2.90	3.24	3.48	3.66	3.81	3.93	4.04	4.13	4.21	4.28	4.35	4.41	4.47
	0.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80
	0.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45

TABLA A.9 Constantes de diagramas de control

n	A_2	A_3	B_3	B_4	D_3	D_4	c_4	d_2
2	1.880	2.659	0.000	3.267	0.000	3.267	0.7979	1.128
3	1.023	1.954	0.000	2.568	0.000	2.575	0.8862	1.693
4	0.729	1.628	0.000	2.266	0.000	2.282	0.9213	2.059
5	0.577	1.427	0.000	2.089	0.000	2.114	0.9400	2.326
6	0.483	1.287	0.030	1.970	0.000	2.004	0.9515	2.534
7	0.419	1.182	0.118	1.882	0.076	1.924	0.9594	2.704
8	0.373	1.099	0.185	1.815	0.136	1.864	0.9650	2.847
9	0.337	1.032	0.239	1.761	0.184	1.816	0.9693	2.970
10	0.308	0.975	0.284	1.716	0.223	1.777	0.9727	3.078
11	0.285	0.927	0.321	1.679	0.256	1.744	0.9754	3.173
12	0.266	0.866	0.354	1.646	0.283	1.717	0.9776	3.258
13	0.249	0.850	0.382	1.618	0.307	1.693	0.9794	3.336
14	0.235	0.817	0.406	1.594	0.328	1.672	0.9810	3.407
15	0.223	0.789	0.428	1.572	0.347	1.653	0.9823	3.472
16	0.212	0.763	0.448	1.552	0.363	1.637	0.9835	3.532
17	0.203	0.739	0.466	1.534	0.378	1.622	0.9845	3.588
18	0.194	0.718	0.482	1.518	0.391	1.609	0.9854	3.640
19	0.187	0.698	0.497	1.503	0.403	1.597	0.9862	3.689
20	0.180	0.680	0.510	1.490	0.415	1.585	0.9869	3.735
21	0.173	0.663	0.523	1.477	0.425	1.575	0.9876	3.778
22	0.167	0.647	0.534	1.466	0.434	1.566	0.9882	3.819
23	0.162	0.633	0.545	1.455	0.443	1.557	0.9887	3.858
24	0.157	0.619	0.555	1.445	0.452	1.548	0.9892	3.895
25	0.153	0.606	0.565	1.435	0.459	1.541	0.9896	3.931

Para $n > 25$: $A_3 \approx 3/\sqrt{n}$, $B_3 \approx 1 - 3/\sqrt{2n}$, y $B_4 \approx 1 + 3/\sqrt{2n}$.

Derivadas parciales

Este apéndice presenta los mecanismos para calcular las derivadas parciales que se necesitan en la sección 3.4. Se inicia recordando que una derivada especifica la razón de cambio de una variable con respecto a la otra. Por ejemplo, el volumen v de una esfera cuyo radio r está dado por $v = 4\pi r^3$. Si a r se le permite aumentar (o disminuir) la razón con la cual v aumenta (o disminuye) está dada por la derivada de v con respecto a r : $dv/dr = 12\pi r^2$.

Las derivadas parciales se necesitan cuando la cantidad cuya razón de cambio se está calculando es una función de más de una variable. Aquí se presenta un ejemplo. El volumen v de un cilindro cuyo radio r y altura h está dado por $v = \pi r^2 h$. Si alguna de las dos, r o h cambian, v también lo hará. Ahora imagine que h es una constante y que a r se le permite aumentar. La razón de aumento en v está dada por la **derivada parcial** de v con respecto a r . Esta derivada se denota por $\partial v/\partial r$, y se calcula exactamente como la derivada ordinaria de v con respecto a r , tratando a h como una constante: $\partial v/\partial r = 2\pi r h$.

Ahora suponga que r es una constante, y que h está aumentando. La razón de aumento en v es la derivada parcial de v con respecto a h , que se denota por $\partial v/\partial h$. Ésta se calcula exactamente como la derivada ordinaria de v con respecto a h , tratando a r como una constante: $\partial v/\partial h = \pi r^2$.

Si v es una función de varias variables, $v = f(x_1, x_2, \dots, x_n)$, entonces la derivada parcial de v con respecto a una de las variables $x_1, x_2, \dots, x_n$ se calcula de la misma forma que la derivada ordinaria, manteniendo a las otras variables constantes.

Los ejemplos B.1 y B.2 muestran que el cálculo de derivadas parciales no es más difícil que el cálculo de derivadas ordinarias.

Ejemplo

B.1

Sea $v = 12x^2y + 3xy^2$. Determine las derivadas parciales de v con respecto a x y y .

Solución

Para calcular $\partial v/\partial x$, manteniendo a y constante, y calculando la derivada con respecto a x . El resultado es

$$\frac{\partial v}{\partial x} = 24xy + 3y^2$$

Para calcular $\partial v / \partial y$, manteniendo a x constante, y calculando la derivada con respecto a y . El resultado es

$$\frac{\partial v}{\partial y} = 12x^2 + 6xy$$

Ejemplo

B.2

Sea $v = \frac{x^3y + y^3z - xz^3}{x^2 + y^2 + z^2}$. Determine las derivadas parciales de v con respecto a x , y , y z .

Solución

Para calcular $\partial v / \partial x$, manteniendo tanto a y como a z constantes, y calculando la derivada con respecto a x , utilizando la regla del cociente:

$$\frac{\partial v}{\partial x} = \frac{(3x^2y - z^3)(x^2 + y^2 + z^2) - (x^3y + y^3z - xz^3)(2x)}{(x^2 + y^2 + z^2)^2}$$

De manera semejante se calculan las derivadas parciales de v con respecto a y y a z :

$$\frac{\partial v}{\partial y} = \frac{(x^3 + 3y^2z)(x^2 + y^2 + z^2) - (x^3y + y^3z - xz^3)(2y)}{(x^2 + y^2 + z^2)^2}$$

$$\frac{\partial v}{\partial z} = \frac{(y^3 - 3xz^2)(x^2 + y^2 + z^2) - (x^3y + y^3z - xz^3)(2z)}{(x^2 + y^2 + z^2)^2}$$

Ejercicios para el apéndice B

En los siguientes ejercicios, calcule todas las derivadas parciales.

1. $v = 3x + 2xy^4$

2. $w = \frac{x^3 + y^3}{x^2 + y^2}$

3. $z = \cos x \sin y^2$

4. $v = e^{xy}$

5. $v = e^x (\cos y + \sin z)$

6. $w = \sqrt{x^2 + 4y^2 + 3z^2}$

7. $z = \ln(x^2 + y^2)$

8. $v = e^{y^2} \cos(xz) + \ln(x^2y + z)$

9. $v = \frac{2xy^3 - 3xy^2}{\sqrt{xy}}$

10. $z = \sqrt{\sin(x^2y)}$

Bibliografía

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. John Wiley & Sons, Nueva York. Un autorizado y completo tratado del tema.

Belsey, D., Kuh, E. y Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, Nueva York. Una presentación de métodos para evaluar la confiabilidad de las estimaciones de regresión.

Bevington, P. y Robinson, D. (2002). *Data Reduction and Error Analysis for the Physical Sciences*, 3a. ed. McGraw-Hill, Boston. Una introducción al análisis de datos con énfasis en la propagación del error y su cálculo.

Bickel, P. y Doksum, K. (2001). *Mathematical Statistics: Basic Ideas and Selected Topics*, vol. 1, 2a. ed. Prentice-Hall, Upper Saddle River, NJ. Un minucioso tratado de los principios matemáticos de la estadística en un nivel bastante avanzado.

Box, G. y Draper, N. (1987). *Empirical Model-Building and Response Surfaces*. John Wiley & Sons, Nueva York. Una excelente introducción práctica al ajuste de curvas y superficies de alto orden de los datos.

Box, G., Hunter, W. y Hunter, J. (1978). *Statistics for Experimenters*. John Wiley & Sons, Nueva York. Una introducción práctica e intuitiva de los principios básicos del análisis de datos y diseño experimental.

Brockwell, R. y Davis, R. (2002). *Introduction to Time Series and Forecasting*, 2a. ed. Springer-Verlag, Nueva York. Un excelente texto introductorio a nivel de licenciatura, más riguroso que Chatfield (2003).

Casella, G. y Berger, R. (2002). *Statistical Inference*, 2a. ed. Duxbury, Pacific Grove, CA. Un riguroso desarrollo de la teoría de la estadística.

Chatfield, C. (2003). *An Analysis of Time Series: An Introduction*, 6a. ed. CRC Press, Boca Raton, FL. Una presentación intuitiva, en un nivel un poco menos avanzado que Brockwell y Davis (2002).

Chatfield, C. (1983). *Statistics for Technology*, 3a. ed., revisada. Chapman y Hall/CRC, Boca Raton, FL. Una obvia y concisa introducción de los principios básicos de la estadística, orientado a los ingenieros y científicos.

Cochran, W. (1977). *Sampling Techniques*, 3a. ed. John Wiley & Sons, Nueva York. Una explicación completa de la teoría del muestreo.

Cook, D. y Weisberg, S. (1994). *An Introduction to Regression Graphics*. John Wiley & Sons, Nueva York. Una presentación de los métodos interactivos para el análisis de datos con modelos lineales.

DeGroot, M. y Schervish, M. (2002). *Probability and Statistics*, 3a. ed. Addison-Wesley, Reading, MA. Una introducción muy entendible en un nivel matemático un poco más alto que este libro.

Draper, N. y Smith, H. (1998). *Applied Regression Analysis*, 3a. ed. John Wiley & Sons, Nueva York. Un tratado amplio y autorizado de la regresión lineal.

Efron, B. y Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, Nueva York. Una clara y completa introducción a los métodos de estimación por bootstrap.

Freedman, D., Pisani, R. y Purves, R. (1998). *Statistics*, 3a. ed. Norton, Nueva York. Una excelente introducción a los principios fundamentales de la estadística.

Hocking, R. (2003). *Methods and Applications of Linear Models: Regression and the Analysis of Variance*, 2a. ed. John Wiley & Sons, Nueva York. Un minucioso tratado de la teoría y aplicaciones de la regresión y del análisis de varianza.

Kenett, R. y Zacks, S. (1998). *Modern Industrial Statistics*. Brooks/Cole, Pacific Grove, California. Un tratado de actualización del tema con énfasis en la ingeniería industrial.

Larsen, R. y Marx, M. (2001). *An Introduction to Mathematical Statistics and Its Applications*, 3a. ed. Prentice-Hall, Upper Saddle River, NJ. Una introducción a la estadística en un nivel matemático superior a este libro. Contiene muy buenos ejemplos.

Lee, P. (1997). *Bayesian Statistics: An Introduction*, 2a. ed. Oxford University Press, Nueva York. Una introducción clara y básica a los métodos estadísticos basados en el punto de vista subjetivo de la probabilidad.

Lehmann, E. (1998). *Nonparametrics: Statistical Methods Based on Ranks*. Pearson, Nueva York. Una minuciosa presentación de los métodos libres de distribución.

Miller, A. (2002). *Subset Selection in Regression*, 2a. ed. Chapman and Hall, Londres. Un fuerte y conciso tratado de los principios básicos de la selección de modelos.

Miller, R. (1997). *Beyond ANOVA: The Basics of Applied Statistics*. Chapman and Hall/CRC, Boca Raton, FL. Un tratado muy práctico e intuitivo de los métodos útiles para el análisis de datos reales, cuando no se satisfacen los supuestos usuales.

Montgomery, D. (2001a). *Design and Analysis of Experiments*, 5a. ed. John Wiley & Sons, Nueva York. Una minuciosa exposición de los métodos de los experimentos factoriales dirigidos a aplicaciones en ingeniería.

Montgomery, D. (2001b). *Introduction to Statistical Quality Control*, 4a. ed. John Wiley & Sons, Nueva York. Una completa e inteligible introducción al tema.

- Mood, A., Graybill, F. y Boes, D. (1974). *Introduction to the Theory of Statistics*, 3a. ed. McGraw-Hill, Boston. Una introducción clásica a la estadística matemática y una excelente referencia bibliográfica.
- Mosteller, F. y Tukey, J. (1977). *Data Analysis and Regression*. Addison-Wesley, Reading, MA. Una intuitiva y filosófica presentación de algunas ideas prácticas.
- Rice, J. (1995). *Mathematical Statistics and Data Analysis*, 2a. ed. Wadsworth, Belmont, CA. Una buena mezcla de teoría y práctica, en un nivel un poco más alto que este libro.
- Ross, S. (2001). *A First Course in Probability*, 6a. ed. Prentice-Hall, Upper Saddle River, NJ. Una introducción sofisticadamente matemática a la probabilidad.
- Ross, S. (2004). *Introduction to Probability and Statistics for Engineers and Scientists*, 3a. ed. Harcourt/Academic Press, San Diego. Una introducción a un nivel matemático un poco superior a este libro.
- Salsburg, D. (2001). *The Lady Tasting Tea*. W. H. Freeman and Company, Nueva York. Un análisis perspicaz de la influencia de la estadística en la ciencia del siglo xx, con muchas anécdotas fascinantes de estadígrafos famosos. En este libro se puede encontrar la historia en la sección 4.3 acerca del descubrimiento de Student de que el número de partículas en una suspensión seguía una distribución de Poisson.
- Tanur, J., Pieters, R. y Mosteller, F. (eds.) (1989). *Statistics: A Guide to the Unknown*, 3a. ed. Wadsworth/Brooks-Cole, Pacific Grove, CA. Un conjunto de estudios de caso que muestra muchas aplicaciones estadísticas.
- Taylor, J. (1997). *An Introduction to Error Analysis*, 2a. ed. University Science Books, Sausalito, CA. Un minucioso tratado de la propagación del error, junto con una selección de otros temas en los análisis de datos.
- Tufte, E. (2001). *The Visual Display of Quantitative Information*, 2a. ed. Graphics Press, Cheshire, CT. Una clara y fascinante demostración de los principios efectivos de las gráficas estadísticas; contiene numerosos ejemplos.
- Tukey, J. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA. Una riqueza de técnicas para resumir y describir datos.
- Wackerly, D., Mendenhall, W. y Scheaffer, R. (2002). *Mathematical Statistics with Applications*, 6a. ed. Duxbury, Pacific Grove, CA. Una introducción a la estadística en un nivel matemático un poco superior a este libro.
- Weisberg, S. (1985). *Applied Linear Regression*, 2a. ed. John Wiley & Sons, Nueva York. Una introducción concisa de las aplicaciones de los modelos de regresión lineal, incluyendo diagnósticos, construcción de modelos e interpretación de resultados.

Respuestas a los ejercicios impares

Sección 1.1

1.
 - a) La población consta de todos los tornillos de la remesa. Es tangible.
 - b) La población consta de todas las mediciones que se hicieron en el resistor con el óhmetro. Es conceptual.
 - c) La población consta de todos los habitantes del pueblo. Es tangible.
 - d) La población consta de todas las soldaduras que pudieron haberse hecho con el proceso. Es conceptual.
 - e) La población consta de todas las partes fabricadas ese día. Es tangible.
3.
 - a) Falso
 - b) Verdadero
5.
 - a) No. Lo importante es la proporción poblacional de defectos: el porcentaje de la muestra es sólo una aproximación. La proporción poblacional para el nuevo proceso puede, en efecto, ser mayor o menor que la del proceso anterior.
 - b) No. La proporción poblacional para el nuevo proceso puede ser 10% o mayor, aun cuando la proporción muestral fue de sólo 9%.
 - c) Encontrando dos botellas defectuosas en la muestra.
7. Un buen conocimiento del proceso que produce los datos.

Sección 1.2

1. Falso.
3. No. En las muestras 1, 2, 4, la media es $7/3$, que no aparece en toda la muestra.
5. El tamaño de la muestra puede ser cualquier número impar.
7. La media y la desviación estándar aumentan 5%.
9. La media es 2.45, la mediana es 3 y la desviación estándar es 1.2999.

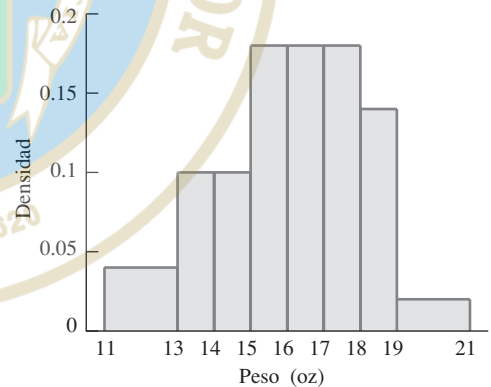
11. a) Sí, la media es 2.45.
 b) Sí, la mediana es 3.
 c) No, la desviación estándar de la muestra depende del tamaño de la muestra.
13. a) Todo se multiplicaría por 2.54.
 b) No exactamente lo mismo, puesto que las mediciones serían un poco diferentes la segunda vez.
15. No existe un cuarto cuartil. Hay tres cuartiles; son los puntos que dividen los datos en cuatro grupos iguales.

Sección 1.3

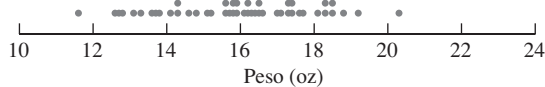
1. a)

Tallos	Hojas
11	6
12	678
13	13678
14	13368
15	126678899
16	122345556
17	013344467
18	1333558
19	2
20	3

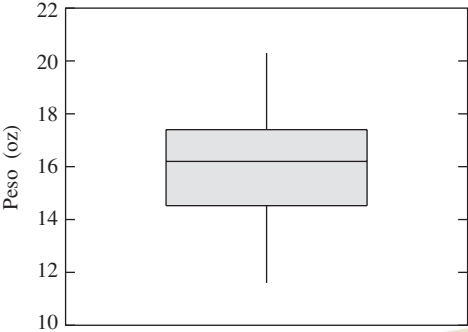
b) Aquí se muestra un histograma. Son posibles otras elecciones para los puntos finales.



c)



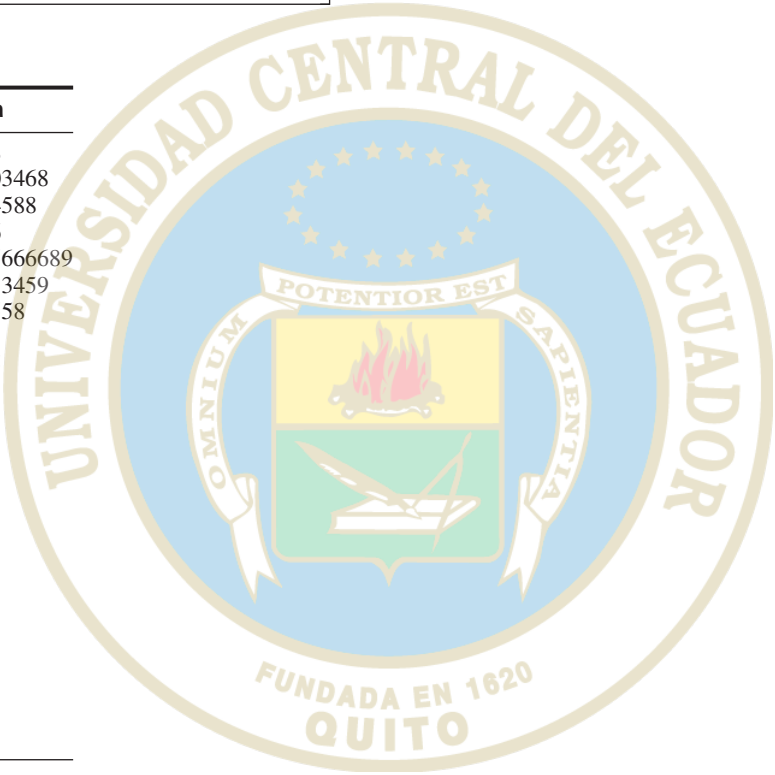
d)



El diagrama de puntos no muestra datos atípicos.

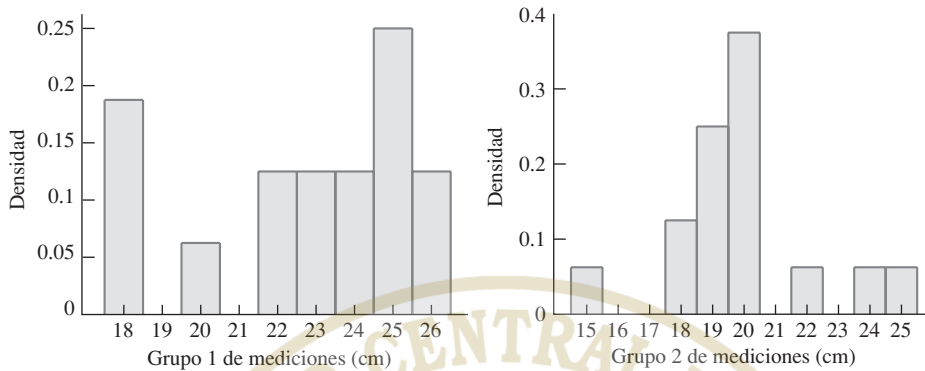
3.

Tallo	Hoja
1	1588
2	00003468
3	0234588
4	0346
5	2235666689
6	00233459
7	113558
8	568
9	1225
10	1
11	
12	2
13	06
14	
15	
16	
17	1
18	6
19	9
20	
21	
22	
23	3

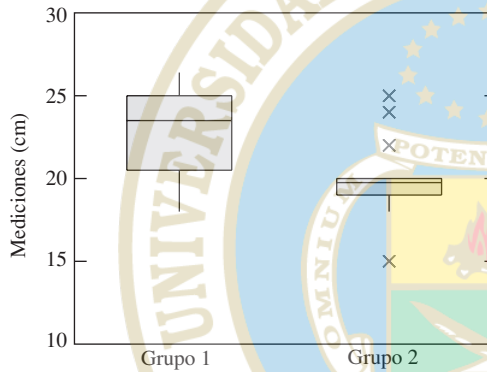


Hay 23 tallos en esta gráfica; una de sus ventajas sobre la de la figura 1.6 consiste en que los valores están dados con dígitos en los décimos en lugar de un dígito. Una desventaja es que hay muchos tallos, y muchos de éstos están vacíos.

5. a) Aquí se muestran los histogramas en cada grupo. Otras elecciones para los puntos extremos son posibles.



b)

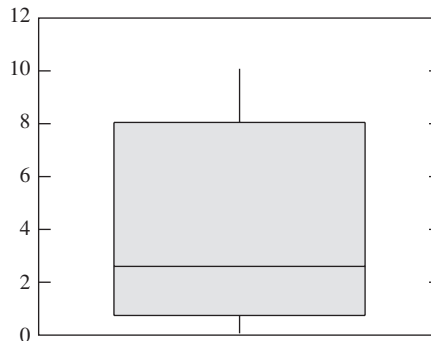


- c) Las mediciones en el grupo 1 son por lo general mayores que en el grupo 2. Las mediciones en el grupo 1 no están lejos de ser simétricas, aunque el diagrama de caja sugiere un ligero sesgo a la izquierda, puesto que la mediana está más cerca del tercer cuartil que del primero. No hay datos atípicos. La mayoría de las mediciones del grupo 2 están muy concentradas en el rango angosto y sesgadas a la izquierda dentro del rango. Las cuatro mediciones restantes son datos atípicos.

7. a) Cercano al 25% b) 130-135 mm c) 12%

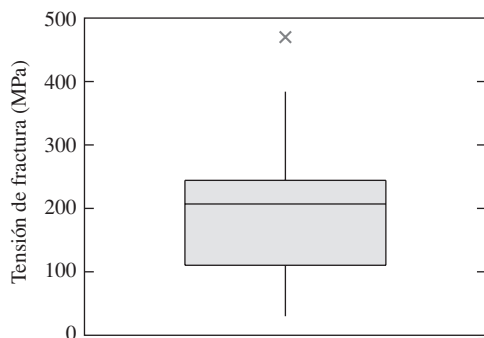
9. ii)

11. a) A: 4.60, B: 3.86 b) Sí



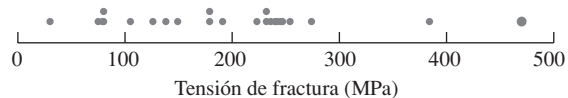
- c) No. El valor mínimo de -2.235 es un “dato atípico” ya que está a más de 1.5 veces del rango intercuartil debajo del primer cuartil. El bigote inferior se extiende al dato menor que no es un dato atípico, pero el valor del punto no está dado.

13. a)



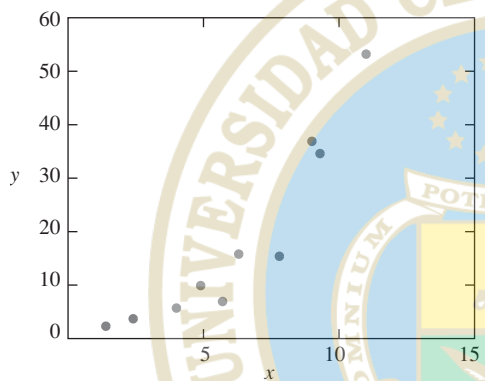
b) El diagrama de caja indica que el valor 470 es un dato atípico.

c)



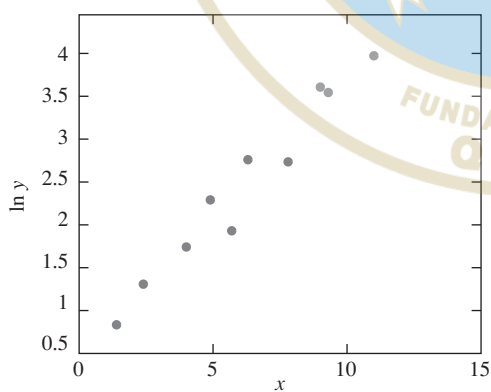
d) El diagrama de caja indica que el valor 384 está alejado del volumen de los datos y se pudiera considerar un dato atípico.

15. a)



La relación es no lineal.

b)



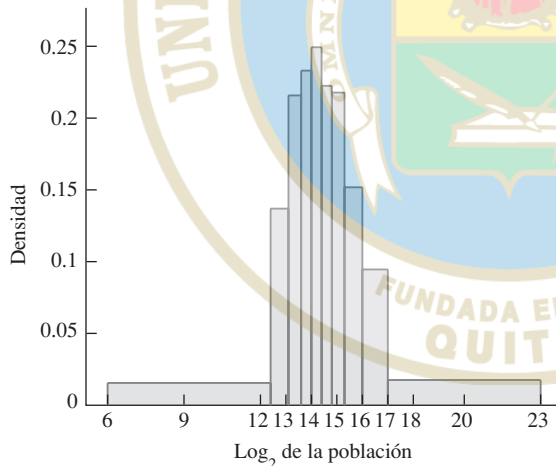
La relación es aproximadamente lineal.

c) Sería más fácil trabajar con x y $\ln y$, porque la relación es aproximadamente lineal.

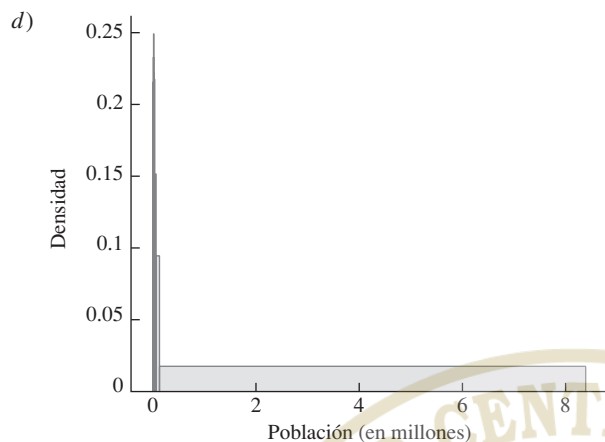
Ejercicios adicionales para el capítulo 1

1. a) La media se dividirá entre 2.2.
b) La desviación estándar estará dividida entre 2.2.
3. a) Falso b) Verdadero c) Falso d) Verdadero
5. a) No es posible decir cuánto cambia la media.
b) Si hay más de dos números aleatorios en la lista, la mediana no cambia. Si hay sólo dos números en la lista, la mediana cambia, pero no se puede decir cuánto.
c) No es posible decir cuánto cambia la desviación estándar.
7. a) La media disminuye por 0.774. b) La media cambia a 24.226.
c) La mediana no cambia.
d) No es posible decir cuánto cambia la desviación estándar.
9. El enunciado i) es verdadero.
11. a) Incorrecto b) Correcto c) Incorrecto d) Correcto
13. a) Sesgado a la izquierda. El 85o. percentil está mucho más cercano a la mediana (50o. percentil) que el 15o. percentil. Por tanto, el histograma parece tener una cola izquierda más larga que la cola derecha.
b) Sesgado a la derecha. El 15o. percentil está mucho más cercano a la mediana (50o. percentil) que el 85o. percentil. Por tanto, el histograma parece tener una cola derecha más larga que la cola izquierda.

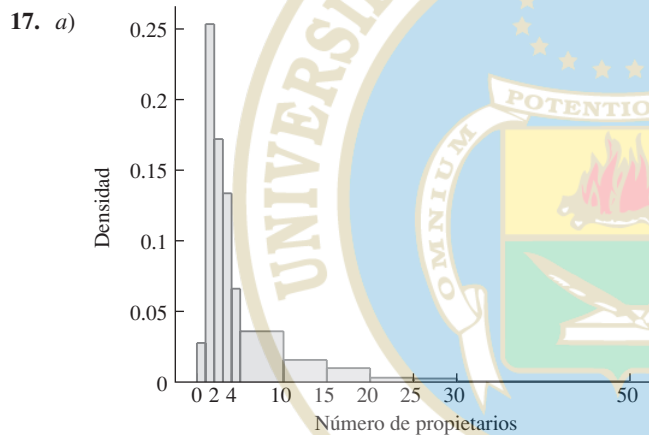
15. a)



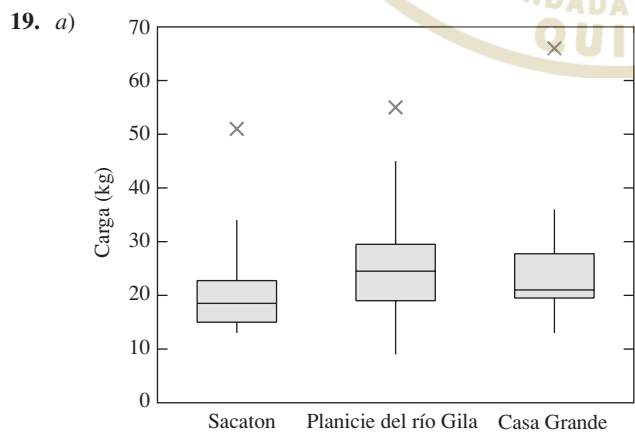
- b) 0.14
- c) Aproximadamente simétrico



Los datos usando la escala real están muy sesgados a la derecha, por lo que es imposible apreciar las características del histograma.



b) 3.35 c) 1.88 d) 7.70 e) 5.95 f) 43.34



- b) Cada muestra contiene un dato atípico.
- c) En el diagrama de caja de Sacaton, la mediana está aproximadamente a la mitad entre el primer y el tercer cuartiles, ello sugiere que los datos entre estos cuartiles son bastante simétricos. El bigote superior de la caja es mucho mayor que el inferior, y hay un dato atípico en el lado superior. Esto último indica que los datos, como un todo, están sesgados a la derecha. En los datos del diagrama de la planicie del río Gila, la mediana está aproximadamente entre el primer y el tercer cuartiles, lo que sugiere que los datos entre estos cuartiles son bastante simétricos. El bigote superior es ligeramente más largo que el inferior, y hay un dato atípico en el lado superior. Lo anterior sugiere que los datos, como un todo, están un poco sesgados a la derecha. En el diagrama de caja de Casa Grande la mediana está muy cerca del primer cuartil. Esto sugiere que hay varios valores muy cercanos entre sí, aproximadamente a un cuarto de distancia de los datos. Los dos bigotes tienen aproximadamente la misma longitud, lo que sugiere que las colas son aproximadamente iguales, excepto para el dato atípico en el lado superior.

Sección 2.1

1. 0.92
3. a) 0.6 b) 0.9
5. a) Falso b) Verdadero
7. 0.94
9. a) Falso b) Verdadero c) Falso d) Verdadero

Sección 2.2

1. a) 64 b) 8 c) 24
3. 210
5. 1 048 576
7. a) $36^8 = 2.8211 \times 10^{12}$ b) $36^8 - 26^8 = 2.6123 \times 10^{12}$ c) 0.9260
9. 0.5238

Sección 2.3

1. a) 2/10 b) 2/9 c) 1/9
3. Dado que un estudiante es un graduado de ingeniería, es casi cierto que el estudiante tomó un curso de cálculo. Por tanto $P(B|A)$ es cercano a 1. Dado que un estudiante tomó un curso de cálculo es mucho menos cierto que es un graduado de ingeniería, ya que muchos graduados de otras carreras toman cálculo. Por tanto, $P(A|B)$ es mucho menor que 1, por lo que $P(B|A) > P(A|B)$.
5. a) 0.018 b) 0.728 c) 0.272
7. a) 0.8 b) 0.125 c) 0.12 d) 0.167 e) 0.88 f) 0.205 g) 0.795
9. a) 0.98 b) 0.02 c) 0.72 d) 0.18
11. a) Los sistemas de medición fallan en forma independiente.
b) Una causa de falla, un incendio, ocasionaría que ambos sistemas de medición fallen. Por tanto, no fallan de manera independiente.

c) Demasiado baja. El cálculo correcto utilizaría P (segundo sistema de medición falla | primer sistema de medición falla) en lugar de P (segundo sistema de medición falla). Porque existe una posibilidad de que ambos sistemas fallen en un incendio, la condición de que el primer sistema de medición falle aumenta la probabilidad de que el segundo sistema de medición falle también. Por tanto, $P(\text{segundo sistema de medición falla} | \text{primer sistema de medición falla}) > P(\text{segundo sistema de medición falla})$.

13. a) $3/10$ b) $2/9$ c) $1/15$ d) $7/30$ e) $3/10$
f) No. $P(B) \neq P(B|A)$ [o $P(A \cap B) \neq P(A)P(B)$]

15. $n = 10\,000$. Los dos componentes son una muestra aleatoria simple proveniente de la población. Cuando la población es grande, los elementos en una muestra aleatoria simple son casi independientes.

17. a) 0.011 b) 0.0033

19. a) 5.08×10^{-5} b) 0.9801 c) 0.0001 d) 0.9801

21. a) 0.9904 b) 0.1 c) 0.2154 d) 7

23. Para demostrar que A^c y B son independientes, se muestra que $P(A^c \cap B) = P(A^c)P(B)$. Ahora $B = (A^c \cap B) \cup (A \cap B)$, y $(A^c \cap B)$ y $(A \cap B)$ son mutuamente excluyentes. Por tanto $P(B) = P(A^c \cap B) + P(A \cap B)$; en consecuencia, $P(A^c \cap B) = P(B) - P(A \cap B)$. Puesto que A y B son independientes, $P(A \cap B) = P(A)P(B)$. Por tanto, $P(A^c \cap B) = P(B) - P(A)P(B) = P(B)[1 - P(A)] = P(A^c)P(B)$. Para mostrar que A y B^c son independientes, basta con intercambiar A y B en el argumento anterior. Para mostrar que A^c y B^c son independientes, sustituya B con B^c en el argumento anterior, y utilice el hecho de que A y B^c son independientes.

Sección 2.4

1. a) Discreta b) Continua c) Discreta d) Continua e) Discreta

3. a) 2.3 b) 1.81 c) 1.345 d)

y	10	20	30	40	50
$p(y)$	0.4	0.2	0.2	0.1	0.1

 e) 23 f) 181 g) 13.45

5. a) $c = 0.1$ b) 0.2 c) 3 d) 1 e) 1

7. a)

x	$p_1(x)$
0	0.2
1	0.16
2	0.128
3	0.1024
4	0.0819
5	0.0655

 b)

x	$p_2(x)$
0	0.4
1	0.24
2	0.144
3	0.0864
4	0.0518
5	0.0311

c) $p_2(x)$ parece ser el mejor modelo. Todas sus probabilidades están bastante cercanas a las proporciones de los días observados en los datos. En contraste, las probabilidades de 0 y 1 para $p_1(x)$ son mucho menores que las proporciones observadas.

d) No, esto no es correcto. Los datos son una muestra aleatoria simple, y un modelo representa la población. Las muestras aleatorias simples por lo general no reflejan a la población exactamente.

9. a) 2 b) 0.81 c) 0.09 d) 0.9 e) 0.162

11. a) 0.444 b) $50 \mu\text{g}$ c) $14.142 \mu\text{g}$ d) $F(x) = \begin{cases} 0 & x < 10 \\ (x^2/2 - 10x + 50)/1800 & 10 \leq x < 70 \\ 1 & x \geq 70 \end{cases}$ e) $52.426 \mu\text{g}$

13. a) 5 segundos b) 5 segundos c) $F(t) = \begin{cases} 0 & t < 0 \\ 1 - e^{-0.2t} & t \geq 0 \end{cases}$ d) 0.8647 e) 3.466 segundos
f) 11.5129 segundos

15. Con este proceso la probabilidad de que un anillo satisfaga la especificación es 0.641. Con el proceso del ejercicio 14, la probabilidad es 0.568. Por tanto, este proceso es mejor que el del ejercicio 14.

17. a) 10 b) 2

Sección 2.5

1. a) $\mu = 21.0, \sigma = 1.0$ b) $\mu = 4.8, \sigma = 0.583$ c) $\mu = 42.9, \sigma = 1.62$
3. $\mu = 3500$ horas, $\sigma = 44.7$ horas
5. a) 0.625 pulg. b) 0.0112 pulg.
7. a) 0.650 b) 0.158
9. a) 150 cm b) 0.447 cm
11. a) 0.2993 b) 0.00288

Sección 2.6

1. a) 0.08 b) 0.36 c) 0.86 d) 0.35 e) 0.60 f) 0.65 g) 0.40
3. a) $p_{Y|X}(0|1) = 0.577, p_{Y|X}(1|1) = 0.308, p_{Y|X}(2|1) = 0.115$
b) $p_{X|Y}(0|1) = 0.522, p_{X|Y}(1|1) = 0.348, p_{X|Y}(2|1) = 0.130$
c) 0.538 d) 0.609
5. a) $\mu_{X+Y} = 3.90$ b) $\sigma_{X+Y} = 1.179$ c) $P(X+Y=4) = 0.35$
7. a) $100X + 200Y$ b) 595 ms c) 188.35
9. a) $p_X(0) = 0.10, p_X(1) = 0.20, p_X(2) = 0.30, p_X(3) = 0.25, p_X(4) = 0.15, p_X(x) = 0$ si $x \neq 0, 1, 2, 3, 4$
b) $p_Y(0) = 0.13, p_Y(1) = 0.21, p_Y(2) = 0.29, p_Y(3) = 0.22, p_Y(4) = 0.15, p_Y(y) = 0$ si $y \neq 0, 1, 2, 3, 4$
c) No. La probabilidad conjunta de la función de masa no es igual al producto de las marginales. Por ejemplo, $p_{X,Y}(0,0) = 0.05 \neq p_X(0)p_Y(0)$.
d) $\mu_X = 2.15, \mu_Y = 2.05$ e) $\sigma_X = 1.1948, \sigma_Y = 1.2440$
f) $\text{Cov}(X, Y) = 1.0525$ g) $\rho_{X,Y} = 0.7081$
11. a) $p_{Y|X}(0|3) = 0, p_{Y|X}(1|3) = 0.08, p_{Y|X}(2|3) = 0.32, p_{Y|X}(3|3) = 0.40, p_{Y|X}(4|3) = 0.20$
b) $p_{X|Y}(0|4) = 0, p_{X|Y}(1|4) = 0, p_{X|Y}(2|4) = 2/15, p_{X|Y}(3|4) = 1/3, p_{X|Y}(4|4) = 8/15$
c) $E(Y|X=3) = 2.72$ d) $E(X|Y=4) = 3.4$

13. a) $\mu_Z = 2.55$ b) $\sigma_Z = 1.4309$ c) $P(Z = 2) = 0.2$

15. a) $p_{Y|X}(0|3) = 0.25$, $p_{Y|X}(1|3) = 0.25$, $p_{Y|X}(2|3) = 0.5$
 b) $p_{X|Y}(0|1) = 0.125$, $p_{X|Y}(1|1) = 0.25$, $p_{X|Y}(2|1) = 0.5$, $p_{X|Y}(3|1) = 0.125$
 c) 1.25 d) 1.625

17. a) -0.000193 b) -0.00232

19. a) $f_X(x) = \begin{cases} x + 1/2 & 0 < x < 1 \\ 0 & \text{para otros valores de } x \end{cases}$ $f_Y(y) = \begin{cases} y + 1/2 & 0 < y < 1 \\ 0 & \text{para otros valores de } y \end{cases}$

b) $f_{Y|X}(y|0.75) = \begin{cases} (4y+3)/5 & 0 < y < 1 \\ 0 & \text{para otros valores de } y \end{cases}$

c) $E(Y|X = 0.75) = 0.5667$

21. a) 0.3 b) 0.45 c) 0.135 d) $\mu_X = 10$ e) $\mu_Y = 5$ f) $\mu_A = 50$

23. a) $f(x, y) = \begin{cases} e^{-x-y} & x > 0 \text{ y } y > 0 \\ 0 & \text{de otro modo} \end{cases}$

b) $P(X \leq 1 \text{ y } Y > 1) = e^{-1} - e^{-2} = 0.2325$

c) 1 d) 2 e) $1 - 3e^{-2} = 0.5940$

25. a) $\mu = 40.25$, $\sigma = 0.11$ b) $n \approx 52$

27. a) $0.3X + 0.7Y$ b) $\mu = \$6$, $\sigma = \$2.52$

c) $\mu = \$6$, $\sigma = 0.03\sqrt{1.4K^2 - 140K + 10\,000}$ d) $K = \$50$

e) Para cualquier correlación ρ , el riesgo es $0.03\sqrt{K^2 + (100 - K)^2 + 2\rho K(100 - K)}$. Si $\rho \neq 1$ la cantidad se minimiza cuando $K = 50$.

29. a) $\text{Cov}(aX, bY) = \mu_{aX \cdot bY} - \mu_{aX}\mu_{bY} = \mu_{abXY} - a\mu_X b\mu_Y = ab\mu_{XY} - ab\mu_X\mu_Y = ab(\mu_{XY} - \mu_X\mu_Y) = ab\text{Cov}(X, Y)$.

b) $\rho_{aX, bY} = \text{Cov}(aX, bY)/(\sigma_{aX}\sigma_{bY}) = ab\text{Cov}(X, Y)/(ab\sigma_X\sigma_Y) = \text{Cov}(X, Y)/(\sigma_X\sigma_Y) = \rho_{X, Y}$.

Ejercicios adicionales para el capítulo 2

1. 0.9997

3. a) 0.15 b) 0.6667

5. 0.271

7. 0.82

9. 1/3

11. a) 0.3996 b) 0.0821 c) $f_X(x) = \begin{cases} \frac{1}{2}e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases}$ d) $f_Y(y) = \begin{cases} \frac{1}{3}e^{-y/3} & y > 0 \\ 0 & y \leq 0 \end{cases}$

e) Sí, $f(x, y) = f_X(x)f_Y(y)$.

13. a) 0.0436 b) 0.0114 c) 0.7377

15. 1/3

17. a) $\mu = 6$, $\sigma^2 = 9$ b) $\mu = 4$, $\sigma^2 = 10$ c) $\mu = 0$, $\sigma^2 = 10$ d) $\mu = 16$, $\sigma^2 = 328$

19. a) Para la concentración de aditivo: $p(0.02) = 0.22$, $p(0.04) = 0.19$, $p(0.06) = 0.29$, $p(0.08) = 0.30$ y $p(x) = 0$ para $x \neq 0.02, 0.04, 0.06$ o 0.08 . Para la fuerza de tensión: $p(100) = 0.14$, $p(150) = 0.36$, $p(200) = 0.50$ y $p(x) = 0$ para $x \neq 100, 150$ o 200 .
- b) No, X y Y no son independientes. Por ejemplo, $P(X = 0.02 \cap Y = 100) = 0.05$, pero $P(X = 0.02)P(Y = 100) = (0.22)(0.14) = 0.0308$. c) 0.947 d) 0.867 e) La concentración sería 0.06.
21. a) $p_{Y|X}(100|0.06) = 0.138$, $p_{Y|X}(150|0.06) = 0.276$, $p_{Y|X}(200|0.06) = 0.586$
 b) $p_{X|Y}(0.02|100) = 0.357$, $p_{X|Y}(0.04|100) = 0.071$, $p_{X|Y}(0.06|100) = 0.286$, $p_{X|Y}(0.08|100) = 0.286$
 c) 172.4 d) 0.0500
23. a) $\mu = 3.75$, $\sigma = 6.68$ b) $\mu = 2.90$, $\sigma = 4.91$ c) $\mu = 1.08$, $\sigma = 1.81$
 d) Bajo el escenario A, 0.85; bajo el escenario B, 0.89 y bajo el escenario C, 0.99.

25. a) La probabilidad conjunta de la función de masa es

x	y		
	0	1	2
0	0.0667	0.2000	0.0667
1	0.2667	0.2667	0
2	0.1333	0	0

- b) 0.8 c) 0.6 d) 0.6532 e) 0.6110 f) -0.2133 g) -0.5345
27. a) $\mu_X = 9/14 = 0.6429$ b) $\sigma_X^2 = 199/2940 = 0.06769$ c) $\text{Cov}(X, Y) = -5/588 = -0.008503$
 d) $\rho_{X,Y} = -25/199 = -0.1256$
29. a) $p_X(0) = 0.6$, $p_X(1) = 0.4$, $p_X(x) = 0$ si $x \neq 0$ o 1 . b) $p_Y(0) = 0.4$, $p_Y(1) = 0.6$, $p_Y(y) = 0$ si $y \neq 0$ o 1 .
 c) Sí. Es razonable suponer que el conocimiento del resultado de una moneda no ayudará a predecir el resultado de la otra.
 d) $p(0, 0) = 0.24$, $p(0, 1) = 0.36$, $p(1, 0) = 0.16$, $p(1, 1) = 0.24$, $p(x, y) = 0$ para otros valores de (x, y) .
31. a) $p_{X,Y}(x, y) = 1/9$ para $x = 1, 2, 3$ y $y = 1, 2, 3$. b) $p_X(1) = p_X(2) = p_X(3) = 1/3$. p_Y es igual.
 c) $\mu_X = \mu_Y = 2$ d) $\mu_{XY} = 4$ e) $\text{Cov}(X, Y) = 0$.

Sección 3.1

1. ii)
3. a) Verdadero b) Falso c) Falso d) Verdadero
5. a) No, no se puede determinar la desviación estándar del proceso a partir de una sola medición.
 b) Sí, el sesgo se puede calcular igual a 2 lb, ya que la lectura es 2 lb cuando el peso verdadero es 0.
7. a) Sí, la incertidumbre se puede calcular con la desviación estándar de las cinco mediciones, que es $21.3 \mu\text{g}$.
 b) No, el sesgo no se puede calcular, porque se desconoce el valor verdadero.
9. Se puede tener una medición más precisa restando el sesgo de $26.2 \mu\text{g}$, obteniendo $100.8 \mu\text{g}$ arriba de 1 kg.
11. a) No, están en orden creciente, lo que no es casi nada común para una muestra aleatoria simple.
 b) No, puesto que no son una muestra aleatoria simple de una población de posibles mediciones, no se puede calcular la incertidumbre.

Sección 3.2

1. a) 0.6 b) 0.45 c) 1.3
3. 1.01 ± 0.30 mm
5. 378.0 ± 0.5 pulg³
7. 1.04 ± 0.04 cm²/mol
9. a) $64.04 \pm 0.39^\circ\text{F}$ b) $64.04 \pm 0.11^\circ\text{F}$
11. a) La incertidumbre en el promedio de las nueve mediciones es aproximadamente igual a $s/\sqrt{9} = 0.081/3 = 0.027$ cm.
b) La incertidumbre en una sola medición es aproximadamente igual a s , que es 0.081 cm.
13. a) 87.0 ± 0.7 mL b) 0.5 mL c) 25
15. a) A 65°C , la producción es 70.14 ± 0.28 . A 80°C , la producción es 90.50 ± 0.25 . b) 20.36 ± 0.38
17. a) 0.016 b) 0.0089
c) La incertidumbre en $\frac{1}{2}\bar{X} + \frac{1}{2}\bar{Y}$ es 0.0091. La incertidumbre en $\frac{10}{15}\bar{X} + \frac{5}{15}\bar{Y}$ es 0.011.
d) $c = 0.24$; la incertidumbre mínima es 0.0078.

Sección 3.3

1. a) 3.2 b) 0.1 c) 0.025 d) 0.1 e) 21.8 f) 0.26
3. 9.80 ± 0.39 m/s²
5. a) 1.7289 ± 0.0058 s b) 9.79 ± 0.11 m/s²
7. a) 0.2555 ± 0.0005 m/s b) 0.256 ± 0.026 m/s c) 0.2555 ± 0.0002 m/s
9. a) 2.3946 ± 0.0011 g/mL
11. a) 0.27% b) 0.037% c) 4.0% d) 0.5%
13. $9.802 \text{ m/s}^2 \pm 6.0\%$
15. a) $1.856 \text{ s} \pm 0.29\%$ b) $9.799 \text{ m/s}^2 \pm 0.54\%$
17. a) $0.2513 \text{ m/s} \pm 0.33\%$ b) $0.2513 \text{ m/s} \pm 2.0\%$ c) $0.2513 \text{ m/s} \pm 0.57\%$
19. $2.484 \text{ g/mL} \pm 0.19\%$

Sección 3.4

1. a) 250 ± 16 b) 125 ± 10 c) 17.50 ± 0.56
3. a) 14.25 ± 0.25 MPa b) Reduciendo la incertidumbre en P_1 a 0.2 MPa.
5. a) 1.320 ± 0.075 cm b) Reduciendo la incertidumbre en p a 0.1 cm.
7. 0.259 ± 0.014 m/s

9. a) 32.6 ± 3.4 MPa b) Reduciendo la incertidumbre en k a 0.025 mm^{-1} .
 c) La implementación del proceso reduciría la incertidumbre en τ sólo a 3.0 MPa. Es probable que no valga la pena implementar el nuevo proceso para una reducción tan pequeña.
11. a) 710.68 ± 0.15 g b) Reduciendo la incertidumbre en b a 0.1 g.
13. a) $2\,264 \pm 608 \text{ N/mm}^2$ b) R
15. $0.0626 \pm 0.0013 \text{ min}^{-1}$
17. a) No, ambos implican las cantidades h y r . b) $2.68c \pm 0.27c$
19. $283.49 \text{ mm/s} \pm 2.5\%$
21. $1.41 \text{ cm} \pm 6.4\%$
23. $0.487 \pm 1.7\%$
25. $3\,347.9 \text{ N/mm}^2 \pm 29\%$
27. a) $17.59 \text{ } \mu\text{m} \pm 18\%$ b) $5.965 \text{ } \mu\text{m}^3 \pm 30\%$ c) $2.95c \pm 11\%$ d) No

Ejercicios adicionales para el capítulo 3

1. a) 1.8 b) 4.1 c) 15 d) 4.3
3. a) 0.14 mm b) 0.035 mm
5. a) $(1.854 \pm 0.073) \times 10^6 \text{ W}$ b) 3.9% c) Reduciendo la incertidumbre en H a 0.05.
7. a) $6.57 \pm 0.17 \text{ kcal}$ b) 2.6% c) Reduciendo la incertidumbre en la masa a 0.005 g.
9. a) $26.32 \pm 0.33 \text{ mm/año}$ b) $3.799 \pm 0.048 \text{ años}$
11. $5.70 \pm 0.17 \text{ mm}$
13. a) 1.4% b) Reduciendo la incertidumbre en l a 0.5%
15. a) Sí, la fuerza calculada es 80 000 lb en ambos casos.
 b) No, en el método de cable dúctil los cuadrados de las incertidumbres de los 16 alambres se suman, para obtener $\sigma = \sqrt{16 \times 20^2} = 80$. Para el método de cable quebradizo, la incertidumbre en la fuerza del alambre más débil se multiplica por el número de alambres para obtener $\sigma = 16 \times 20 = 320$.
17. a) $113.1 \pm 6.1 \text{ m}^3/\text{s}$ b) $100.5 \pm 5.4 \text{ m}^3/\text{s}$ c) Sí, la incertidumbre relativa es 5.4%.
19. a) $10.04 \pm 0.95 \text{ s}^{-1}$ b) $10.4 \pm 1.2 \text{ s}^{-1}$ c) 0.78 d) 0.63
21. a) $32\,833 \pm 36 \text{ m}^2$ b) $12\,894 \pm 14 \text{ m}^2$
 c) Esto no es correcto. Sea s la longitud de un lado del cuadrado. Ya que tanto S como C se calculan en función de s , no son independientes. Con el fin de comparar a σ_A en forma correcta, se debe expresar A directamente en función de s : $A = s^2 + 2\pi s^2/8 = s^2(1 + \pi/4)$. Por lo que $\sigma_A = (dA/ds)\sigma_s = 2s(1 + \pi/4)\sigma_s = 65 \text{ m}^2$.
23. a) $P_3 = 11.16871 \pm 0.10 \text{ MPa}$ b) 11.16916
 c) No. La diferencia entre los dos cálculos es mucho menor que la incertidumbre.

Sección 4.1

1. a) $\mu = 0.55, \sigma^2 = 0.2475$
 b) No. Una variable aleatoria de Bernoulli tiene valores posibles 0 y 1. Los valores posibles de Y son 0 y 2.
 c) $\mu = 1.10, \sigma^2 = 0.99$
3. a) 0.05 b) 0.20 c) 0.23 d) Sí e) No
 f) No. Si la superficie se decolora y se agrieta, entonces $X = 1, Y = 1$ y $Z = 1$, pero $X + Y = 2$.
5. a) 1/2 b) 1/2 c) 1/4 d) Sí e) Sí
 f) Sí, si ambas monedas salen "caras", entonces $X = 1, Y = 1$ y $Z = 1$, por lo que $Z = XY$. Si no, entonces $Z = 0$, y ya sea X, Y , o ambas, también son iguales a 0, por lo que nuevamente $Z = XY$.
7. a) Puesto que los valores posibles de X y Y son 0 y 1, los valores posibles del producto $Z = XY$ son también 0 y 1. Por tanto, Z es una variable aleatoria de Bernoulli.
 b) $p_Z = P(Z = 1) = P(XY = 1) = P(X = 1 \text{ y } Y = 1) = P(X = 1)P(Y = 1) = p_X p_Y$.

Sección 4.2

1. a) 0.2090 b) 0.2322 c) 0.1064 d) 0.0085 e) 3.2 f) 1.92
3. a) 0.1172 b) 5 c) 2.5 d) 1.58
5. a) 0.0039 b) 0.2188 c) 0.1445 d) 0.9648
7. a) 0.120 ± 0.032 b) 0.050 ± 0.015 c) 0.070 ± 0.036
9. a) 0.96 b) 0.0582
11. a) 1.346×10^{-4}
 b) Sí, sólo aproximadamente 13 o 14 de cada 100 000 muestras de tamaño 10 tendrían siete o más unidades defectuosas.
 c) Sí, debido a que siete unidades en una muestra de tamaño 10 es un número inusualmente grande para un buen cargamento.
 d) 0.4557 e) No, en aproximadamente 45% de las muestras de tamaño 10, dos o más unidades estarían defectuosas.
 f) No, debido a que dos defectuosas en una muestra de tamaño 10 no es un número inusualmente grande para un buen cargamento.
13. a) $Y = 7X + 300$ b) \$930 c) \$21
15. 0.225 ± 0.064

Sección 4.3

1. a) 0.0733 b) 0.0183 c) 0.0916 d) 0.9084 e) 4 f) 2
3. a) 0.2240 b) 0.4232 c) 0.5974 d) 3 e) 1.73
5. a) 0.0916 b) 0.1048 c) 0.2381
7. ii)
9. 78 ± 12
11. a) 12.5 b) 7.0 c) 2.5 d) 1.9 e) 5.5 ± 3.1

13. a) 7.295×10^{-3} b) Si la media de la concentración es siete partículas por mL, entonces sólo aproximadamente siete de cada mil muestras de 1 mL contendrían una o menos partículas.
 c) Sí, debido a que una partícula en una muestra de 1 mL es un número inusualmente pequeño si la media de la concentración es siete partículas por mL. d) 0.4497 e) No. Si la media de la concentración es siete partículas por mL, entonces aproximadamente 45% de todas las muestras de 1 mL contendrán seis o menos partículas.
 f) No, debido a que seis partículas en una muestra de 1 mL no es un número inusualmente pequeño si la media de la concentración es de siete partículas por mL.
15. 0.271 ± 0.019

Sección 4.4

1. 0.4196
 3. 0.0314
 5. a) 0.1244 b) 7.5 c) 11.25
 7. iv)
 9. a) 0.3 b) $\mu_x = 1.2$ c) 0.7483
 11. a) 0.00890 b) 0.1275

Sección 4.5

1. a) 0.8023 b) 0.2478 c) 0.4338 d) 0.7404
 3. a) 0.0073 b) ≈ 420 c) 91o. percentil d) 0.4186
 5. a) 0.0764 b) 9.062 GPa c) 12.303 GPa
 7. a) 0.0336 b) Sí, la proporción de días se suspende; en este caso sería de sólo 0.0228.
 9. a) 0.06 cm b) 0.01458 cm c) 0.2451 d) 0.0502 cm e) 0.7352
 f) El diámetro de la perforación tendría una media de 15.02 cm. La probabilidad de satisfacer la especificación será entonces de 0.8294.
 11. a) 0.0475 b) 12.07 oz c) 0.0215 oz
 13. a) 7.8125 N/m^2 b) 4.292 N/m^2 c) 76.65 N/m^2
 15. a) 0.7088 b) 0.2912 c) 0.0485
 17. a) la media es 114.8 J; la desviación estándar es 5.006 J.
 b) Sí, sólo 0.15% de los pernos tendrían torsiones de ruptura menores de 100 J.
 c) La media es 117.08 J; la desviación estándar es 8.295 J. Aproximadamente 2% de pernos tendrían torsiones de ruptura menores de 100 J, por lo que el cargamento no sería aceptado.
 d) Los pernos del inciso c) son más resistentes.
 e) El método efectivamente no es válido para los pernos del inciso c). Esta muestra contiene un dato atípico (140), por lo que la distribución normal no se debe usar.

Sección 4.6

1. a) 3.5966 b) 0.5293 c) 3.3201 d) 5.5400
3. a) 25.212 b) 3.9828 c) 24.903 d) 0.2148 e) 27.666
5. a) 46.711 N/mm b) 33.348 N/mm c) Clavos con rosca anular. La probabilidad es 0.3372 contra 0.0516 para clavos con rosca en espiral. d) 0.0985. e) Un clavo con rosca en espiral. Sólo aproximadamente 0.01% de los clavos con rosca anular tienen resistencias así de pequeñas, mientras que aproximadamente 4.09% de los clavos con rosca en espiral la tienen. Se puede tener la certeza de que éste fue un clavo con rosca en espiral.
7. a) \$1.0565 b) 0.0934 c) \$1.0408 d) 0.2090

Sección 4.7

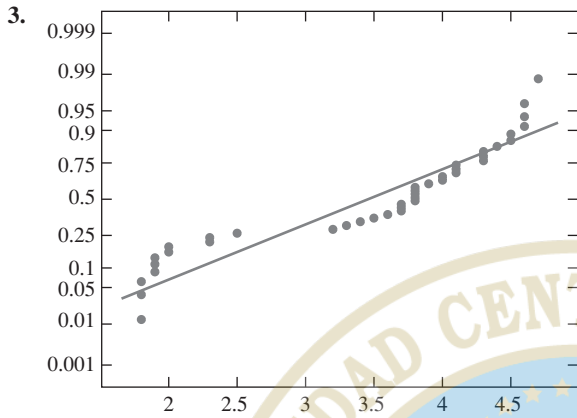
1. a) 2.2222 b) 4.9383 c) 0.2592 d) 1.5403
3. a) 4 micrones b) 4 micrones c) 0.5276 d) 0.0639 e) 2.7726 micrones f) 5.5452 micrones
g) 18.4207 micrones
5. No. Si las duraciones estuvieran distribuidas exponencialmente, la proporción de los componentes usados que duran más de cinco años sería la misma que la proporción de los nuevos componentes que duran más de cinco años, debido a la propiedad de falta de memoria.
7. a) 1/3 año b) 1/3 año c) 0.0498 d) 0.2212 e) 0.9502
9. a) 0.6065 b) 0.0821 c) El tiempo del primer reemplazo sería mayor de 100 horas si y sólo si cada uno de los focos durara más de 100 horas. d) 0.9179 e) $P(T \leq t) = 1 - e^{-0.025t}$.
f) Sí, $T \sim \text{Exp}(0.025)$, g) 40 horas h) $T \sim \text{Exp}(n\lambda)$

Sección 4.8

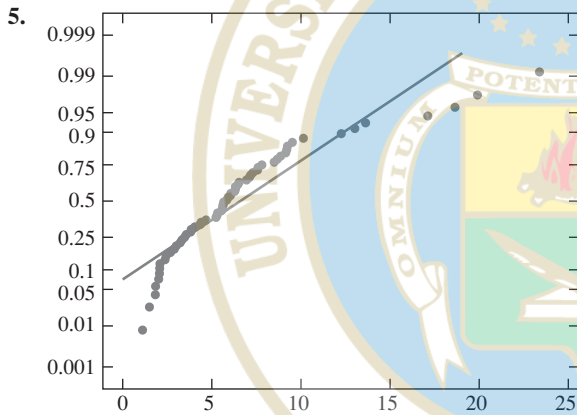
1. a) 8 b) 4 c) 0.00175 d) 0.9344
3. a) 0.6667 b) 1.4907 c) 0.8231 d) 0.0208 e) 0.0550
5. a) 0.8490 b) 0.5410 c) 1 899.2 horas d) 8.761×10^{-4}
7. a) 0.3679 b) 0.2978 c) 0.4227
9. a) 0.3679 b) 0.1353
- c) La duración del sistema será mayor de cinco horas si y sólo si las duraciones de ambos componentes fueran mayores de cinco horas. d) 0.8647 e) $P(T \leq t) = 1 - e^{-0.08t}$ f) Sí, $T \sim \text{Weibull}(2, 0.2828)$

Sección 4.9

1. a) No b) No c) Sí



Estos datos no parecen provenir de una distribución aproximadamente normal.



Los datos de MP no parecen provenir de una distribución aproximadamente normal.

7. Sí, si los logs de los datos de MP provienen de una población normal, entonces los datos de MP provienen de una población lognormal, y viceversa.

Sección 4.10

1. a) 0.0793 b) 0.0170
3. 0.9772
5. 0.5793
7. a) 0.0951 b) 0.0344 c) 0.181 (0.278 es una raíz espuria.)
9. a) 0.6578 b) 0.4714 c) 0.6266 d) 48.02 mL
11. a) 0.0002 b) Sí. Sólo aproximadamente dos de 10 000 muestras de tamaño 1 000 tendrán 75 o más losas no satisfactorias si se ha alcanzado el objetivo. c) No, debido a que 75 losas no satisfactorias en una muestra de 1 000 es un número inusualmente grande si se ha alcanzado el objetivo.

d) 0.3594 e) No. Más de un tercio de las muestras de tamaño 1 000 tendrán 53 o más losas no satisfactorias si se ha alcanzado el objetivo. f) Sí, debido a que 53 de las losas no satisfactorias en una muestra de 1 000 no es un número inusualmente grande si se ha alcanzado el objetivo.

Sección 4.11

1. a) $X \sim \text{Bin}(100, 0.03)$, $Y \sim \text{Bin}(100, 0.05)$ b) Las respuestas variarán. c) ≈ 0.72 d) ≈ 0.18
e) La distribución se desvía un poco de lo normal.
3. a) $\mu_A = 6$ exactamente (los resultados de la simulación serán aproximados), $\sigma_A^2 \approx 0.25$.
b) ≈ 0.16 c) La distribución es aproximadamente normal.
5. a) ≈ 0.25 b) ≈ 0.25 c) ≈ 0.61
7. a, b, c) Las respuestas variarán. d) ≈ 0.025
9. a) Las respuestas variarán. b) ≈ 2.7 c) ≈ 0.34 d) ≈ 1.6
e) El tiempo de vida del sistema no está distribuido aproximadamente normal. f) Sesgado a la derecha.
11. a) Las respuestas variarán. b) $\approx 10\,090$ c) $\approx 1\,250$ d) ≈ 0.58 e) ≈ 0.095
f) La distribución difiere un poco de la normal.
13. a) $\hat{\lambda} = 0.25616$ b, c, d) Las respuestas variarán. e) Sesgo ≈ 0.037 , $\sigma_{\hat{\lambda}} \approx 0.12$.

Ejercicios adicionales para el capítulo 4

1. 0.9744
3. a) 0.2503 b) 0.4744 c) 0.1020 d) 0.1414 e) 0.8508
5. a) 0.9044 b) 0.00427 c) 0.00512
7. a) 0.6826 b) $z = 1.28$ c) 0.0010
9. a) 0.0668 b) 0.6687 c) 0.0508
11. a) 0.8830 b) 0.4013 c) 0.0390 (0.1065 es una raíz espuria.)
13. a) 28.0 ± 3.7 b) 28 mL
15. a) 0.0749 b) 4.7910 cm c) 4
17. a) 0.4889 b) 0.8679
19. 62
21. a) 0.4090 b) No. Más de 40% de las muestras tendrán un peso total de 914.8 oz o menos si la afirmación es cierta.
c) No, debido a que un peso total de 914.8 oz no es inusualmente pequeño si la afirmación es cierta.
d) ≈ 0 e) Sí. Casi ninguna de las muestras tendrá un peso total de 910.3 oz o menos si la afirmación es verdadera.
f) Sí, porque un peso total de 910.3 oz es inusualmente pequeño si la afirmación es cierta.
23. a) $P(X > s) = P(\text{Primeras } s \text{ tiradas son fracasos}) = (1 - p)^s$.
b) $P(X > s + t | X > s) = P(X > s + t \text{ y } X > s) / P(X > s) = P(X > s + t) / P(X > s) = (1 - p)^{s+t} / (1 - p)^s = (1 - p)^t = P(X > t)$. Observe que si $X > s + t$, éste debe ser el caso que $X > s$, que es la razón de que $P(X > s + t \text{ y } X > s) = P(X > s + t)$.

- c) Sea X el número de lanzamientos de la moneda de un centavo necesario para obtener la primera cara. Entonces $P(X > 5 | X > 3) = P(X > 2) = 1/4$. La probabilidad de que la moneda de cinco centavos salga “cruz” dos veces también es $1/4$.

$$25. a) \frac{P(X=x)}{P(X=x-1)} = \frac{e^{-\lambda} \lambda^x / x!}{e^{-\lambda} \lambda^{x-1} / (x-1)!} = \frac{e^{-\lambda} \lambda^x (x-1)!}{e^{-\lambda} \lambda^{x-1} x!} = \frac{\lambda}{x}.$$

$$b) P(X=x) \geq P(X=x-1) \text{ si y sólo si } \frac{\lambda}{x} \geq 1 \text{ si y sólo si } x \leq \lambda.$$

Sección 5.1

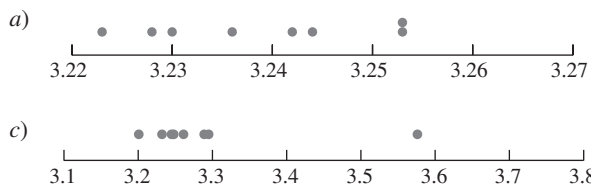
1. a) 1.645 b) 1.37 c) 2.81 d) 1.15
3. Aumenta, disminuye
5. a) (145.10, 154.90) b) (143.55, 156.45) c) 76.98% d) 601 e) 1041
7. a) (1.538, 1.582) b) (1.534, 1.586) c) 92.66% d) 385 e) 543
9. a) (27.04, 30.96) b) (26.42, 31.58) c) 86.64% d) 312 e) 540
11. a) 22.62 b) 99.38%
13. a) 70.33 b) 99.29%
15. 280
17. (83.11, 84.19)
19. a) Falso b) Verdadero c) Falso
21. El supervisor está subestimando la confianza. El enunciado de que la media del costo es menor de \$160 es un límite superior de confianza unilateral con un nivel de confianza de 97.5%.

Sección 5.2

1. a) 0.74 b) (0.603, 0.842) c) 74 d) (0.565, 0.879) e) 130
3. a) (0.629, 0.831) b) (0.645, 0.815) c) 300 d) 210 e) 0.0217
5. 0.339
7. a) (0.0529, 0.1055) b) 697 c) (0.0008, 0.556)
9. a) (0.490, 0.727) b) (0.863, 1) c) (0.572, 0.906)
11. a) (0.271, 0.382) b) 658 c) 748

Sección 5.3

1. a) 1.860 b) 2.776 c) 2.763 d) 12.706
3. a) 95% b) 98% c) 90% d) 99% e) 99.9%
5. (5.303, 6.497)
7. Sí, no hay datos atípicos. Un intervalo de confianza de 95% es (203.81, 206.45).



b) Sí, el intervalo de confianza es (3.2247, 3.2526)

d) No, el conjunto de datos contiene un dato atípico.

11. (12.318, 13.762)

13. (0.515, 1.985)

15. a) 2.3541 b) 0.888 c) 3.900

Sección 5.4

1. (5.589, 8.211)

3. (9.683, 11.597)

5. (11.018, 32.982)

7. a) (−1.789, 2.589)

b) No, ya que 0 se encuentra en el intervalo de confianza, se podría considerar que es un valor creíble para la diferencia de las medias de dureza.

9. No es posible. Los tiempos en cama despiertos y dormidos no son independientes.

Sección 5.5

1. (0.0591, 0.208)

3. (−0.232, 0.00148)

5. a) (−0.0446, 0.103) b) Si se muestrearan 100 chips más del proceso menos costoso, el ancho del intervalo de confianza sería aproximadamente ± 0.0721 . Si se hubiesen muestreado 50 chips más del proceso más costoso, el ancho del intervalo de confianza sería aproximadamente ± 0.0638 . Si se hubiesen muestreado 50 chips más del proceso menos costoso y 25 chips más del proceso más costoso, el ancho del intervalo de confianza sería aproximadamente ± 0.0670 . Por tanto, el mayor aumento en la precisión se lograría muestreando 50 chips más del proceso más costoso.

7. No. Las proporciones de la muestra provienen de la muestra en vez de dos muestras independientes.

9. (−0.0481, 0.226)

11. No, éstas no son muestras aleatorias simples.

Sección 5.6

1. $(-0.0978, 0.00975)$
3. $(7.798, 30.602)$
5. $(20.278, 25.922)$
7. $(0.0447, 0.173)$
9. $(2.500, 33.671)$
11. $(38.931, 132.244)$
13. $(0.614, 11.386)$

Sección 5.7

1. $(2.090, 11.384)$
3. $(-0.0456, 0.00558)$
5. $(-30.260, 26.260)$
7. $(9.350, 10.939)$
9. a) $(0.747, 2.742)$ b) 80%

Sección 5.8

1. a) $X^* \sim N(8.5, 0.2^2)$, $Y^* \sim N(21.2, 0.3^2)$ b) Las respuestas variarán. c) $\sigma_P \approx 0.18429$
 d) Sí, P está de manera normal aproximadamente distribuida. e) $(13.063, 13.785)$ si la propagación del error se utiliza para encontrar la desviación estándar.
3. a) Sí, A está de manera normal aproximadamente distribuida. b) $\sigma_A \approx 0.24042$ c) $(6.140, 7.083)$ si la propagación del error se utiliza para encontrar la desviación estándar.
5. a) $N(0.27, 0.40^2/349)$ y $N(1.62, 1.70^2/143)$. Debido a que los valores 0.27 y 1.62 son las medias muestrales, sus varianzas son iguales a las varianzas poblacionales divididas entre los tamaños de la muestra. b) No, R está de manera normal aproximadamente distribuida. c) $\sigma_R \approx 0.70966$ d) No es adecuado, ya que R está de manera normal aproximadamente distribuida.
7. a, b, c) Las respuestas variarán.
9. a) La probabilidad de cobertura de Agresti-Coull ≈ 0.98 ; para el intervalo tradicional ≈ 0.89 . La media del ancho para Agresti-Coull ≈ 0.51 ; para el intervalo tradicional ≈ 0.585 . b) La probabilidad de cobertura de Agresti-Coull ≈ 0.95 ; para el intervalo tradicional ≈ 0.95 . La media del ancho para Agresti-Coull ≈ 0.42 ; para el intervalo tradicional ≈ 0.46 . c) La probabilidad de cobertura de Agresti-Coull ≈ 0.96 ; para el intervalo tradicional ≈ 0.92 . La media del ancho para Agresti-Coull ≈ 0.29 ; para el intervalo tradicional ≈ 0.305 . d) El método tradicional tiene una probabilidad de cobertura cercana a 0.95 para $n = 17$, pero menor que 0.95 tanto para $n = 10$ como para $n = 40$. e) Agresti-Coull tiene mayor probabilidad de cobertura para tamaños de muestra de 10 y de 40, casi igual que para 17. f) El método de Agresti-Coull.

Ejercicios adicionales para el capítulo 5

1. (1.942, 19.725)
3. (0.0374, 0.0667)
5. (0.084, 0.516)
7. a) (0.0886, 0.241) b) 584
9. El intervalo más angosto (4.20, 5.83), es el intervalo de confianza de 90%, el intervalo más ancho, (3.57, 6.46), es el intervalo de confianza de 99%, y (4.01, 6.02) es el intervalo de confianza de 95%.
11. (-0.420, 0.238)
13. 441
15. a) Falso b) Falso c) Verdadero d) Falso
17. a) (36.804, 37.196) b) 68% c) Las mediciones provienen de una población normal.
d) (36.774, 37.226)
19. a) Debido a que X se distribuye normalmente con media $n\lambda$, se tiene que para una proporción $1 - \alpha$ de todas las muestras posibles, $-z_{\alpha/2}\sigma_X < X - n\lambda < z_{\alpha/2}\sigma_X$. Multiplicando por -1 y sumando a X en toda la desigualdad se obtiene $X - z_{\alpha/2}\sigma_X < n\lambda < X + z_{\alpha/2}\sigma_X$, que es el resultado deseado.
b) Ya que n es constante, $\sigma_{X/n} = \sigma_X/n = \sqrt{n\lambda}/n = \sqrt{\lambda/n}$. Por tanto $\sigma_{\hat{\lambda}} = \sigma_{X/n}$.
c) Se divide la desigualdad del inciso a) entre n .
d) Sustituyendo $\sqrt{\hat{\lambda}/n}$ en lugar de $\sigma_{\hat{\lambda}}$ en el inciso c) se demuestra que para una proporción $1 - \alpha$ de todas las posibles muestras $\hat{\lambda} - z_{\alpha/2}\sqrt{\hat{\lambda}/n} < \lambda < \hat{\lambda} + z_{\alpha/2}\sqrt{\hat{\lambda}/n}$. Por tanto, el intervalo $\hat{\lambda} \pm z_{\alpha/2}\sqrt{\hat{\lambda}/n}$ es un intervalo de confianza para λ a un nivel $1 - \alpha$.
e) (53.210, 66.790)
21. a) 234.375 ± 19.639 b) (195.883, 272.867)
c) Hay algo de desviación de la normalidad en las colas de la distribución. El 95% intermedio sigue la curva normal cercanamente, por lo que el intervalo de confianza es razonablemente bueno.
23. a, b, c) Las respuestas variarán.

Sección 6.1

1. a) 0.0014 b) 0.14%
3. a) 0.2584 b) 25.84%
5. a) ≈ 0 b) Si la media de la producción diaria fuera 740 toneladas o más, la probabilidad de observar una media muestral tan pequeña como el valor de 715 que realmente fue observado es casi 0. Por tanto, se tiene el convencimiento de que la media de la producción diaria no es de 740 toneladas o más, si no que, por el contrario, es menor.
7. a) 0.3300 b) Si la media de la velocidad del aire es 40 cm/s, hay 33% de probabilidad de que una muestra tenga una media menor o igual que el valor observado de 39.6. Debido a que 33% no es una probabilidad pequeña, es creíble que la media de la velocidad del aire sea 40 cm/s.
9. ii)
11. $P = 0.0456$
13. a) 0.2153 b) 2.65 c) 0.0040

Sección 6.2

1. $P = 0.5$
3. iv)
5. a) Verdadero b) Falso c) Verdadero d) Falso
7. a) No. El P -valor es 0.196, que es mayor que 0.05.
b) La media muestral es 73.2461. La media muestral es, por tanto, más cercana a 73 que a 73.5. El P -valor para la hipótesis nula $\mu = 73$, por consiguiente, será mayor que el P -valor para la hipótesis nula $\mu = 73.5$, que es 0.196. Por tanto, la hipótesis nula $\mu = 73$ no se puede rechazar al nivel de 5%.
9. a) $H_0: \mu \leq 8$ b) $H_0: \mu \leq 60\,000$ c) $H_0: \mu = 10$
11. a) ii) La balanza está fuera de calibración. b) iii) La balanza podría estar calibrada. c) No. La balanza está calibrada sólo si $\mu = 10$. La evidencia más fuerte en favor de esta hipótesis ocurriría si $\bar{X} = 10$, pero puesto que hay incertidumbre en $\bar{X}$, no se puede tener la seguridad aun cuando $\mu = 10$.
13. No, ella no puede concluir que la hipótesis nula sea verdadera.
15. i)
17. a) Sí. Las cantidades mayores que el límite superior de confianza tendrán P -valores menores que 0.05. Por tanto, $P < 0.05$.
b) No, se necesitaría conocer el límite superior de confianza del 99% para determinar si $P < 0.01$.
19. Si, se puede calcular el P -valor exactamente. Ya que 95% del límite superior de confianza es 3.45, se sabe que $3.40 + 1.645s/\sqrt{n} = 3.45$. Por tanto, $s/\sqrt{n} = 0.0304$. El puntaje z es $(3.40 - 3.50)/0.0304 = -3.29$. El P -valor es 0.0005, que es menor que 0.01.

Sección 6.3

1. Sí, $P = 0.0040$.
3. No, $P = 0.1251$.
5. No, $P = 0.2033$.
7. No, $P = 0.1251$.
9. Sí, $P = 0.0011$.
11. a) 0.69 b) -0.49 c) 0.3121

Sección 6.4

1. a) $t_2 = 0.6547$, $0.50 < P < 0.80$. La balanza podría estar calibrada adecuadamente.
b) No se puede realizar la prueba t , ya que la desviación estándar de la muestra no se puede calcular a partir de una muestra de tamaño 1.
3. a) $H_0: \mu \leq 35$ vs. $H_1: \mu > 35$ b) $t_5 = 2.4495$, $0.025 < P < 0.050$
c) Sí, el P -valor es pequeño, por lo que se concluye que $\mu > 35$.

5. No, $t_6 = 1.7085$, $0.10 < P < 0.20$.
7. Sí, $t_{17} = -2.4244$, $0.01 < P < 0.025$.
9. Sí, $t_3 = -4.0032$, $0.01 < P < 0.025$.
11. a) 6.0989 b) 9.190 c) 17.384 d) -1.48

Sección 6.5

1. Sí, $P = 0.0002$
3. No, $P = 0.1336$
5. a) $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_1: \mu_1 - \mu_2 > 0$, $P = 0.2119$. No se puede concluir que la calificación promedio de las preguntas de una cola sea mayor.
 b) $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_1: \mu_1 - \mu_2 \neq 0$, $P = 0.4238$. No se puede concluir que la calificación promedio de las preguntas de una cola difiera de la calificación promedio de las preguntas de dos colas.
7. a) Sí, $P = 0.0217$. b) No, $P = 0.2514$.
9. Sí, $P = 0.0006$.
11. a) i) 11.128, ii) 0.380484 b) 0.424, similar al P -valor calculado con el estadístico t .
 c) $(-0.3967, 5.7367)$

Sección 6.6

1. a) $H_0: p_1 - p_2 \geq 0$ vs. $H_1: p_1 - p_2 < 0$ b) $P = 0.1492$ c) Máquina 1
3. Sí, $P = 0.0018$.
5. Sí, $P = 0.0367$.
7. $P = 0.0643$. La evidencia sugiere que el empaque de gran resistencia reduce la proporción de órdenes deterioradas, pero no puede ser concluyente.
9. No, $P = 0.7114$.
11. No, porque las dos muestras no son independientes.
13. a) 0.660131 b) 49 c) 1.79 d) 0.073

Sección 6.7

1. a) Sí, $t_3 = 2.5740$, $0.025 < P < 0.050$. b) No, $t_3 = 0.5148$, $0.25 < P < 0.40$.
3. $t_{16} = -2.0143$, $0.05 < P < 0.10$. La hipótesis nula es sospechosa.
5. No, $t_7 = 0.3444$, $0.50 < P < 0.80$.
7. No, $t_{15} = 1.0024$, $0.10 < P < 0.25$.
9. Sí, $t_9 = 2.5615$, $0.02 < P < 0.05$.

11. a) Sí, $t_{16} = 6.1113$, $P < 0.0005$. b) Sí, $t_{16} = 3.659$, $0.001 < P < 0.005$.
13. No, $t_{14} = 1.0236$, $0.20 < P < 0.50$.
15. a) 0.197 b) 0.339 c) -1.484 d) -6.805

Sección 6.8

1. Sí, $t_6 = 7.0711$, $P < 0.001$.
3. Sí, $t_9 = 2.6434$, $0.02 < P < 0.05$.
5. Sí, $t_4 = 4.7900$, $0.001 < P < 0.005$.
7. Sí, $t_6 = 3.7591$, $0.002 < P < 0.010$.
9. a) Sea μ_R la media del número de millas por galón para taxis que utilizan neumáticos radiales, y sea μ_B la media del número de millas por galón para taxis que utilizan neumáticos con cinturón. Las hipótesis nula y alternativa adecuadas son $H_0: \mu_R - \mu_B \leq 0$ contra $H_1: \mu_R - \mu_B > 0$. El valor del estadístico de prueba es $t_9 = 8.9532$, por lo que $P < 0.0005$.
- b) Las hipótesis nula y alternativa adecuadas son $H_0: \mu_R - \mu_B \leq 2$ contra $H_1: \mu_R - \mu_B > 2$. El valor del estadístico de prueba es $t_9 = 3.3749$, por lo que $0.001 < P < 0.005$.
11. a) 1.1050 b) 2.8479 c) 4.0665 d) 3.40

Sección 6.9

1. a) Sí. $S_+ = 25$, $P = 0.0391$. b) No. $S_+ = 7$, $P > 0.1094$.
c) No. $S_+ = 23$, $P > 2(0.0781) = 0.1562$.
3. a) No. $S_+ = 134$, $z = -0.46$, $P = 0.3228$. b) Sí. $S_+ = 249.5$, $z = 2.84$, $P = 0.0023$.
c) Sí. $S_+ = 70.5$, $z = -2.27$, $P = 0.0232$.

5. Diferencia	0.01	0.01	-0.01	0.03	0.05	-0.05	-0.07	-0.11	-0.13	0.15
Signo del rango	2	2	-2	4	5.5	-5.5	-7	-8	-9	10

$S_+ = 2 + 2 + 4 + 5.5 + 10 = 23.5$. De la tabla, $P > 2(0.1162) = 0.2324$. No se rechaza.

7. Sí. $W = 34$, $P = 2(0.0087) = 0.0174$.

9. No. $W = 168$, $z = 0.31$, $P = 0.7566$.

Sección 6.10

1. a) $H_0: p_1 = 0.90$, $p_2 = 0.05$, $p_3 = 0.05$ b) 900, 50, 50 c) $\chi^2_2 = 21.7778$
d) $P < 0.005$. Los verdaderos porcentajes difieren de 90%, 5% y 5%.

3. Los valores esperados son

	Capacidad excesiva neta				
	< 0%	0-10%	11-20%	21-30%	> 30%
Pequeña	81.7572	44.7180	14.0026	4.9687	27.5535
Grande	99.2428	54.2820	16.9974	6.0313	33.4465

$\chi^2_4 = 12.9451$, $0.01 < P < 0.05$. Es razonable concluir que las distribuciones difieren.

5. a) 10.30 13.35 13.35
 6.96 9.02 9.02
 9.74 12.62 12.62
 b) $\chi^2_4 = 6.4808, P > 0.10$. No hay evidencia de que los renglones y las columnas no sean independientes.
7. iii)
9. $\chi^2_9 = 8.0000, P > 0.10$. No hay evidencia de que el generador de números aleatorios no esté funcionando adecuadamente.
11. Sí, $\chi^2_{11} = 41.3289, P < 0.005$.

Sección 6.11

1. 251
3. a) 0.01 b) 0.02
5. No, $F_{8,13} = 1.0429, P > 0.20$.

Sección 6.12

1. a) Verdadero b) Falso c) Falso
3. a) $H_0: \mu \geq 90$ contra $H_1: \mu < 90$ b) $\bar{X} < 89.3284$
 c) Ésta no es una región de rechazo adecuada. La región de rechazo consistiría de los valores para $\bar{X}$ que haría el P -valor de la prueba menor que una frecuencia de umbral elegida. Esta región de rechazo consiste de los valores para los cuales el P -valor será mayor que algún nivel.
 d) Ésta es una región de rechazo adecuada. El nivel de la prueba es 0.0708.
 e) Ésta no es una región de rechazo adecuada. La región de rechazo consiste de los valores para $\bar{X}$ que harían el P -valor de la prueba menor que un nivel de umbral elegida. La región de rechazo contiene los valores de $\bar{X}$ para los cuales el P -valor será grande.
5. a) Error de tipo I b) Decisión correcta c) Decisión correcta d) Error de tipo II
7. Al nivel 1%

Sección 6.13

1. a) Verdadero b) Verdadero c) Falso d) Falso
3. Aumenta
5. a) $H_0: \mu \geq 50\,000$ contra $H_1: \mu < 50\,000$ es verdadera. b) El nivel es 0.1151; la potencia es 0.4207.
 c) 0.2578 d) 0.4364 e) 618
7. ii)
9. a) Dos colas b) $p = 0.5$ c) $p = 0.4$
 d) Menor que 0.7. La potencia para un tamaño de muestra de 150 es 0.691332, y la potencia para un tamaño de muestra menor de 100 sería menor que éste.
 e) Mayor que 0.6. La potencia para un tamaño de muestra de 150 es 0.691332, y la potencia para un tamaño de muestra mayor de 100 sería mayor que éste.
 f) Mayor que 0.65. La potencia contra la alternativa $p = 0.4$ es 0.691332, y la alternativa $p = 0.3$ está más alejada de la nula que $p = 0.4$. Por lo que la potencia contra la alternativa $p = 0.3$ es mayor que 0.691332.

- g) Es imposible decirlo del resultado. La potencia contra la alternativa $p = 0.45$ será menor que la potencia contra $p = 0.4$ que es 0.691332. Pero no se puede decir sin calcularlo si es menor de 0.65.
11. a) Dos colas b) Menos de 0.9. El tamaño de la muestra de 60 es el menor que produce la potencia mayor que o igual a la potencia objetivo de 0.9.
- c) Mayor de 0.9. La potencia es mayor de 0.9 contra una diferencia de 3, por lo que será mayor que 0.9 contra cualquier diferencia mayor que 3.

Sección 6.14

1. a) El P -valor con el ajuste de Bonferroni es 0.012. Puesto que este valor es pequeño, se puede concluir que este ajuste reduce la proporción de piezas defectuosas.
- b) El P -valor con el ajuste de Bonferroni es 0.018. Puesto que este valor no es tan pequeño, no se puede concluir que este ajuste reduzca la proporción de piezas defectuosas.
3. 0.0025
5. a) No. Si la media del amperaje de quemada es igual a 15 A cada día, la probabilidad de rechazar H_0 es 0.05 cada día. El número de veces en 200 días que H_0 es rechazada es entonces una variable aleatoria binomial con $n = 200$, $p = 0.05$. La probabilidad de rechazar H_0 10 o más veces en 200 días es entonces aproximadamente igual a 0.5636. Por lo que no sería inusual rechazar H_0 10 veces en 200 ensayos si H_0 es siempre verdadera.
- b) Sí. Si la media del amperaje de quemada es igual a 15 A cada día, la probabilidad de rechazar H_0 es 0.05 cada día. El número de veces en 200 días que H_0 es rechazada es entonces una variable aleatoria binomial con $n = 200$, $p = 0.05$. La probabilidad de rechazar H_0 20 o más veces en 200 días es entonces aproximadamente igual a 0.0010. Por lo que sería bastante inusual rechazar H_0 20 veces en 200 ensayos si H_0 es siempre verdadera.

Sección 6.15

1. a) $V = 26.323$, $\sigma_V = 0.3342$ b) $z = 3.96$, $P \approx 0$. c) Sí, V está aproximadamente normalmente distribuida.
3. a) ii) y iv) b) i), ii) y iv)
5. No, el valor 103 es un dato atípico.
7. a) $s_A^2 = 200.28$, $s_B^2 = 39.833$, $s_A^2/s_B^2 = 5.02$. b) No, la prueba F requiere la suposición de que los datos están normalmente distribuidos. Estos datos contienen un dato atípico (103), por lo que no se debe usar la prueba F . c) $P \approx 37$.
9. a) El estadístico de prueba es $t = \frac{\bar{X} - 7}{s/\sqrt{7}}$. H_0 se rechazará si $|t| > 2.447$. b) ≈ 0.60 .

Ejercicios adicionales para el capítulo 6

1. Éste requiere una prueba para la diferencia entre dos medias. Los datos no están apareados. Sea μ_1 la media poblacional del costo anual para automóviles que usan combustible regular, y μ_2 la media poblacional del costo anual para automóviles que usan combustible premium. Entonces las hipótesis nula y alternativa adecuadas son $H_0 : \mu_1 - \mu_2 \geq 0$ contra $H_1 : \mu_1 - \mu_2 < 0$. El estadístico de prueba es la diferencia en la media muestral de los costos entre los dos grupos. La tabla z se debe utilizar para encontrar el P -valor.
3. Éste requiere una prueba para una proporción poblacional. Sea p la proporción poblacional de las piezas defectuosas bajo el nuevo programa. Las hipótesis nula y alternativa adecuadas son $H_0 : p \geq 0.10$ contra $H_1 : p < 0.10$. El estadístico de prueba es una proporción muestral de las piezas defectuosas. La tabla z se debe utilizar para determinar el P -valor.

5. a) $H_0: \mu \geq 16$ contra $H_1: \mu < 16$ b) $t_9 = -2.7388$ c) $0.01 < P < 0.025$, rechaza a H_0 .
7. a) $H_0: \mu_1 - \mu_2 = 0$ contra $H_1: \mu_1 - \mu_2 \neq 0$ b) $t_6 = 2.1187$ c) $0.05 < P < 0.10$, H_0 es sospechosa.
9. Sí. $z = 4.61$, $P \approx 0$.
11. a) Rechaza a H_0 si $\bar{X} \geq 100.0196$ o si $\bar{X} \leq 99.9804$. b) Rechaza a H_0 si $\bar{X} \geq 100.01645$ o si $\bar{X} \leq 99.98355$.
c) Sí d) No e) 13.36%
13. a) 0.05 b) 0.1094
15. El P -valor con el ajuste de Bonferroni es 0.1228. No se puede concluir que la tasa de fallas en la línea 3 sea menor de 0.10.
17. a) Ambas muestras tienen una mediana de 20. b) $W = 281.5$, $z = 2.03$, $P = 0.0424$. El P -valor es bastante pequeño. Si la hipótesis nula establece que las medianas poblacionales eran iguales, proporcionaría una razonable fuerte evidencia de que las medianas poblacionales eran efectivamente diferentes.
c) No, la muestra X está pesadamente sesgada a la derecha, mientras que la muestra Y es fuertemente bimodal. No parece razonable suponer que estas muestras provengan de poblaciones con la misma forma.
19. a) Sea μ_A la media del cociente empuje/peso para el combustible A, y sea μ_B la media del cociente empuje/peso para el combustible B. Las hipótesis nula y alternativa adecuadas son $H_0: \mu_A - \mu_B \leq 0$ contra $H_1: \mu_A - \mu_B > 0$.
b) Sí, $t_{29} = 2.0339$, $0.025 < P < 0.05$.
21. a) Sí b) La conclusión no está justificada. El ingeniero está concluyendo que H_0 es verdadera debido a que la prueba falla al rechazarla.
23. No, $\chi^2_2 = 2.1228$, $P > 0.10$.

Sección 7.1

1. 0.8214
3. a) El coeficiente de correlación es adecuado. Los puntos están aproximadamente agrupados alrededor de una recta.
b) El coeficiente de correlación no es adecuado. La relación es una curva, no es lineal.
c) El coeficiente de correlación no es adecuado. La gráfica contiene datos atípicos.
5. Más de 0.6
7. a) Entre la temperatura y la producción, $r = 0.7323$; entre la tasa de agitación y la producción, $r = 0.7513$; entre la temperatura y la tasa de agitación, $r = 0.9064$.
b) No, el resultado se podría deber a la confusión, ya que la correlación entre la temperatura y la tasa de agitación está alejada de 0.
c) No, el resultado se podría deber a la confusión, ya que la correlación entre la temperatura y la tasa de agitación está alejada de 0.
9. a) $(-0.8632, 0.1739)$ b) $z = -2.31$, $P = -0.0104$, se puede concluir que $\rho < 0.3$.
c) $U = -1.6849$, $0.10 < P < 0.20$, es creíble que $\rho = 0$.
11. a) $y = 2$ b) $y = -3$ c) $y = -1.1699$ d) $y = -9.8301$ e) Para que la correlación sea igual a -1 , los puntos se encontrarían en una recta con pendiente negativa. No existe un valor para y para este caso.

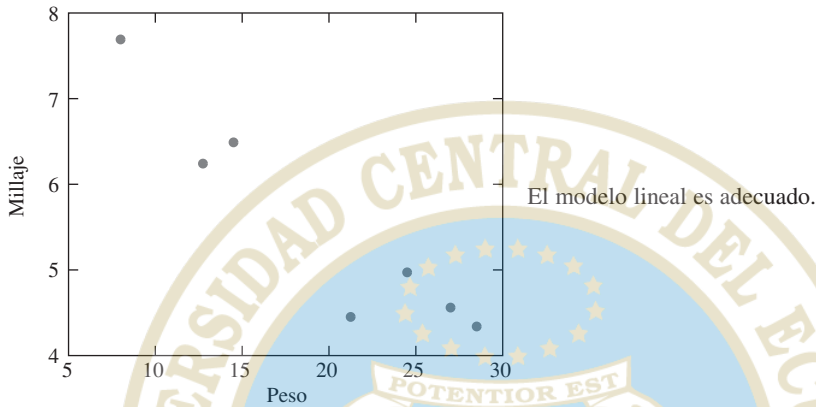
Sección 7.2

1. a) 319.27 lb b) 5.65 lb

3. 0.8492

5. a) 18.869 pulg. b) 70.477 pulg. c) No, algunos de los hombres cuyos puntos se encuentran debajo de la recta de mínimos cuadrados tendrán brazos más cortos.

7. a)



b) $y = 8.5593 - 0.15513x$.

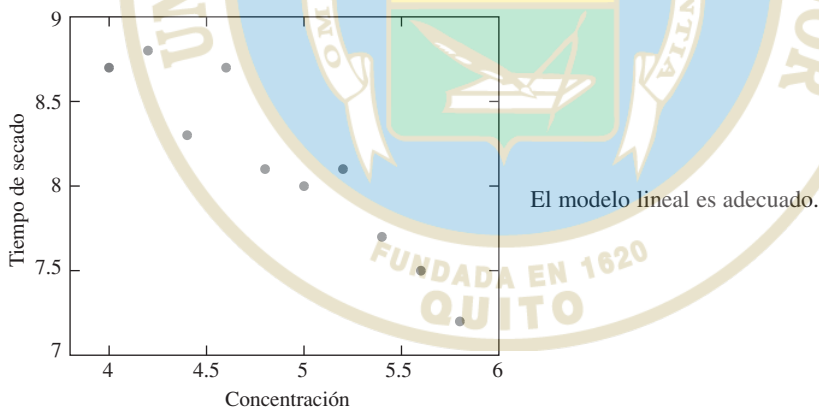
f) Millas por galón

c) 0.776 millas/galón

d) 6.23 millas/galón

e) Millas por galón por tonelada

7. a)



b) $y = 12.193 - 0.833x$ c) (8.86, -0.16), (8.69, 0.11), (8.53, -0.23), (8.36, 0.34), (8.19, -0.09), (8.03, -0.03), (7.86, 0.24), (7.69, 0.01), (7.53, -0.03), (7.36, -0.16) d) Disminuye en 0.0833 horas.

e) 8.53 horas f) No, ya que el 7% está fuera del rango de concentraciones presentes en los datos. g) 4.79%

h) No se puede especificar esa concentración. De acuerdo con la recta de mínimos cuadrados, una concentración de 7.43% daría como resultado un tiempo de secado de seis horas. Sin embargo, puesto que 7.43% está fuera del rango de los datos, esta predicción no es confiable.

11. $y = 20 + 10x$

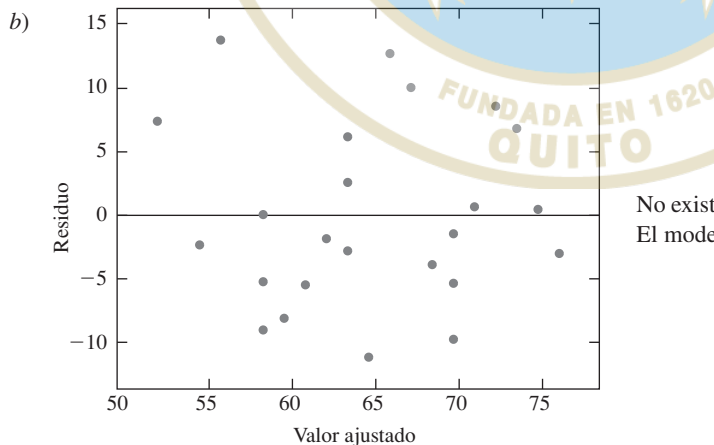
13. iii)

Sección 7.3

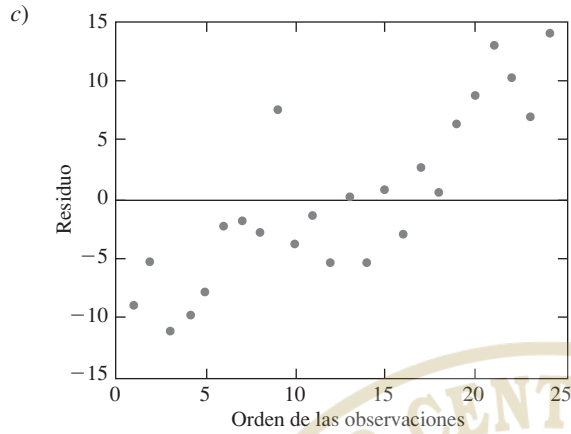
1. a) $\hat{\beta}_0 = 4.7829, \hat{\beta}_1 = 3.0821$ b) 0.2169 c) Para β_0 : (4.409, 5.157) para β_1 : (2.948, 3.216)
d) No. $t_{23} = 1.271, 0.10 < P < 0.25$. e) Sí. $t_{23} = -3.965, P < 0.0005$. f) (9.098, 9.714)
g) (8.063, 10.749) h) El intervalo de confianza es más útil ya que considera la longitud verdadera del resorte, mientras que el intervalo de predicción considera la siguiente medición de la longitud.
3. a) La pendiente es 0.84451; el intercepto es 44.534.
b) Sí, el P -valor para la pendiente es ≈ 0 , por lo que el caballo de fuerza está relacionado con NO_x .
c) 52.979 mg/s
d) 0.920 e) (74.6, 82.0) f) No. Un razonable rango de los valores pronosticados está dado por el intervalo de predicción del 95%, que es (29.37, 127.26).
5. a) $H_0: \beta_C - \beta_E = 0$ b) Sí. $z = 2.10, P = 0.0358$
7. a) $y = 2.11 - 0.776x$ b) Para β_0 : (2.07, 2.16) para β_1 : (-0.829, -0.722). c) 0.795
d) (0.742, 0.847) e) (0.675, 0.914)
9. El intervalo de confianza en 1.5 sería más corto. El intervalo de confianza en 1.8 sería más largo.
11. 1.388
13. a) 0.256 b) 0.80 c) 1.13448 d) 0.001
15. a) 553.71 b) 162.06 c) Debajo
d) Hay mayor cantidad de dispersión vertical en el lado derecho de la gráfica que en el lado izquierdo.

Sección 7.4

1. a) $\ln y = -0.4442 + 0.79833 \ln x$ b) 330.95 c) 231.96 d) (53.19, 1 009.89)
3. a) $y = 20.162 + 1.269x$

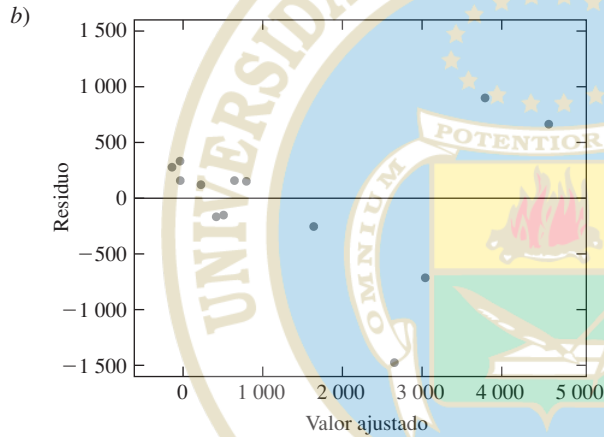


No existe un patrón aparente en la gráfica de residuos.
El modelo lineal es adecuado.



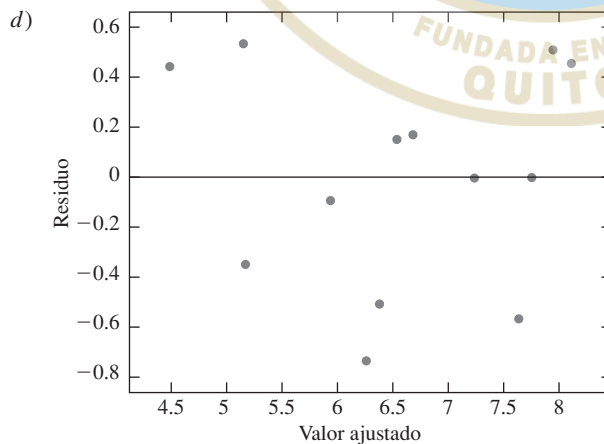
Los residuos aumentan con el tiempo. El modelo lineal no es adecuado. El tiempo, u otras variables relacionadas con el tiempo, se deben incluir en el modelo.

5. a) $y = -235.32 + 0.695x$.



La gráfica de residuos muestra un patrón con residuos positivos en los valores ajustados altos y bajos y residuos negativos en medio. El modelo no es adecuado.

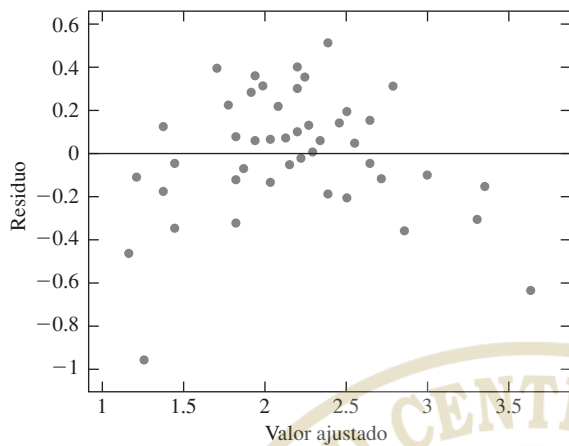
c) $\ln y = -0.0745 + 0.925 \ln x$.



La gráfica de residuos no muestra un patrón obvio. El modelo es adecuado.

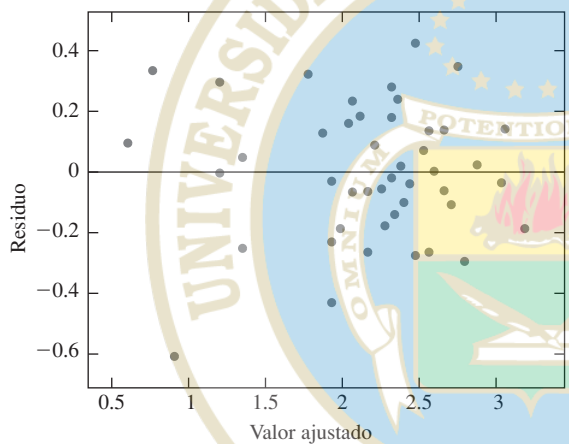
e) El modelo log es más adecuado. El intervalo de predicción del 95% es (197.26, 1 559.76).

a)



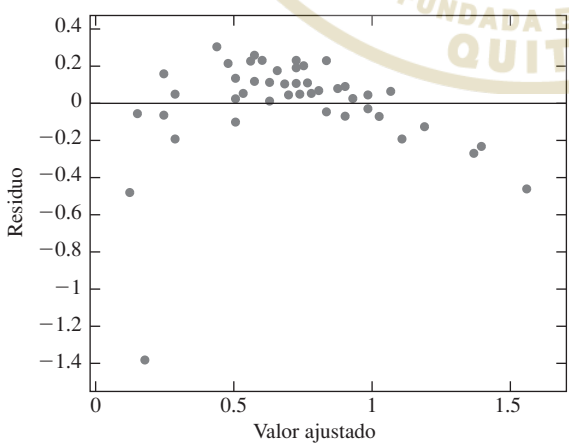
La recta de mínimos cuadrados es
 $y = 0.833 + 0.235x$.

b)



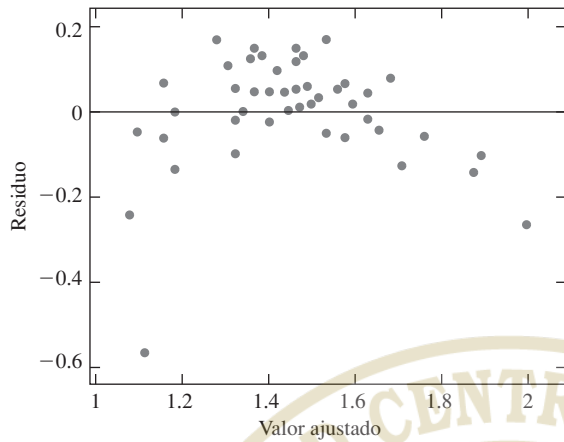
La recta de mínimos cuadrados es
 $y = 0.199 + 1.207 \ln x$.

c)



La recta de mínimos cuadrados es
 $y = -0.0679 + 0.137x$.

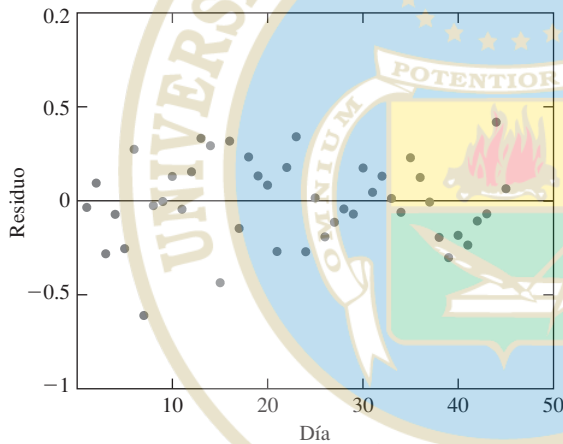
d)



La recta de mínimos cuadrados es $\sqrt{y} = 0.956 + 0.0874x$.

e) El modelo $y = 0.199 + 1.207 \ln x$ ajusta mejor. Su gráfica de residuos muestra el patrón mínimo.

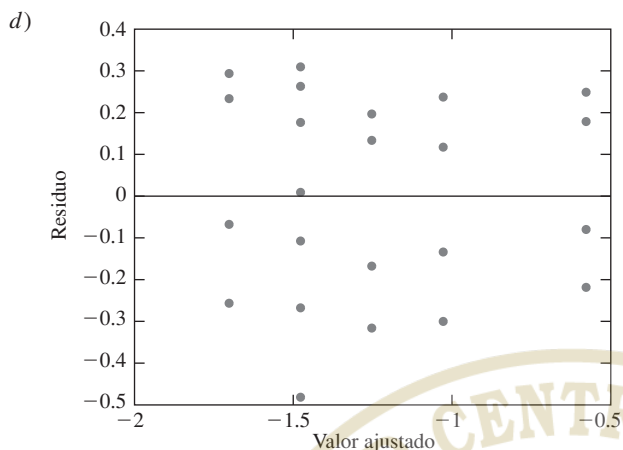
f)



Los residuos no muestran un patrón con el tiempo.

g) 2.14 h) (1.689, 2.594)

9. a) El modelo es $\log_{10} y = \beta_0 + \beta_1 \log_{10} x + \varepsilon$. Observe que el logaritmo natural ($\ln$) se podría usar en lugar de $\log_{10}$, pero los logaritmos comunes son más convenientes, ya que las presiones parciales se expresan como potencias de 10.



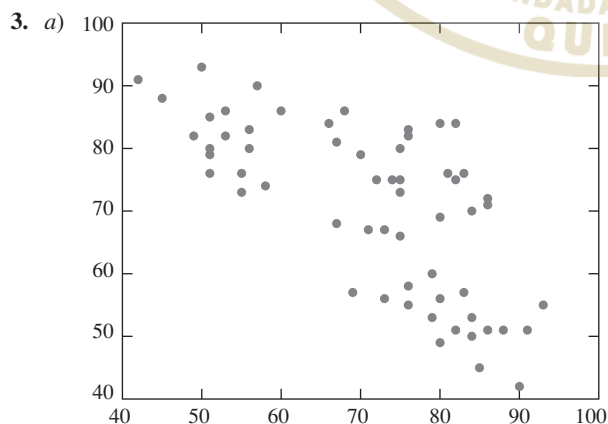
La recta de mínimos cuadrados es $\log_{10} y = -3.277 - 0.225 \log_{10} x$. El modelo lineal parece ajustar bastante bien.

c) La teoría establece que el coeficiente β_1 de $\log_{10} x$ en el modelo lineal es igual a -0.25 . El valor estimado es $\hat{\beta} = -0.225$. Se determina si los datos son consistentes con la teoría mediante la prueba de hipótesis $H_0 : \beta_1 = -0.25$ contra $H_1 : \beta_1 \neq -0.25$. El valor del estadístico de prueba es $t_{21} = 0.821$, por lo que $0.20 < P < 0.50$. No se rechaza H_0 , por lo que los datos son consistentes con la teoría.

11. a) $y = 2\,049.87 - 4.270x$ b) (12, 2 046) y (13, 1 954) son datos atípicos. La recta de mínimos cuadrados con (12, 2 046) eliminado es $y = 2\,021.85 - 2.861x$. La recta de mínimos cuadrados con (13, 1 954) eliminado es $y = 2\,069.30 - 5.236x$. La recta de mínimos cuadrados con ambos datos atípicos eliminados es $y = 2\,040.88 - 3.809x$.
 c) Las pendientes de las rectas de mínimos cuadrados están notoriamente afectadas por los datos atípicos. Su rango es de -2.861 a -5.236 .
13. a) Una ley física. b) Sería mejor repetir el experimento. Si los resultados de un experimento violan una ley física, entonces algo está equivocado en el experimento, y no puede fijar éste al transformar las variables.

Ejercicios adicionales para el capítulo 7

11. a) $\hat{\beta}_0 = -0.0390$, $\hat{\beta}_1 = 1.017$ b) 0 c) El coeficiente de absorción molar M . d) Probando $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, $t_3 = -0.428$ y $0.50 < P < 0.80$, por lo que los datos son consistentes con la ley de Beer-Lambert.



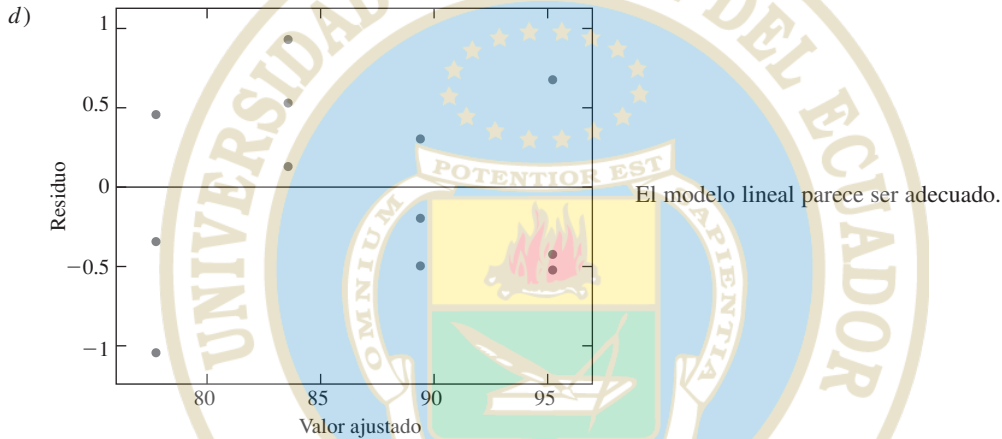
- b) $T_{i+1} = 120.18 - 0.696T_i$ c) $(-0.888, -0.503)$ d) 71.48 minutos e) $(68.40, 74.56)$
 f) $(45.00, 97.95)$

5. a) $\hat{\beta}_0 = 0.8182, \hat{\beta}_1 = 0.9418$ b) No. $t_9 = 1.274, 0.20 < P < 0.50$. c) Sí. $t_9 = -5.358, P < 0.001$.
 d) Sí, puesto que se puede concluir que $\beta_1 \neq 1$, se puede concluir que la máquina está fuera de calibración.
 e) $(18.58, 20.73)$ f) $(75.09, 73.23)$ g) No, cuando el valor verdadero es 20, el resultado del inciso e) muestra que un intervalo de confianza de 95% para la media de los valores medidos es $(18.58, 20.73)$. Por tanto, es creíble que la media de la medición será 20, por lo que la máquina está calibrada.

7. a) $y = 0.041496 + 0.0073664x$ b) $(-0.00018, 0.01492)$ c) $(0.145, 0.232)$ d) $(0.0576, 0.320)$

9. a) $\ln y = \beta_0 + \beta_1 \ln x$, donde $\beta_0 = \ln k$ y $\beta_1 = r$
 b) La recta de mínimos cuadrados es $\ln y = -1.7058 + 0.65033 \ln x$. Por tanto, $\hat{r} = 0.65033$ y $\hat{k} = e^{-1.7058} = 0.18162$.
 c) $T_3 = 4.660, P = 0.019$. No, esto no es creíble.

11. a) $\hat{\beta}_0 = 60.263, \hat{\beta}_1 = 0.11653, s^2 = 0.38660$. b) Sí. $t_{10} = 80.956, P \approx 0$. c) Sí. $t_{10} = 36.294, P \approx 0$.



- e) $(0.10938, 0.12369)$ f) $(86.083, 86.883)$ g) $(85.041, 87.925)$

13. ii)

15. a) 145.63 b) Sí. $r = -\sqrt{R-Sq} = -0.988$. Observe que r es negativo porque la pendiente de la recta de mínimos cuadrados es negativa. c) 145.68.

17. a) Se necesita minimizar la suma de cuadrados $S = \sum (y_i - \hat{\beta}x_i)^2$. Se toma la derivada con respecto a $\hat{\beta}$ y se hace igual a 0, obteniendo $-2 \sum x_i(y_i - \hat{\beta}x_i) = 0$. Entonces $\sum x_i y_i - \hat{\beta} \sum x_i^2 = 0$, por lo que $\hat{\beta} = \sum x_i y_i / \sum x_i^2$.

b) Sea $c_i = x_i / \sum x_i^2$. Entonces $\hat{\beta} = \sum c_i y_i$, por lo que $\sigma_{\hat{\beta}}^2 = \sum c_i^2 \sigma^2 = \sigma^2 \sum x_i^2 / (\sum x_i^2)^2 = \sigma^2 / \sum x_i^2$.

19. De la respuesta del ejercicio 18, se sabe que $\sum_{i=1}^n (x_i - \bar{x}) = 0$, $\sum_{i=1}^n \bar{x}(x_i - \bar{x}) = 0$, y $\sum_{i=1}^n x_i(x_i - \bar{x}) = \sum_{i=1}^n (x_i - \bar{x})^2$. Ahora

$$\begin{aligned} \mu_{\hat{\beta}_0} &= \sum_{i=1}^n \left[\frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \mu_{y_i} \\ &= \sum_{i=1}^n \left[\frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] (\beta_0 + \beta_1 x_i) \end{aligned}$$

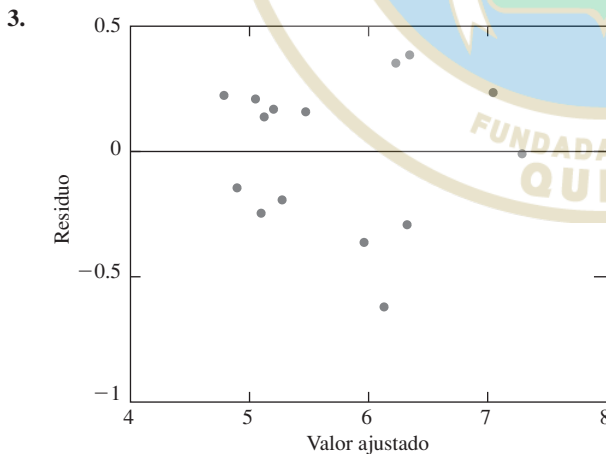
$$\begin{aligned}
&= \beta_0 \sum_{i=1}^n \frac{1}{n} + \beta_1 \sum_{i=1}^n \frac{x_i}{n} - \beta_0 \frac{\sum_{i=1}^n \bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} - \beta_1 \frac{\sum_{i=1}^n x_i \bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \beta_0 + \beta_1 \bar{x} - 0 - \beta_1 \bar{x} \frac{\sum_{i=1}^n x_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \beta_0 + \beta_1 \bar{x} - 0 - \beta_1 \bar{x} \\
&= \beta_0
\end{aligned}$$

21.

$$\begin{aligned}
\sigma_{\hat{\beta}_0}^2 &= \sum_{i=1}^n \left[\frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]^2 \sigma^2 \\
&= \sum_{i=1}^n \left[\frac{1}{n^2} - \frac{2\bar{x}}{n} \frac{(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} + \bar{x}^2 \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} \right] \sigma^2 \\
&= \left[\sum_{i=1}^n \frac{1}{n^2} - 2\frac{\bar{x}}{n} \frac{\sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} + \bar{x}^2 \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} \right] \sigma^2 \\
&= \left[\frac{1}{n} - 2\frac{\bar{x}}{n} (0) + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \sigma^2 \\
&= \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \sigma^2
\end{aligned}$$

Sección 8.1

1. a) 6.9619 b) 0.03925 c) 0.5413



No hay un patrón obvio en la gráfica de residuos, por lo que el modelo lineal parece ajustar bien.

5. a) 25.465 b) No, los cambios pronosticados dependen de los valores de las otras variables independientes, debido a los términos de interacción. c) 0.9691 d) $F_{9,17} = 59.204$. Sí, la hipótesis nula se puede rechazar.

7. a) 2.3411 L b) 0.06768 L c) Nada está equivocado, las estimaciones constantes FEV₁ respecto de una persona cuyos valores para las otras variables sean todas iguales a cero. Debido a que estos valores están fuera del rango de los datos (por ejemplo, no hay ninguna con estatura cero), la constante necesita no representar un valor real para una persona real. d) FEV₁ ajustado = FEV₁ medido + 0.0013431 · presión.
9. a) 3.572 b) 0.098184 c) Nada está equivocado. Las estimaciones constantes del pH respecto de una pulpa cuyos valores para las otras variables sean iguales a cero. Puesto que estos valores están fuera del rango de los datos (por ejemplo, no pulpa tiene densidad cero), la constante necesita no representar un valor real para una pulpa real. d) (3.4207, 4.0496) e) (2.2333, 3.9416) f) Pulpa B. La desviación estándar de su pronosticado (SE Fit) es más pequeño que para la pulpa A (0.1351 contra 0.2510).
11. a) -2.05 b) 0.3512 c) -0.2445 d) 4.72 e) 13.92 f) 18.316 g) 4.54 h) 9
13. a) 135.92°F b) No. El cambio en el punto de inflamación pronosticado debido al cambio en la concentración del ácido acético depende también de la concentración del ácido butírico, debido a la interacción entre estas dos variables. c) Sí. El punto de inflamación pronosticado cambiará por -13.897°F.
15. a) -9.17819, -1.42732, 123.26818, -151.35236, 50.12800, -11.43494 b) SSE = 40 832.432, SST = 1 990 660
c) $s^2 = 13\,610.811$ d) $R^2 = 0.979487$ e) $F = 71.6257$. Hay dos y tres grados de libertad. f) Sí. A partir de la tabla F , $0.001 < P < 0.01$.
17. a) 2.0711 b) 0.17918 c) PP es más útil, ya que su P -valor es pequeño, mientras que el P -valor de PC es bastante grande. d) El cambio porcentual en el PIB se esperaría que fuese mayor en Suecia, porque el coeficiente de PP es negativo.
19. a) $y = -0.012167 + 0.043258t + 2.9205t^2$ b) (2.830, 3.011) c) (5.660, 6.022) d) $\hat{\beta}_0: t_7 = -1.1766$, $P = 0.278$, $\hat{\beta}_1: t_7 = 1.0017$, $P = 0.350$, $\hat{\beta}_2: t_7 = 76.33$, $P = 0.000$. e) No el P -valor de 0.278 no es lo suficientemente pequeño para rechazar la hipótesis nula de $\beta_0 = 0$. f) No, el P -valor de 0.350 no es lo suficientemente pequeño para rechazar la hipótesis nula de $\beta_1 = 0$.

Sección 8.2

1. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	6.3347	2.1740	2.9138	0.009
x_1	1.2915	0.1392	9.2776	0.000

β_0 difiere de 0 ($P = 0.009$), β_1 difiere de 0 ($P = 0.000$).

b)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	53.964	8.7737	6.1506	0.000
x_2	-0.9192	0.2821	-3.2580	0.004

β_0 difiere de 0 ($P = 0.000$), β_1 difiere de 0 ($P = 0.004$).

c)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	12.844	7.5139	1.7094	0.104
x_1	1.2029	0.1707	7.0479	0.000
x_2	-0.1682	0.1858	-0.90537	0.377

β_0 puede no diferir de 0 ($P = 0.104$), β_1 difiere de 0 ($P = 0.000$) β_2 puede no diferir de 0 ($P = 0.377$).

d) El modelo en el inciso a) es el mejor. Cuando tanto x_1 y x_2 están en el modelo, sólo el coeficiente de x_1 es significativamente diferente de 0. Además, el valor de R^2 es sólo ligeramente mayor (0.819 contra 0.811) para el modelo que contiene tanto a x_1 como a x_2 que para el modelo que sólo contiene a x_1 .

3. a) La gráfica i) proviene del ingeniero B, y la gráfica ii), del ingeniero A. Se sabe esto porque las variables x_1 y x_2 son significativamente diferentes de 0 para el ingeniero A, pero no para el ingeniero B; por tanto, el ingeniero B es quien diseñó que el experimento tenga las variables dependientes casi colineales.

b) El experimento del ingeniero A produce el experimento con los resultados más reales. En el experimento del ingeniero B, las dos variables son casi colineales.

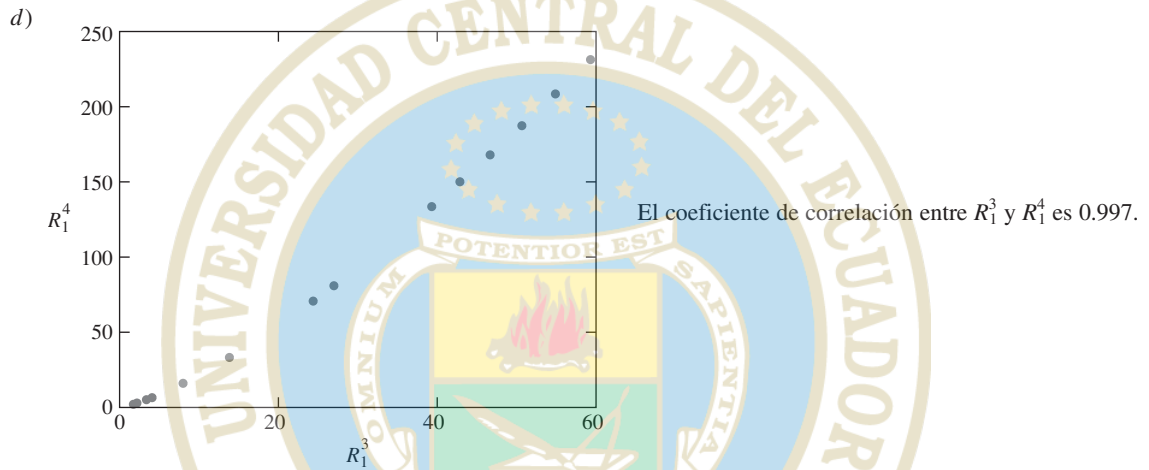
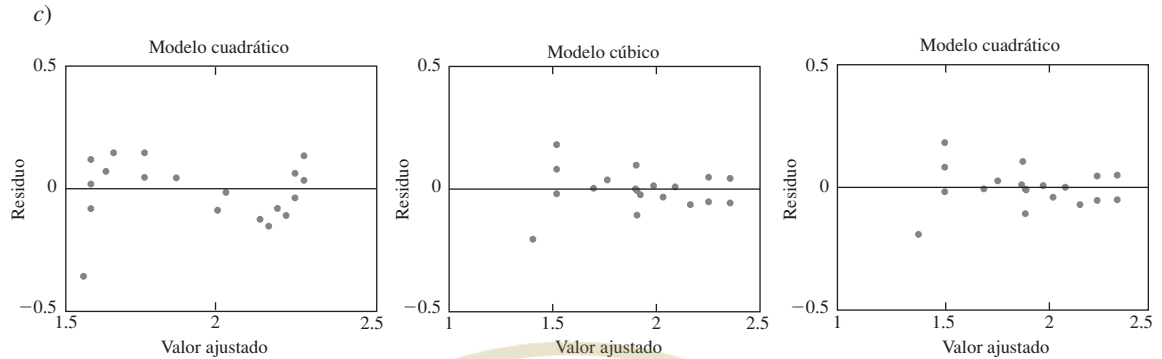
5. a) Para $R_1 < 4$, la recta de mínimos cuadrados es $R_2 = 1.23 + 0.264R_1$. Para $R_1 \geq 4$, la recta de mínimos cuadrados es $R_2 = -0.190 + 0.710R_1$.

b) La relación es obviamente no lineal cuando $R_1 < 4$.

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	1.2840	0.26454	4.8536	0.000
R_1	0.21661	0.23558	0.91947	0.368
R_1^2	0.0090189	0.044984	-0.20049	0.843

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	-1.8396	0.56292	-3.2680	0.004
R_1	4.4987	0.75218	5.9809	0.000
R_1^2	-1.7709	0.30789	-5.7518	0.000
R_1^3	0.22904	0.039454	5.8053	0.000

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	-2.6714	2.0117	-1.3279	0.200
R_1	6.0208	3.6106	1.6675	0.112
R_1^2	-2.7520	2.2957	-1.1988	0.245
R_1^3	0.49423	0.61599	0.80234	0.432
R_1^4	-0.02558	0.05930	-0.43143	0.671



e) R_1^3 y R_1^4 son casi colineales.

f) El modelo cúbico es mejor. El cuadrático no es adecuado porque la gráfica de residuos muestra un patrón. Las gráficas de residuos para los modelos cúbicos y cuárticos ajustan bien; sin embargo, no hay razón para incluir R_1^4 en el modelo porque éste simplemente confunde el efecto de R_1^3 .

Sección 8.3

1. a) Falso b) Verdadero c) Falso d) Verdadero

3. v)

5. i)

7. La cuarta variable del modelo con el valor más alto de R^2 tiene un menor R^2 que la tercera variable del modelo con el valor más alto de R^2 . Esto es imposible.

9. a) 0.2803 b) Tres grados de libertad en el numerador y 157 en el denominador. c) $P > 0.10$. El modelo reducido es creíble. d) Esto no es correcto. Es posible para un grupo de variables que están fuertemente relacionadas con una variable independiente, aun cuando ninguna de las variables individualmente esté fuertemente relacionada.

e) No hay error. Si y es la variable dependiente, entonces la suma total de cuadrados es $\sum (y_i - \bar{y})^2$. La cantidad no implica las variables independientes.

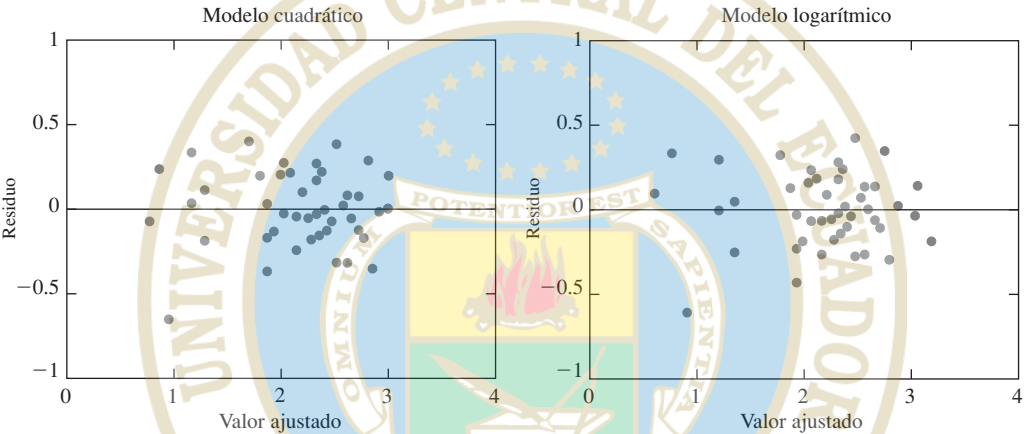
11. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.087324	0.16403	0.53237	0.597
x	0.52203	0.055003	9.4909	0.000
x^2	-0.02333	0.0043235	-5.396	0.000

b)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.19878	0.11677	1.7023	0.096
$\ln x$	1.2066	0.068272	17.673	0.000

c)



Ninguna de las gráficas de residuos revela un fuerte incumplimiento de los supuestos.

d)

x	Modelo cuadrático de predicción	Modelo log de predicción
3.0	1.44	1.52
5.0	2.11	2.14
7.0	2.60	2.55
9.0	2.90	2.85

- e) La prueba F no se puede utilizar; sólo puede emplearse cuando un modelo se forma eliminando una o más de las variables independientes del otro modelo.
- f) Las predicciones de los dos modelos no difieren mucho. El modelo logarítmico tiene sólo una variable independiente en lugar de dos, lo que es una ventaja.

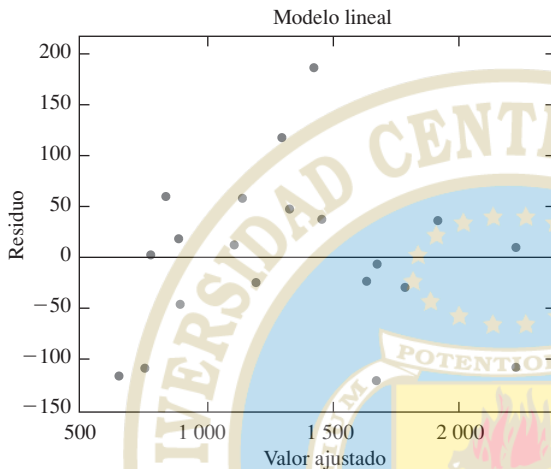
13. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	37.989	53.502	0.71004	0.487
x	1.0774	0.041608	25.894	0.000

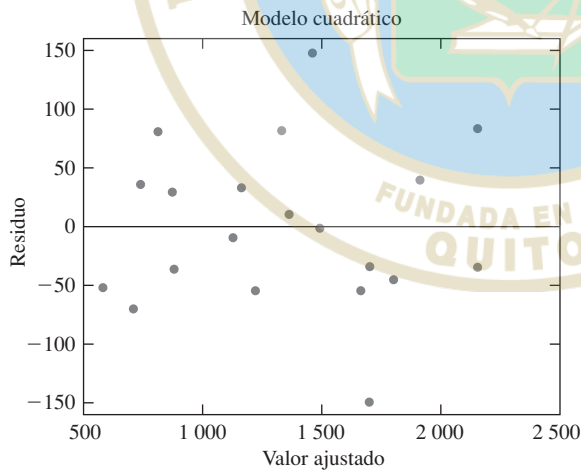
b)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	-253.45	132.93	-1.9067	0.074
x	1.592	0.22215	7.1665	0.000
x^2	-0.00020052	0.000085328	-2.3499	0.031

c)



d)



e) El modelo cuadrático parece el más apropiado. El P -valor para el término cuadrático es bastante pequeño (0.031) y la gráfica de residuos para el modelo cuadrático muestra menos de un patrón. (Sin embargo, hay un par de puntos un poco alejados del resto de la gráfica.)

f) 1 683.5 g) (1 634.7, 1 732.2)

15. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	25.613	10.424	2.4572	0.044
x_1	0.18387	0.12353	1.4885	0.180
x_2	-0.015878	0.0040542	-3.9164	0.006

b)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	14.444	16.754	0.86215	0.414
x_1	0.17334	0.20637	0.83993	0.425

c)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	40.370	3.4545	11.686	0.000
x_2	-0.015747	0.0043503	-3.6197	0.007

d) El modelo que contiene a x_2 como la única variable independiente es mejor. No hay evidencia de que el coeficiente de x_1 difiera de 0.

17. El modelo $y = \beta_0 + \beta_1 x_2 + \varepsilon$ es uno bueno. Una forma de ver esto último es comparar el ajuste de este modelo con el modelo cuadrático completo. La tabla ANOVA para este modelo es

Fuente	GL	SC	MC	F	P
Regresión	5	4.1007	0.82013	1.881	0.193
Error de residuos	9	3.9241	0.43601		
Total	14	8.0248			

La tabla ANOVA para el modelo $y = \beta_0 + \beta_1 x_2 + \varepsilon$ es

Fuente	GL	SC	MC	F	P
Regresión	1	2.7636	2.7636	6.8285	0.021
Error de residuos	13	5.2612	0.40471		
Total	14	8.0248			

A partir de estas dos tablas, el estadístico de prueba para probar la credibilidad del modelo reducido es $\frac{(5.2612 - 3.9241)/(5 - 1)}{3.9241/9} = 0.7667$. La hipótesis nula es $F_{4,9}$, $P > 0.10$. El P -valor grande indica que el modelo reducido es creíble.

Ejercicios adicionales para el capítulo 8

1. a) 24.6% b) 5.43% c) No, se necesita conocer el contenido de oxígeno.
 3. a) 0.207 b) 0.8015 c) 3.82 d) 1.200 e) 2 f) 86.81 g) 43.405 h) 30.14 i) 14

5. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	10.84	0.2749	39.432	0.000
Velocidad	-0.073851	0.023379	-3.1589	0.004
Pausa	-0.12743	0.013934	-9.1456	0.000
Velocidad ²	0.0011098	0.00048887	2.2702	0.032
Pausa ²	0.0016736	0.00024304	6.8861	0.000
Velocidad · Pausa	-0.00024272	0.00027719	-0.87563	0.390

Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	5	31.304	6.2608	56.783	0.000
Error de residuos	24	2.6462	0.11026		
Total	29	33.95			

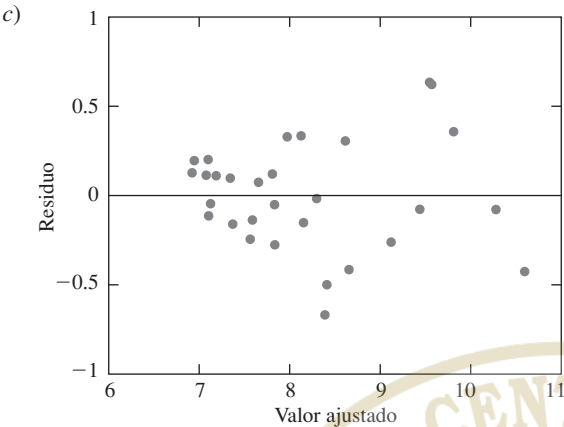
b) Se elimina el término de interacción Velocidad · Pausa.

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	10.967	0.23213	47.246	0.000
Velocidad	-0.079919	0.022223	-3.5961	0.001
Pausa	-0.13253	0.01260	-10.518	0.000
Velocidad ²	0.0011098	0.00048658	2.2809	0.031
Pausa ²	0.0016736	0.0002419	6.9185	0.000

Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	5	31.22	7.8049	71.454	0.000
Error de residuos	25	2.7307	0.10923		
Total	29	33.95			

 Comparando este modelo con el del inciso a), $F_{1,24} = 0.77$, $P > 0.10$.



Hay cierta sugerencia de heteroscedasticidad, pero es difícil de estar seguros sin más datos.

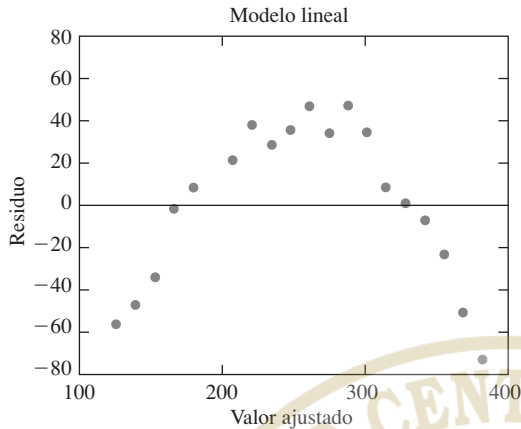
d) No, comparado con el modelo cuadrático completo que contiene Velocidad, Pausa, Velocidad² y Pausa², y Velocidad · Pausa, el estadístico F es $F_{3, 24} = 15.70$, y $P > 0.001$.

e) _____

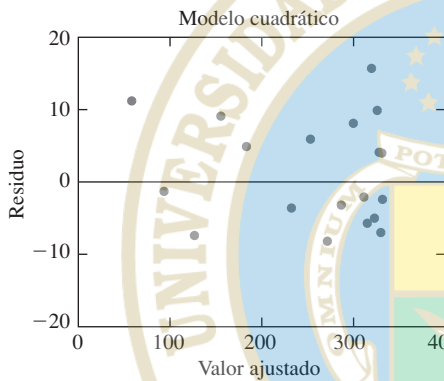
Vars	R-Sq	R-Sq(adj)	C-p	S	P	S	P	S	P
1	61.5	60.1	92.5	0.68318	X				
1	60.0	58.6	97.0	0.69600					X
2	76.9	75.2	47.1	0.53888	X		X		
2	74.9	73.0	53.3	0.56198	X	X			
3	90.3	89.2	7.9	0.35621	X	X	X		
3	87.8	86.4	15.5	0.39903	X		X	X	
4	92.0	90.7	4.8	0.33050	X	X	X	X	
4	90.5	89.0	9.2	0.35858	X	X		X	X
5	92.2	90.6	6.0	0.33205	X	X	X	X	X

f) El modelo que contiene las variables dependientes Velocidad · Pausa, Velocidad² y Pausa², tienen tanto el valor mínimo de C_p como el valor máximo de R^2 ajustado.

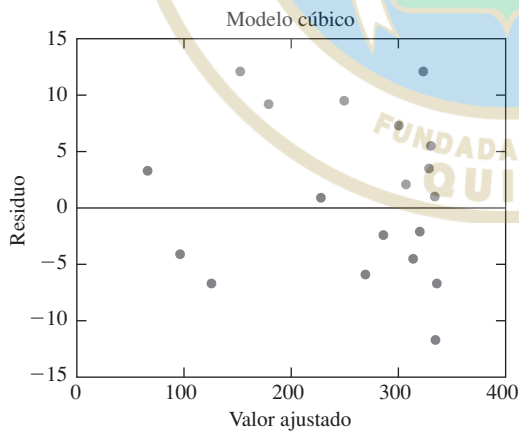
7.



La gráfica de residuos muestra un patrón curvo obvio, por lo que el modelo lineal no es adecuado.



No hay un patrón obvio en la gráfica de residuos, por lo que el modelo cuadrático parece ajustar bien.



No hay un patrón obvio en la gráfica de residuos, por lo que el modelo cúbico parece ajustar bien.

9. a) 182.52, 166.55, 187.56 b) 236.39, 234.18, 237.02 c) 176.80, 163.89, 180.78

d) iv). El resultado no proporciona mucho de dónde elegir entre los modelos de dos variables. En el modelo de tres variables, ninguno de los coeficientes es significativamente diferente de 0, aun cuando fueron significativos en los modelos de dos variables. Esto sugiere colinealidad.

11. a) A continuación se presentan los valores de SSE y sus grados de libertad para modelos de grados 1, 2, 3 y 4.

Lineal	18	2726.55
Cuadrático	17	481.90
Cúbico	16	115.23
Cuártico	15	111.78

Para comparar cuadrático contra lineal, $F_{1,17} = \frac{(2726.55 - 481.90)/(18 - 17)}{481.90/17} = 79.185, P \approx 0.$

Para comparar cúbico contra cuadrático, $F_{1,16} = \frac{(481.90 - 115.23)/(17 - 16)}{115.23/16} = 50.913, P \approx 0.$

Para comparar cuártico contra cúbico, $F_{1,15} = \frac{(115.23 - 111.78)/(16 - 15)}{111.78/15} = 0.463, P > 0.10.$

El modelo cúbico se selecciona con este procedimiento.

- b) El modelo cúbico es $y = 27.937 + 0.48749x + 0.85104x^2 - 0.057254x^3$. La estimación y se maximiza cuando $dy/dx = 0$. $dy/dx = 0.48749 + 1.70208x - 0.171762x^2$. Por tanto, $x = 10.188$.
13. a) Sea y_1 el tiempo de vida de la pintura del patrocinador, y_2 es el tiempo de vida de la pintura del competidor, x_1 es la temperatura en enero, x_2 es la temperatura en julio, x_3 es la precipitación. Entonces un buen modelo para y_1 es $y_1 = -4.2342 + 0.79037x_1 + 0.20554x_2 - 0.082363x_3 - 0.0079983x_1x_2 - 0.0018349x_1^2$. Un buen modelo para y_2 es $y_2 = 6.8544 + 0.58898x_1 + 0.054759x_2 - 0.15058x_3 - 0.0046519x_1x_2 + 0.0019029x_1x_3 - 0.0035069x_1^2$.
- b) $\hat{y}_1 = 13.83, \hat{y}_2 = 13.90$.

15. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.25317	0.0065217	38.819	0.000
x	-0.041561	0.040281	-1.0318	0.320

- b)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.21995	0.0038434	57.23	0.000
x	0.58931	0.06146	9.5886	0.000
x^2	-2.2679	0.2155	-10.524	0.000

- c)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.22514	0.0068959	32.648	0.000
x	0.41105	0.20576	1.9977	0.069
x^2	-0.74651	1.6887	-0.44206	0.666
x^3	-3.6728	4.043	-0.90843	0.382

d)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	0.23152	0.013498	17.152	0.000
x	0.10911	0.58342	0.18702	0.855
x^2	3.4544	7.7602	0.44515	0.665
x^3	-26.022	40.45	-0.64333	0.533
x^4	40.157	72.293	0.55548	0.590

e) El modelo cuadrático. El coeficiente de x^3 en el modelo cúbico no es significativamente diferente de 0. No hay ningún coeficiente de x^4 en el modelo cuártico.

f) 0.258

17. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	-0.093765	0.092621	-1.0123	0.335
x_1	0.63318	2.2088	0.28666	0.780
x_2	2.5095	0.30151	8.3233	0.000
x_1^2	5.318	8.2231	0.64672	0.532
x_2^2	-0.3214	0.17396	-1.8475	0.094
x_1x_2	0.15209	1.5778	0.09639	0.925

Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	5	20.349	4.0698	894.19	0.000
Error de residuos	10	0.045513	0.0045513		
Total	15	20.394			

b) El modelo que contiene las variables x_1 , x_2 y x_2^2 es bueno. A continuación se presentan los coeficientes junto con sus desviaciones estándar, y enseguida la tabla del análisis de varianza.

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	-0.088618	0.068181	-1.2997	0.218
x_1	2.1282	0.30057	7.0805	0.000
x_2	2.4079	0.13985	17.218	0.000
x_2^2	-0.27994	0.059211	-4.7279	0.000

Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	3	20.346	6.782	1683.9	0.000
Error de residuos	12	0.048329	0.0040275		
Total	15	20.394			

El estadístico F para comparar este modelo con el modelo cuadrático completo es

$$F_{2,10} = \frac{(0.048329 - 0.045513)/(12 - 10)}{0.045513/10} = 0.309, P > 0.10, \text{ por lo que es razonable eliminar } x_1^2 \text{ y } x_1x_2 \text{ del modelo}$$

cuadrático completo. Todos los coeficientes restantes son significativamente diferentes de 0, por lo que no sería razonable reducir más el modelo.

- c) El modelo con el mejor R^2 ajustado contiene las variables x_2, x_1^2 y x_2^2 . Este modelo también es el modelo con el valor mínimo de Mallows C_p (2.2). Éste no es el mejor modelo, ya que contiene a x_1^2 , pero no x_1 . El modelo que contiene x_1, x_2 y x_2^2 , sugerido en la respuesta del inciso b), es mejor. Observe que el R^2 ajustado para el modelo en el inciso b) es 0.99704, que casi no difiere del modelo con el máximo valor de R^2 ajustado.

19. a)

Pronosticador	Coefficiente	Desviación estándar	T	P
Constante	1.1623	0.17042	6.8201	0.006
t	0.059718	0.0088901	6.7174	0.007
t^2	-0.00027482	0.000069662	-3.9450	0.029

- b) 17.68 minutos c) (0.0314, 0.0880) d) La tasa de reacción disminuye con el tiempo si $\beta_2 < 0$. Por tanto, se prueba $H_0 : \beta_2 \geq 0$ contra $H_1 : \beta_2 < 0$. El estadístico de prueba es $t_3 = 3.945$, $P = 0.029/2 = 0.0145$. Es razonable concluir que la tasa de reacción disminuye con el tiempo.

21. $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon$.

23. a) El modelo de 17 variables que contiene las variables independientes $x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{11}, x_{13}, x_{14}, x_{16}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}$ y x_{23} tiene un R^2 ajustado igual a 0.98446. El modelo ajustado es

$$\begin{aligned} y = & -1569.8 - 24.909x_1 + 196.95x_2 + 8.8669x_3 - 2.2359x_6 - 0.077581x_7 + 0.057329x_8 \\ & - 1.3057x_9 - 12.227x_{11} + 44.143x_{13} + 4.1883x_{14} + 0.97071x_{16} + 74.775x_{18} \\ & + 21.656x_{19} - 18.253x_{20} + 82.591x_{21} - 37.553x_{22} + 329.8x_{23} \end{aligned}$$

- b) El modelo de ocho variables que contiene las variables independientes $x_1, x_2, x_5, x_8, x_{10}, x_{11}, x_{14}$, y x_{21} tiene un C_p de Mallows igual a 1.7. El modelo ajustado es

$$y = -665.98 - 24.782x_1 + 76.499x_2 + 121.96x_5 + 0.024247x_8 + 20.4x_{10} - 7.1313x_{11} + 2.4466x_{14} + 47.85x_{21}$$

- c) Utilizando un valor de 0.15 tanto para α que se introduce como para α que se elimina, la ecuación elegida para la regresión paso a paso es $y = -927.72 + 142.40x_5 + 0.081701x_7 + 21.698x_{10} + 0.41270x_{16} + 45.672x_{21}$.

- d) El siguiente modelo de 13 variables tiene un R^2 ajustado igual a 0.95402. (También hay modelos de 12 variables cuyo R^2 ajustado es sólo ligeramente menor.)

$$\begin{aligned} z = & 8663.2 - 313.31x_3 - 14.46x_6 + 0.358x_7 - 0.078746x_8 + 13.998x_9 + 230.24x_{10} \\ & - 188.16x_{13} + 5.4133x_{14} + 1928.2x_{15} - 8.2533x_{16} + 294.94x_{19} + 129.79x_{22} - 3020.7x_{23} \end{aligned}$$

- e) El modelo de dos variables $z = -1660.9 + 0.67152x_7 + 134.28x_{10}$ tiene un C_p de Mallows igual a -4.0.

- f) Utilizando un valor de 0.15 tanto para α que se introduce como para α que se elimina, la ecuación elegida por la regresión paso a paso es $z = -1660.9 + 0.67152x_7 + 134.28x_{10}$

- g) El siguiente modelo de 17 variables tiene un R^2 ajustado igual a 0.97783.

$$\begin{aligned} w = & 700.56 - 21.701x_2 - 20.000x_3 + 21.813x_4 + 62.599x_5 + 0.016156x_7 \\ & - 0.012689x_8 + 1.1315x_9 + 15.245x_{10} + 1.1103x_{11} - 20.523x_{13} - 90.189x_{15} \\ & - 0.77442x_{16} + 7.5559x_{19} + 5.9163x_{20} - 7.5497x_{21} + 12.994x_{22} - 271.32x_{23} \end{aligned}$$

- h) El siguiente modelo de 13 variables tiene un C_p de Mallows igual a 8.0.

$$w = 567.06 - 23.582x_2 - 16.766x_3 + 90.482x_5 + 0.0082274x_7 - 0.011004x_8 + 0.89554x_9 \\ + 12.131x_{10} - 11.984x_{13} - 0.67302x_{16} + 11.097x_{19} + 4.6448x_{20} + 11.108x_{22} - 217.82x_{23}$$

- i) Utilizando un valor de 0.15 tanto para α que se introduce como para α que se elimina, la ecuación elegida por la regresión paso a paso es $w = 130.92 - 28.085x_2 + 113.49x_5 + 0.16802x_9 - 0.20216x_{16} + 11.417x_{19} + 12.068x_{21} - 78.371x_{23}$.

Sección 9.1

1. a)

Fuente	GL	SC	MC	F	P
Duración	4	1.3280	0.33200	7.1143	0.002
Error	15	0.7000	0.046667		
Total	19	2.0280			

- b) Sí. $F_{4,15} = 7.1143$, $0.001 < P < 0.01$ ($P = 0.002$).

3. a)

Fuente	GL	SC	MC	F	P
Tratamiento	4	19.009	4.7522	2.3604	0.117
Error	11	22.147	2.0133		
Total	15	41.155			

- b) No. $F_{4,11} = 2.3604$, $P > 0.10$ ($P = 0.117$).

5. a)

Fuente	GL	SC	MC	F	P
Lugar	3	1.4498	0.48327	2.1183	0.111
Error	47	10.723	0.22815		
Total	50	12.173			

- b) No. $F_{3,47} = 2.1183$, $P > 0.10$ ($P = 0.111$).

7. a)

Fuente	GL	SC	MC	F	P
Catalizador	2	61.960	30.980	5.9926	0.037
Error	6	31.019	5.1698		
Total	8	92.979			

- b) Sí. $F_{2,6} = 5.9926$, $0.01 < P < 0.05$ ($P = 0.037$).

9. No, $F_{3,16} = 15.8255$, $P < 0.001$ ($P \approx 4.8 \times 10^{-5}$).

11. a)

Fuente	GL	SC	MC	F	P
Temperatura	3	58.650	19.550	8.4914	0.001
Error	16	36.837	2.3023		
Total	19	95.487			

b) Sí. $F_{3,16} = 8.4914$, $0.001 < P < 0.01$ ($P = 0.0013$).

13. a) $s = 1.517$ b) 18 c) 38

15. a)

Fuente	GL	SC	MC	F	P
Grado	3	1721.4	573.81	9.4431	0.000
Error	96	5833.4	60.765		
Total	99	7554.9			

b) Sí. $F_{3,96} = 9.4431$, $P < 0.001$ ($P \approx 0$).

17. a) 0.85 b) 56 c) 200 d) Mayor que 0.864138. Entre mayor sea la diferencia, mayor será la probabilidad de detectarla.

Sección 9.2

- a) Sí. $F_{5,6} = 46.64$, $P \approx 0$.

b) $q_{6,6,.05} = 5.63$. El valor de MSE es 0.00508. El valor crítico al 5% es, por tanto, $5.63\sqrt{0.00508/2} = 0.284$. Cualquier par que difiera en más de 0.284 se puede concluir que es diferente. Los siguientes pares satisfacen este criterio: A y B, A y C, A y D, A y E, B y C, B y D, B y E, B y F, D y F.

c) $t_{6,.025/15} = 4.698$ (el valor obtenido por interpolación es 4.958). El valor de MSE es 0.00508. El valor crítico al 5% es, por tanto, $4.698\sqrt{2(0.00508)/2} = 0.335$. Cualquier par que difiera en más de 0.335 se puede concluir que es diferente. Los siguientes pares satisfacen este criterio: A y B, A y C, A y D, A y E, B y C, B y D, B y E, B y F, D y F.

d) El método de Tukey-Kramer es más poderoso, ya que su valor crítico es más pequeño (0.284 contra 0.335).

e) Sea el método de Tukey-Kramer o el método de Bonferroni, ambos se pueden utilizar.
- a) MSE = 2.9659, $J_i = 12$ para toda i . Hay siete comparaciones que deben hacerse. Ahora $t_{88,.025/7} = 2.754$, por lo que el valor crítico al 5% es $2.754\sqrt{2.9659(1/12 + 1/12)} = 1.936$. Todas las medias muestrales de las formulaciones no controladas difieren de la media muestral de las formulaciones controladas en más de esta cantidad. Por tanto, se concluye que al nivel de 5% todas las formulaciones no controladas difieren de la formulación controlada.

b) Hay siete comparaciones que deben hacerse. Se debe utilizar el valor de rango de Student $q_{7,88,.05}$. Este valor no está en la tabla, por lo que se usará $q_{7,60,.05} = 4.31$, que es un poco mayor. El valor crítico al 5% es $4.31\sqrt{2.9659/12} = 2.14$. Todas las formulaciones no controladas difieren de la media muestral de la formulación controlada en más de esta cantidad. Por tanto, se concluye que al nivel de 5% todas las formulaciones no controladas difieren de la formulación controlada.

c) El método de Bonferroni es más poderoso, porque está basado en el número real de comparaciones que deben hacerse, que es siete. El método de Tukey-Kramer está basado en el número mayor de comparaciones que se pudieran hacer, que es $(7)(8)/2 = 28$.
- a) $t_{16,.025/6} = 3.0083$ (el valor obtenido por interpolación es 3.080). El valor de MSE es 2.3023. Por tanto, el valor crítico al 5% es $3.0083\sqrt{2(2.3023)/5} = 2.8869$. Se puede concluir que la media para 750°C difiere de las medias para 850°C y 900°C, y que la media para 800°C difiere de la media para 900°C.

b) $q_{4,16,.05} = 4.05$. El valor de MSE es 2.3023. Por tanto, el valor crítico al 5% es $4.05\sqrt{2.3023/5} = 2.75$. Se puede concluir que la media para 750°C difiere de las medias para 850°C y 900°C, y que la media para 800°C difiere de la media para 900°C.

c) El método de Tukey-Kramer es más poderoso, porque su valor crítico es menor.
- a) $t_{16,.025/3} = 2.6730$ (el valor obtenido por interpolación es 2.696). El valor de MSE es 2.3023. Por tanto, el valor crítico al 5% es $2.6730\sqrt{2(2.3023)/5} = 2.5651$. Se puede concluir que la media para 900°C difiere de las medias para 750°C y 800°C.

- b) $q_{4, 16, .05} = 4.05$. El valor de MSE es 2.3023. Por tanto, el valor crítico al 5% es $4.05\sqrt{2.3023/5} = 2.75$. Se puede concluir que la media para 900°C difiere de las medias para 750°C y 800°C.
- c) El método de Bonferroni es más poderoso, porque su valor crítico es menor.
9. a) $t_{47, .025} = 2.012$, MSE = 0.22815, los tamaños de la muestra son 16 y 9. Las medias de la muestra son $\bar{X}_1 = 1.255625$, $\bar{X}_2 = 1.756667$. El intervalo de confianza del 95% es $0.501042 \pm 2.012\sqrt{0.22815(1/16 + 1/9)}$, o (0.1006, 0.9015).
- b) Los tamaños de la muestra son $J_1 = 16$, $J_2 = 9$, $J_3 = 14$, $J_4 = 12$. MSE = 0.22815. Se debe usar el valor del rango de Student $q_{4, 47, .05}$. Este valor no está en la tabla, por lo que se usará $q_{4, 40, .05} = 3.79$, que es sólo un poco mayor. Los valores de $q_{4, 40, .05}\sqrt{(MSE/2)(1/J_i + 1/J_j)}$ se presentan en la siguiente tabla de la izquierda y los valores de las diferencias $|\bar{X}_i - \bar{X}_j|$ se presentan en la tabla de la derecha.

	1	2	3	4		1	2	3	4
1	—	0.53336	0.46846	0.48884	1	0	0.50104	0.16223	0.18354
2	0.53336	—	0.54691	0.56446	2	0.50104	0	0.33881	0.3175
3	0.46846	0.54691	—	0.50358	3	0.16223	0.33881	0	0.02131
4	0.48884	0.56446	0.50358	—	4	0.18354	0.31750	0.02131	0

Ninguna de las diferencias excede su valor crítico, por lo que cualquiera de las medias de tratamiento difiere.

11. a) $t_{8, .025} = 2.306$, MSE = 1.3718. Las medias muestrales son $\bar{X}_1 = 1.998$ y $\bar{X}_3 = 5.300$. Los tamaños de las muestras son $J_1 = 5$ y $J_3 = 3$. Por tanto, el intervalo de confianza del 95% es $3.302 \pm 2.306\sqrt{1.3718(1/5 + 1/3)}$, o (1.330, 5.274).
- b) Las medias muestrales son $\bar{X}_1 = 1.998$, $\bar{X}_2 = 3.0000$, $\bar{X}_3 = 5.300$. Los tamaños de la muestra son $J_1 = 5$, $J_2 = J_3 = 3$. El punto superior al 5% del rango de Student es $q_{3, 8, .05} = 4.04$. El valor crítico para $|\bar{X}_1 - \bar{X}_2|$ y para $|\bar{X}_1 - \bar{X}_3|$ es $4.04\sqrt{(1.3718/2)(1/5 + 1/3)} = 2.44$, y el valor crítico al 5% para $|\bar{X}_2 - \bar{X}_3|$ es $4.04\sqrt{(1.3718/2)(1/3 + 1/3)} = 2.73$. Por tanto, las medias 1 y 3 difieren al nivel de 5%.
13. a) MStr = 19.554 por lo que $F = 19.554/3.85 = 5.08$. Hay 3 y 16 grados de libertad, por lo que $0.01 < P < 0.05$. La hipótesis nula de no diferencia se rechaza al nivel de 5%.
- b) $q_{4, 16, .05} = 4.05$, por lo que el catalizador cuyas medias difieren en más de $4.05\sqrt{3.85/5} = 3.55$ son significativamente diferentes al nivel de 5%. Los dos catalizadores 1 y 2 difieren significativamente del catalizador 4.
15. Cualquier valor de MSE satisface $5.099 < \text{MSE} < 6.035$.

Sección 9.3

1. a) 3 b) 2 c) 6 d) 24

e)

Fuente	GL	SC	MC	F	P
Aceite	3	1.0926	0.36420	5.1314	0.007
Anillo	2	0.9340	0.46700	6.5798	0.005
Interacción	6	0.2485	0.041417	0.58354	0.740
Error	24	1.7034	0.070975		
Total	35	3.9785			

- f) Sí. $F_{6, 24} = 0.58354$, $P > 0.10$ ($P = 0.740$).
- g) No, algunos de los efectos principales del tipo de aceite son diferentes de cero. $F_{3, 24} = 5.1314$, $0.001 < P < 0.01$ ($P = 0.007$).
- h) No, algunos de los efectos principales del tipo de anillo son diferentes de cero. $F_{2, 24} = 6.5798$, $0.001 < P < 0.01$ ($P = 0.005$).

3. a)

Fuente	GL	SC	MC	F	P
Temperatura molde	4	69 738	17 434.5	6.7724	0.000
Aleación	2	8958	4 479.0	1.7399	0.187
Interacción	8	7275	909.38	0.35325	0.939
Error	45	115 845	2 574.3		
Total	59	201 816			

b) Sí. $F_{8,45} = 0.35325$, $P > 0.10$ ($P = 0.939$).c) No, algunos de los principales efectos de la temperatura de molde son distintos de cero. $F_{4,45} = 6.7724$, $P < 0.001$ ($P \approx 0$).d) Sí. $F_{3,45} = 1.7399$, $P > 0.10$, ($P = 0.187$).

5. a)

Efectos principales de la solución		Efectos principales de la temperatura		Interacciones	
				Solución	Temperatura
					25°C 37°C
NaCl	-9.1148	25°C	1.8101	NaCl	-0.49983 0.49983
Na ₂ HPO ₄	9.1148	37°C	-1.8101	Na ₂ HPO ₄	0.49983 -0.49983

b)

Fuente	GL	SC	MC	F	P
Solución	1	1 993.9	1 993.9	5.1983	0.034
Temperatura	1	78.634	78.634	0.20500	0.656
Interacción	1	5.9960	5.9960	0.015632	0.902
Error	20	7 671.4	383.57		
Total	23	9 750.0			

c) Sí. $F_{1,20} = 0.015632$, $P > 0.10$ ($P = 0.902$).d) Sí, puesto que el modelo aditivo es creíble. La media de la tensión producida difiere entre Na₂HPO₄ y NaCl: $F_{1,20} = 5.1983$, $0.01 < P < 0.05$ ($P = 0.034$).e) No hay evidencia de que la temperatura afecte la tensión producida: $F_{1,20} = 0.20500$, $P > 0.10$ ($P = 0.656$).

7. a)

Efectos principales del peso		Efectos principales del ácido láurico		Interacciones	
				Fracción = 10	Fracción = 30
15	1.9125	10	2.8375	Peso = 15	-3.9125 3.9125
25	-1.9125	30	-2.8375	Peso = 25	3.9125 -3.9125

b)

Fuente	GL	SC	MC	F	P
Peso	1	29.261	29.261	7.4432	0.053
Ácido láurico	1	64.411	64.411	16.384	0.016
Interacción	1	122.46	122.46	31.151	0.005
Error	4	15.725	3.9312		
Total	7	231.86			

c) No. $F_{1,4} = 31.151$, $0.001 < P < 0.01$ ($P = 0.005$).

d) No, debido a que el modelo aditivo es rechazado.

e) No, debido a que el modelo aditivo es rechazado.

9. a)

Efectos principales del material		Efectos principales del largo		Interacciones		
				Largo		
				Corto	Medio	Largo
CPTi-ZrO ₂	0.044367	Corto	-0.018533	CPTi-ZrO ₂	0.0063333	-0.023767
Aleación Ti-ZrO ₂	-0.044367	Medio	-0.024833	Aleación Ti-ZrO ₂	0.0063333	0.023767
		Largo	0.043367			-0.017433

b)

Fuente	GL	SC	MC	F	P
Material del cono	1	0.059052	0.059052	23.630	0.000
Longitud	2	0.028408	0.014204	5.6840	0.010
Interacción	2	0.0090089	0.0045444	1.8185	0.184
Error	24	0.059976	0.002499		
Total	29	0.15652			

- c) Sí, puede ser creíble que las interacciones sean iguales a 0. El valor del estadístico de prueba es 1.8185, su distribución nula es $F_{2,24}$ y $P > 0.10$ ($P = 0.184$).
- d) Sí, debido a que el modelo aditivo es creíble. La media del coeficiente de fricción difiere entre CPTi-ZrO₂ y Aleación Ti-ZrO₂: $F_{1,24} = 23.630$, $P < 0.001$.
- e) Sí, debido a que el modelo aditivo es creíble. La media del coeficiente de fricción no es la misma para todas las longitudes: $F_{2,24} = 5.6840$, $P \approx 0.01$. Para determinar qué pares de efectos difieren, se utiliza $q_{3,24,.05} = 3.53$. Se calcula $3.53\sqrt{0.002499/10} = 0.056$. Se concluye que el efecto de la longitud del largo difiere de las longitudes cortas y medianas, pero no se puede concluir que los efectos de las longitudes cortas y medianas difieren entre sí.

11. a)

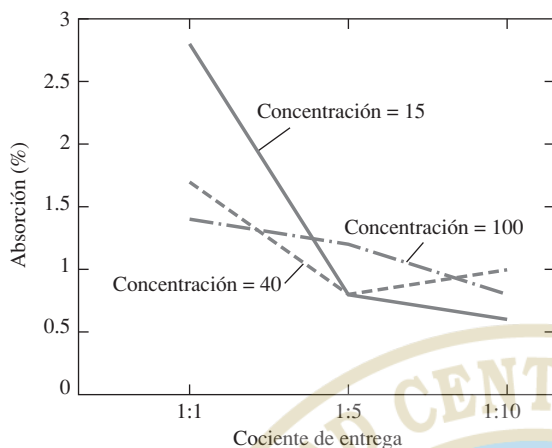
				Interacciones			
Efectos principales de la concentración		Efectos principales del cociente de entrega		Concentración	Cociente de entrega		
					1:1	1:5	1:10
15	0.16667	1:1	0.73333	15	0.66778	-0.30222	-0.36556
40	-0.067778	1:5	-0.30000	40	-0.20111	-0.064444	0.26556
100	-0.098889	1:10	-0.43333	100	-0.46667	0.36667	0.10000

b)

Fuente	DL	SC	MC	F	P
Concentración	2	0.37936	0.18968	3.8736	0.040
Cociente de entrega	2	7.34	3.67	74.949	0.000
Interacción	4	3.4447	0.86118	17.587	0.000
Error	18	0.8814	0.048967		
Total	26	12.045			

- c) No. El valor del estadístico de prueba es 17.587, su distribución nula es $F_{4,18}$, y $P \approx 0$.

d)



Las pendientes de los segmentos de la recta son bastante diferentes entre sí, lo que indica el alto grado de interacción.

13. a)

Efectos principales del sistema de unión			Interacciones		
			Longitud		
Efectos principales de la longitud			Sistema de unión		
			Un cuarto	La mitad	Completo
Clavo	-1.3832	Un cuarto -7.1165	Clavo	0.48317	-0.51633
Adhesivo	1.3832	La mitad -2.5665	Adhesivo	-0.48317	0.51633
		Completo 9.683			

b)

Fuente	GL	SC	MC	F	P
Sistema de unión	1	114.79	114.79	57.773	0.000
Longitud	2	3 019.8	1 509.9	759.94	0.000
Interacción	2	10,023	5.0115	2.5223	0.090
Error	54	107.29	1.9869		
Total	59	3 251.9			

c) El modelo aditivo es apenas creíble: $F_{2,54} = 2.5223$, $0.05 < P < 0.10$ ($P = 0.090$).

d) Sí, el método del sistema de unión afecta la carga crítica de doblaje: $F_{1,54} = 57.773$, $P \approx 0$.

e) Sí, el miembro de lado afecta la carga crítica de doblaje: $F_{2,54} = 759.94$, $P \approx 0$. Para determinar cuál de los efectos difiere a un nivel de 5% se usaría $q_{3,54,.05}$. Este valor no se encuentra en la tabla A.8, por lo que se aproxima a éste con $q_{3,40,.05} = 3.44$. Se calcula $3.44\sqrt{1.9869/20} = 1.08$. Se concluye que los efectos de un cuarto, la mitad y completo difieren entre sí.

15. a)

Fuente	GL	SC	MC	F	P
Placa	2	114 661.4	57 330.7	11 340.1	0.000
Operador	2	136.78	68.389	13.53	0.002
Interacción	4	6.5556	1.6389	0.32	0.855
Error	9	45.500	5.0556		
Total	17	114 850.3			

b) Hay diferencias entre los operadores. $F_{2,9} = 13.53$, $0.01 < P < 0.001$ ($P = 0.002$).

17. a)

Fuente	GL	SC	MC	F	P
PVAL	2	125.41	62.704	8.2424	0.003
DCM	2	1 647.9	823.94	108.31	0.000
Interacción	4	159.96	39.990	5.2567	0.006
Error	18	136.94	7.6075		
Total	26	2 070.2			

- b) Puesto que los términos de interacción no son iguales a 0 ($F_{4,18} = 5.2567$, $P = 0.006$), no se pueden interpretar los efectos principales. Por tanto, se calculan las medias de las celdas. Éstas son

PVAL	DCM (ml)		
	50	40	30
0.5	97.8	92.7	74.2
1.0	93.5	80.8	75.4
2.0	94.2	88.6	78.8

Se concluye que un nivel DCM de 50 ml produce mayor eficiencia de encapsulación que en los otros niveles. Si DCM = 50, la concentración PVAL no tiene mucho efecto. Observe que para DCM = 50, la eficiencia de encapsulación se maximiza en la menor concentración PVAL, pero para DCM = 30 se maximiza a la mayor concentración PVAL. Ésta es la fuente de la interacción significativa.

Sección 9.4

1. a) El abono es el factor bloqueado, la tierra es el factor de tratamiento.

b)

Fuente	GL	SC	MC	F	P
Tierra	3	1.178	0.39267	18.335	0.000
Bloque	4	5.047	1.2617	58.914	0.000
Error	12	0.257	0.021417		
Total	19	6.482			

- c) Sí, $F_{3,12} = 18.335$, $P \approx 0$.

- d) $q_{4,12,.05} = 4.20$, $MSAB = 0.021417$ y $J = 5$. Por tanto, el valor crítico de 5% es $4.20\sqrt{0.021417/5} = 0.275$. Las medias muestrales son $\bar{X}_A = 6.32$, $\bar{X}_B = 6.02$, $\bar{X}_C = 6.28$, $\bar{X}_D = 6.70$. Por tanto, se puede concluir que D difiere de A, B y C, y que A difiere de B.

3. a)

Fuente	GL	SC	MC	F	P
Concentración	2	756.7	378.35	5.0185	0.015
Bloque	3	504.7	168.23	2.2315	0.110
Interacción	6	415.3	69.217	0.91809	0.499
Error	24	1 809.4	75.392		
Total	35	3 486.1			

- b) Sí. El P -valor para la interacción es grande (0.499).

- c) Sí. El P -valor para la concentración es pequeño (0.015).

5. a)

Fuente	GL	SC	MC	F	P
Variedad	9	339 032	37 670	2.5677	0.018
Bloque	5	1 860 838	372 168	25.367	0.000
Error	45	660 198	14 671		
Total	59	2 860 069			

b) Sí, $F_{9,45} = 2.5677$, $P = 0.018$.

7. a)

Fuente	GL	SC	MC	F	P
Centrales de abastecimiento de agua	3	1 253.5	417.84	4.8953	0.005
Bloque	14	1 006.1	71.864	0.84193	0.622
Error	42	3 585.0	85.356		
Total	59	5 844.6			

b) Sí, $F_{3,42} = 4.8953$, $P = 0.005$. c) Para determinar cuáles efectos difieren al 5%, se debe usar $q_{4,42,.05}$. Este valor no se encuentra en la tabla A.8, por lo que se aproxima con $q_{4,40,.05} = 3.79$. El valor crítico del 5% es $3.79\sqrt{85.356/15} = 9.04$. Las medias muestrales son $\bar{X}_A = 34.000$, $\bar{X}_B = 22.933$, $\bar{X}_C = 24.800$, $\bar{X}_D = 31.467$. Se puede concluir que A difiere tanto de B como de C.

d) El P -valor para el factor bloqueado es grande (0.622), lo que sugiere que el factor bloqueado (tiempo) tiene sólo un efecto pequeño en el resultado. Por tanto, podría ser razonable despreciar al factor bloqueado y realizar un ANOVA de un sentido.

9. a) Un motor de cada tipo se deberá probar cada día. El orden con el cual se probarán en un día dado debe ser aleatorio. Éste es un diseño de bloques aleatorios, en el cual los días son los bloques. Éste no es un diseño completamente aleatorio, ya que la aleatorización ocurre sólo dentro de los bloques.

b) El estadístico de prueba es
$$\frac{\sum_{i=1}^5 (\bar{X}_{i.} - \bar{X}_{..})^2}{\sum_{j=1}^4 \sum_{i=1}^5 (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2 / 12}.$$

Sección 9.5

1.

	A	B	C	D
1	-	-	-	-
ad	+	-	-	+
bd	-	+	-	+
ab	+	+	-	-
cd	-	-	+	+
ac	+	-	+	-
bc	-	+	+	-
abcd	+	+	+	+

Los pares de pseudónimos son $\{A, BCD\}$, $\{B, ACD\}$, $\{C, ABD\}$, $\{D, ABC\}$, $\{AB, CD\}$, $\{AC, BD\}$ y $\{AD, BC\}$

3. a)

Variable	Efecto	GL	Suma de cuadrados	Media cuadrática	F	P
A	23.0000	1	2 116.0000	2 116.0000	264.5000	0.000
B	-5.0000	1	100.0000	100.0000	12.5000	0.008
C	1.5000	1	9.0000	9.0000	1.1250	0.320
AB	7.5000	1	225.0000	225.0000	28.1250	0.001
AC	3.0000	1	36.0000	36.0000	4.5000	0.067
BC	0.0000	1	0.0000	0.0000	0.0000	1.000
ABC	0.5000	1	1.0000	1.0000	0.1250	0.733
Error		8	64.0000	8.0000		
Total		15	2 551.0000			

b) Los efectos principales A y B y la interacción AB son más importantes. La interacción AC es frontera.

c) La media de la producción es mayor cuando la temperatura es 180°C.

5. a)

Variable	Efecto
A	3.3750
B	23.625
C	1.1250
AB	-2.8750
AC	-1.3750
BC	-1.6250
ABC	1.8750

b) No, puesto que el diseño es sin réplica, no hay suma de cuadrados del error.

c) No, ninguno de los términos de interacción son casi tan grandes como el efecto principal del factor B.

d) Si el modelo aditivo se sabe que vale, entonces la siguiente tabla ANOVA muestra que los efectos principales de B no es igual a 0, mientras que los efectos principales de A y C podría ser igual a 0.

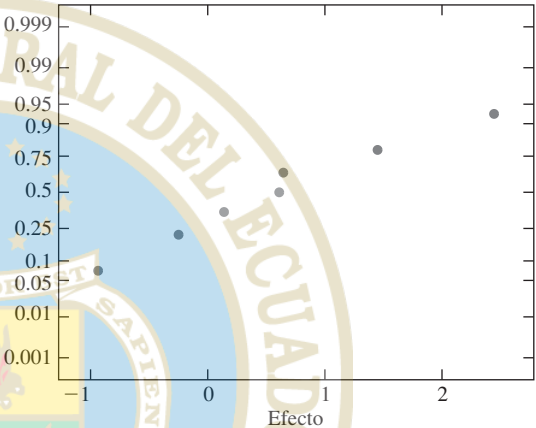
Variable	Efecto	GL	Suma de cuadrados	Media cuadrática	F	P
A	3.3750	1	22.781	22.781	2.7931	0.170
B	23.625	1	1 116.3	1 116.3	136.86	0.000
C	1.1250	1	2.5312	2.5312	0.31034	0.607
Error		4	32.625	8.1562		
Total		7	1 174.2			

7. a)

Variable	Efecto
A	2.445
B	0.140
C	-0.250
AB	1.450
AC	0.610
BC	0.645
ABC	-0.935

b) No, debido a que el diseño es sin réplica, no hay suma de los cuadrados del error.

c) Las estimaciones se encuentran cerca de una recta, por lo que ninguno de los factores se puede decir obviamente que influya en la resistencia.



9. a)

Variable	Efecto
A	1.2
B	3.25
C	-16.05
D	-2.55
AB	2
AC	2.9
AD	-1.2
BC	1.05
BD	-1.45
CD	-1.6
ABC	-0.8
ABD	-1.9
ACD	-0.15
BCD	0.8
ABCD	0.65

b) El factor C es el único que realmente se distingue.

11. a)

Variable	Efecto	GL	Suma de cuadrados	Media cuadrática	F	P
A	14.245	1	811.68	811.68	691.2	0.000
B	8.0275	1	257.76	257.76	219.5	0.000
C	-6.385	1	163.07	163.07	138.87	0.000
AB	-1.68	1	11.29	11.29	9.6139	0.015
AC	-1.1175	1	4.9952	4.9952	4.2538	0.073
BC	-0.535	1	1.1449	1.1449	0.97496	0.352
ABC	-1.2175	1	5.9292	5.9292	5.0492	0.055
Error		8	9.3944	1.1743		
Total		15	1 265.3			

- b) Todos los efectos principales son significativos, así como la interacción *AB*. Únicamente la interacción *BC* tiene un *P*-valor que es razonablemente grande. Los tres factores parecen importantes, y parecen interaccionar considerablemente entre sí.

13. ii)

Ejercicios adicionales para el capítulo 9

1.

Fuente	GL	SC	MC	F	P
Yeso	3	0.013092	0.0043639	0.28916	0.832
Error	8	0.12073	0.015092		
Total	11	0.13383			

El valor del estadístico es $F_{3,8} = 0.28916$; $P > 0.10$ ($P = 0.832$). No hay evidencia de que el pH difiera con la cantidad de yeso agregado.

3.

Fuente	GL	SC	MC	F	P
Día	2	1.0908	0.54538	22.35	0.000
Error	36	0.87846	0.024402		
Total	38	1.9692			

Se puede concluir que la media del contenido del azúcar entre los tres días ($F_{2,36} = 22.35$, $P \approx 0$).

5. a) No. Las varianzas no son constantes a través de los grupos. En particular, hay un dato atípico en el grupo 1.
b) No, por las mismas razones que en el inciso a).

c)

Fuente	GL	SC	MC	F	P
Grupo	4	5.2029	1.3007	8.9126	0.000
Error	35	5.1080	0.14594		
Total	39	10.311			

Se concluye que la media de los tiempos de disolución difiere entre los grupos ($F_{4,35} = 8.9126$, $P \approx 0$).

7. La recomendación no es buena. El ingeniero está tratando de interpretar los efectos principales sin considerar las interacciones. El *P*-valor pequeño para las interacciones indica que se deben considerar. Observando las medias de la celda,

es obvio que si se usa el diseño 2, entonces el material menos costoso se desarrolla igual de bien que el material más costoso. Por tanto, la mejor recomendación es utilizar el diseño 2 con el material menos costoso.

9. a)

Fuente	GL	SC	MC	F	P
Base	3	13 495	4 498.3	7.5307	0.000
Instrumento	2	90 990	45 495	76.164	0.000
Interacción	6	12 050	2 008.3	3.3622	0.003
Error	708	422 912	597.33		
Total	719	539 447			

b) No, esto no es adecuado debido a que hay interacciones entre los efectos renglón y columna ($F_{6, 708} = 3.3622$, $P = 0.003$).

11. a) Sí, $F_{4, 15} = 8.7139$, $P = 0.001$. b) $q_{5, 20} = 4.23$, $MSE = 29.026$, $J = 4$. Por tanto, el valor crítico 5% es $4.23\sqrt{29.026/4} = 11.39$. Las medias muestrales para los cinco canales son $\bar{X}_1 = 44.000$, $\bar{X}_2 = 44.100$, $\bar{X}_3 = 30.900$, $\bar{X}_4 = 28.575$, $\bar{X}_5 = 44.425$. Por tanto, podemos concluir que los canales 3 y 4 difieren de los canales 1, 2 y 5.

13. No, $F_{4, 289} = 1.5974$, $P > 0.10$ ($P = 0.175$).

15. a) $s = 5.388$ b) 10 c) 22

17. a)

Variable	Efecto	Variable	Efecto	Variable	Efecto	Variable	Efecto
A	3.9875	AB	-0.1125	BD	-0.0875	ACD	0.4875
B	2.0375	AC	0.0125	CD	0.6375	BCD	-0.3125
C	1.7125	AD	-0.9375	ABC	-0.2375	ABCD	-0.7125
D	3.7125	BC	0.7125	ABD	0.5125		

b) Los efectos principales son notablemente mayores que las interacciones, y los efectos principales para A y D son notablemente mayores que aquellos para B y C.

c)

Variable	Efecto	GL	Suma de cuadrados	Media cuadrática	F	P
A	3.9875	1	63.601	63.601	68.415	0.000
B	2.0375	1	16.606	16.606	17.863	0.008
C	1.7125	1	11.731	11.731	12.619	0.016
D	3.7125	1	55.131	55.131	59.304	0.001
AB	-0.1125	1	0.050625	0.050625	0.054457	0.825
AC	0.0125	1	0.000625	0.000625	0.00067231	0.980
AD	-0.9375	1	3.5156	3.5156	3.7818	0.109
BC	0.7125	1	2.0306	2.0306	2.1843	0.199
BD	-0.0875	1	0.030625	0.030625	0.032943	0.863
CD	0.6375	1	1.6256	1.6256	1.7487	0.243
Interacción		5	4.6481	0.92963		
Total		15	158.97			

Se puede concluir que cada uno de los factores A, B, C y D tienen un efecto en el resultado.

- d) Los estadísticos F se calculan al dividir la media cuadrática para cada efecto (igual a su suma de los cuadrados) entre el error medio cuadrático 1.04. Los grados de libertad para cada estadístico F son 1 y 4. Los resultados se resumen en la tabla siguiente.

Variable	Efecto	GL	Suma de cuadrados	Media cuadrática	F	P
A	3.9875	1	63.601	63.601	61.154	0.001
B	2.0375	1	16.606	16.606	15.967	0.016
C	1.7125	1	11.731	11.731	11.279	0.028
D	3.7125	1	55.131	55.131	53.01	0.002
AB	-0.1125	1	0.050625	0.050625	0.048678	0.836
AC	0.0125	1	0.000625	0.000625	0.00060096	0.982
AD	-0.9375	1	3.5156	3.5156	3.3804	0.140
BC	0.7125	1	2.0306	2.0306	1.9525	0.235
BD	-0.0875	1	0.030625	0.030625	0.029447	0.872
CD	0.6375	1	1.6256	1.6256	1.5631	0.279
ABC	-0.2375	1	0.22563	0.22563	0.21695	0.666
ABD	0.5125	1	1.0506	1.0506	1.0102	0.372
ACD	0.4875	1	0.95063	0.95063	0.91406	0.393
BCD	-0.3125	1	0.39062	0.39062	0.3756	0.573
ABCD	-0.7125	1	2.0306	2.0306	1.9525	0.235

- e) Sí. Ninguno de los P -valores para las interacciones de tercer o más alto orden es pequeño.
 f) Se puede concluir que cada uno de los factores A, B, C y D tienen un efecto en el resultado.

19. Sí, $F_{2,107} = 9.4427$, $P < 0.001$.

21. a)

Fuente	GL	SC	MC	F	P
H ₂ SO ₄	2	457.65	228.83	8.8447	0.008
CaCl ₂	2	38 783	19 391	749.53	0.000
Interacción	4	279.78	69.946	2.7036	0.099
Error	9	232.85	25.872		
Total	17	39 753			

- b) El P -valor para la interacción es 0.099. No se puede descartar el modelo aditivo.
 c) Sí, $F_{2,9} = 8.8447$, $0.001 < P < 0.01$ ($P = 0.008$).
 d) Sí, $F_{2,9} = 749.53$, $P \approx 0.000$.

23. Sí, $F_{6,39} = 20.302$, $P \approx 0$.

Sección 10.1

- a) Continua b) Discreta c) Binaria d) Continua
- a) está en control b) tiene una capacidad alta
- a) Falso b) Falso c) Verdadero d) Falso

Sección 10.2

1. a) $LCL = 0$, $UCL = 10.931$ b) $LCL = 0$, $UCL = 4.721$ c) $LCL = 20.258$, $UCL = 27.242$
d) $LCL = 20.358$, $UCL = 27.142$
3. a) $LCL = 0$, $UCL = 0.277$. La varianza está en control.
b) $LCL = 1.034$, $UCL = 1.186$. El proceso está fuera de control por primera vez en la muestra 17.
c) Los límites 1σ son 1.085, 1.135; los límites 2σ son 1.0596, 1.1604. El proceso está fuera de control la primera vez en la muestra 8, donde dos de las tres últimas muestras están debajo del límite de control 2σ .
5. a) 43.86 b) 12.03 c) 0.985 d) 17
7. a) 0.126 b) 0.237 c) 0.582 d) 256
9. a) $LCL = 0.0588$, $UCL = 3.863$. La varianza está en control.
b) $LCL = 197.292$, $UCL = 202.340$. El proceso está en control.
c) Los límites 1σ son 198.975, 200.657; los límites 2σ son 198.133, 201.499. El proceso está en control.
11. a) $LCL = 0$, $UCL = 0.971$. La varianza está en control.
b) $LCL = 9.147$, $UCL = 10.473$. El proceso está en control.
c) Los límites 1σ son 9.589, 10.031; los límites 2σ son 9.368, 10.252. El proceso está fuera de control por primera vez en la muestra 9, donde dos de las tres últimas muestras están debajo del límite de control 2σ .
13. a) $LCL = 0$, $UCL = 6.984$. La varianza está fuera de control en la muestra 8. Después de eliminar esta muestra, $\bar{\bar{X}} = 150.166$, $\bar{R} = 6.538$, $\bar{s} = 2.911$. Los nuevos límites para el diagrama S son 0 y 6.596. Ahora la varianza está en control.
b) $LCL = 145.427$, $UCL = 154.905$. El proceso está en control.
c) Los límites 1σ son 148.586, 151.746; los límites 2σ son 147.007, 153.325. El proceso está en control (recuerde que la muestra 8 se ha eliminado).

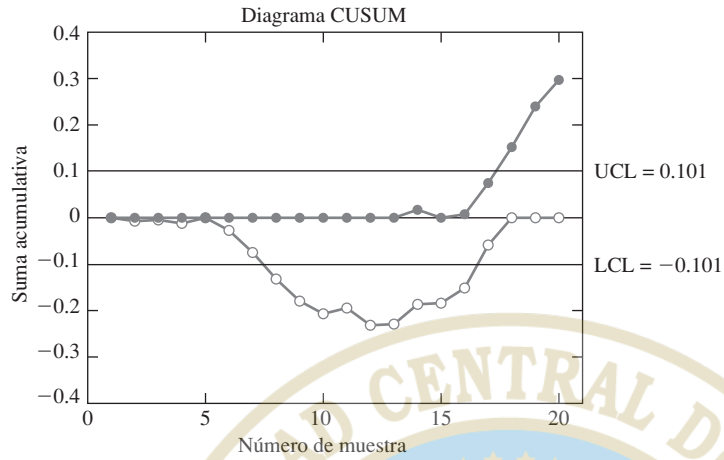
Sección 10.3

1. La recta central es 0.0547, LCL es 0.00644, UCL es 0.1029.
3. Sí, los límites 3σ son 0.0254 y 0.2234.
5. iv)
7. Estaba fuera de control. El UCL es 45.82.

Sección 10.4

1. a) No se necesita eliminar muestras. b) $\sigma_{\bar{x}} = (0.577)(0.131)/3 = 0.0252$

c)



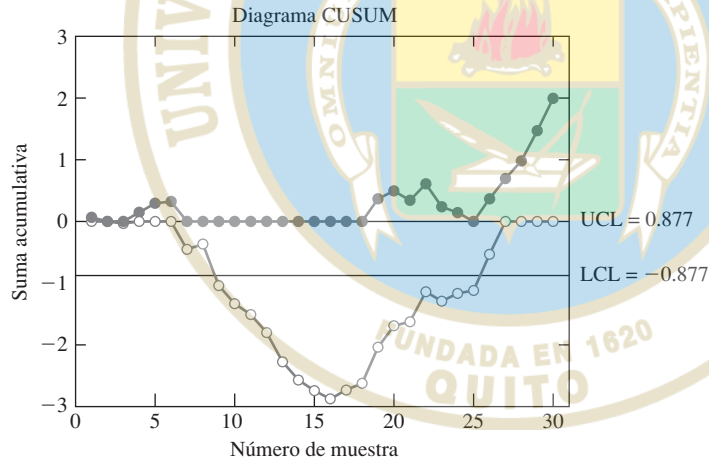
d) El proceso está fuera de control en la muestra 8.

e) Las reglas de Western Electric especifican que el proceso está fuera de control en la muestra 8.

3. a) No es necesario eliminar muestras.

$$b) \sigma_{\bar{x}} = (0.577)(1.14)/3 = 0.219$$

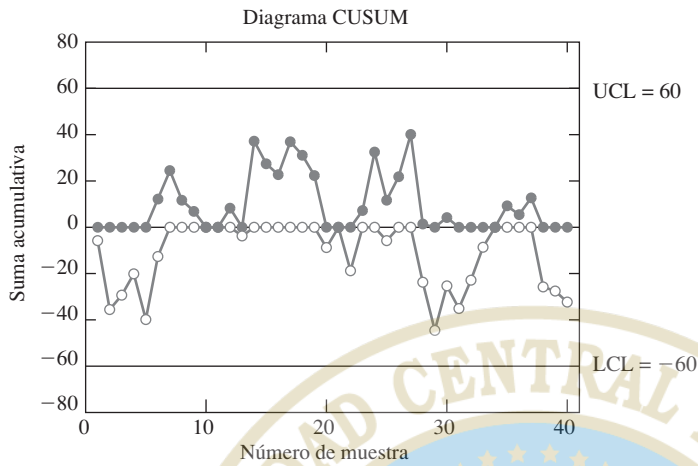
c)



d) El proceso está fuera de control en la muestra 9.

e) Las reglas de Western Electric especifican que el proceso está fuera de control en la muestra 9.

5. a)



b) El proceso está en control.

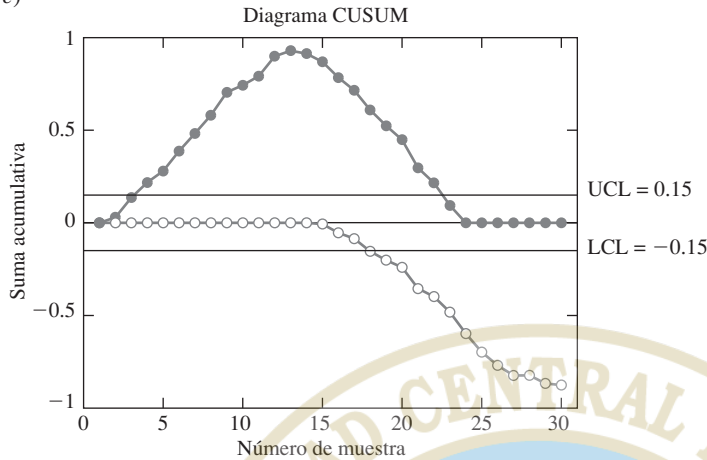
Sección 10.5

1. a) $C_{pk} = 2.303$ b) Sí. Puesto que $C_{pk} > 1$, la capacidad del proceso es aceptable.
3. a) 0.20 b) 3.071
5. a) $\mu \pm 3.6\sigma$ b) 0.0004 c) Probablemente. La aproximación normal quizá sea imprecisa en las colas.

Ejercicios adicionales para el capítulo 10

1. La recta central es 0.0596, LCL es 0.0147, UCL es 0.1045.
3. a) LCL = 0, UCL = 0.283. La varianza está en control.
 b) LCL = 4.982, UCL = 5.208. El proceso está fuera de control en la muestra 3.
 c) Los límites 1σ son 5.057, 5.133; los límites 2σ son 5.020, 5.170. El proceso está fuera de control por primera vez en la muestra 3, donde una media de la muestra está arriba del límite de control superior 3σ .
5. a) No es necesario eliminar muestras. b) $\sigma_{\bar{x}} = (1.023)(0.110)/3 = 0.0375$

c)



d) El proceso no está en control en la muestra 4. e) La reglas de Western Electric especifican que el proceso está fuera de control en la muestra 3.

7. a) LCL = 0.0170, UCL = 0.0726. b) Muestra 12
c) No, esta causa especial mejora el proceso. Ésta debe preservarse en lugar de eliminarse.

Apéndice B

1. $\frac{\partial v}{\partial x} = 3 + 2y^4$, $\frac{\partial v}{\partial y} = 8xy^3$
2. $\frac{\partial w}{\partial x} = \frac{3x^2}{x^2 + y^2} - \frac{2x(x^3 + y^3)}{(x^2 + y^2)^2}$, $\frac{\partial w}{\partial y} = \frac{3y^2}{x^2 + y^2} - \frac{2y(x^3 + y^3)}{(x^2 + y^2)^2}$
3. $\frac{\partial z}{\partial x} = -\operatorname{sen} x \operatorname{sen} y^2$, $\frac{\partial z}{\partial y} = 2y \cos x \cos y^2$
4. $\frac{\partial v}{\partial x} = ye^{xy}$, $\frac{\partial v}{\partial y} = xe^{xy}$
5. $\frac{\partial v}{\partial x} = e^x (\cos y + \operatorname{sen} z)$, $\frac{\partial v}{\partial y} = -e^x \operatorname{sen} y$, $\frac{\partial v}{\partial z} = e^x \cos z$
6. $\frac{\partial w}{\partial x} = \frac{x}{\sqrt{x^2 + 4y^2 + 3z^2}}$, $\frac{\partial w}{\partial y} = \frac{4y}{\sqrt{x^2 + 4y^2 + 3z^2}}$, $\frac{\partial w}{\partial z} = \frac{3z}{\sqrt{x^2 + 4y^2 + 3z^2}}$
7. $\frac{\partial z}{\partial x} = \frac{2x}{x^2 + y^2}$, $\frac{\partial z}{\partial y} = \frac{2y}{x^2 + y^2}$
8. $\frac{\partial v}{\partial x} = \frac{2xy}{x^2y + z} - ze^{y^2} \operatorname{sen}(xz)$, $\frac{\partial v}{\partial y} = \frac{x^2}{x^2y + z} + 2ye^{y^2} \cos(xz)$, $\frac{\partial v}{\partial z} = \frac{1}{x^2y + z} - xe^{y^2} \operatorname{sen}(xz)$
9. $\frac{\partial v}{\partial x} = \sqrt{\frac{y^5}{x}} - \frac{3}{2} \sqrt{\frac{y^3}{x}}$, $\frac{\partial v}{\partial y} = 5\sqrt{xy^3} - \frac{9}{2}\sqrt{xy}$
10. $\frac{\partial z}{\partial x} = \frac{xy \cos(x^2y)}{\sqrt{\operatorname{sen}(x^2y)}}$, $\frac{\partial z}{\partial y} = \frac{x^2 \cos(x^2y)}{2\sqrt{\operatorname{sen}(x^2y)}}$



ÍNDICE ANALÍTICO

- 2³ Experimento factorial
 efectos, suma de los cuadrados de los, 696
 errores, suma de los cuadrados de los, 696
 estimación de los efectos, 695
 notación, 693
 prueba *F*, 696-697
 tabla de signos, 694
 tabla del análisis de la varianza, 697
- 2^p Experimento factorial
 efectos, suma de los cuadrados de los, 700
 errores, suma de los cuadrados de los, 700
 estimación de los efectos, 700
 prueba *F*, 700
 sin réplica, 700
 tabla de signos, 701
- A**
Aleatorización dentro de los bloques, 685
Análisis de confiabilidad, 81
Análisis de la varianza
 unidireccional, *véase* unidireccional, análisis de la
Análisis de la varianza bidireccional
 diseño balanceado, 659
 diseño completo, 659
 diseño factorial completo, 659
 efecto columna, 661
 efecto renglón, 661
 efectos estimados de efectos, 663
 efectos principales, 661
 error de la suma de cuadrados, 665
 F prueba, 667
 factor columna, 659
 factor renglón, 659
 hipótesis nula, 664
 identidad del análisis de la varianza, 666
 interacción, 661
 interacción suma de cuadrados, 665
 interpretación de efectos principales, 669-672
 media cuadrática, 667
 media de la celda, 662
 media de la columna, 663
 media de renglón, 663
 modelo aditivo, 662
 suma de cuadrados por columna, 665
 suma de cuadrados por renglón, 665
 suma total de cuadrados, 665
 supuestos para el, 665
 tabla del análisis de la varianza, 666
 una observación por celda, 676
- Análisis de la varianza unidireccional
 hipótesis nula, 626
 identidad del análisis de la varianza, 636
 Modelo de efectos fijos, 624, 639
 Modelos de efectos aleatorios, 639-640
 prueba *F*, 630
 suma de los cuadrados de los errores, 627
 suma total de los cuadrados, 632
 supuestos en, 629
 tabla del análisis de la varianza, 632
 tratamiento de la suma de los cuadrados, 627
Aproximadamente normal, *véase* Teorema del límite central
- B**
Bondad del ajuste, 500-502
- C**
Capacidad del proceso, 755-759
 vs. control del proceso, 725
Causas atribuibles, 724
Causas azarosas, 723
Causas comunes, 723
Causas especiales, 724
Coeficiente de correlación, 476-477
 cómo funciona el, 478, 481
 mide sólo la relación lineal, 482
 midiendo la bondad del ajuste, 500-502
 y datos atípicos, 482-483
 y la proporción de la varianza explicada por la
 regresión, 502
Coeficiente de determinación
 en regresión lineal simple, 501
 en regresión lineal múltiple, 560
 y la proporción de la varianza explícita por regresión, 502
Coeficiente de variación, 174
Coeficientes de la regresión, 492
 intervalos de confianza para los, 513, 563-564
 pruebas de hipótesis para los, 513-515, 563-564

Coeficientes mínimos cuadrados
 desviación estándar de los, 510
 distribuidos normalmente, 510
 en la regresión lineal simple, 493
 en la regresión múltiple, 557
 insesgados, 510
 relación con el coeficiente de correlación, 499
 Colinealidad, 577-579
 Combinación lineal de mediciones,
 incertidumbre en la, 163
 límite superior para la incertidumbre en la, 169
 Combinación lineal de variables aleatorias, 114, 142
 media de la, 114, 142
 varianza de la, 116, 142-143
 Combinaciones, 65-66
 Comparaciones múltiples
 análisis de la varianza bidireccional, 674-676
 análisis de la varianza unidireccional, 650-655
 diseños de bloques completamente aleatorios,
 688-689
 método de Bonferroni, 459-460, 650
 método de Tukey, 675, 689
 método de Tukey-Kramer, 652
 Confusión entre variables,
 y regresión múltiple, 574-577
 Conteo
 combinaciones, 65-66
 permutaciones, 63-65
 principio fundamental del, 63
 Corrección por continuidad, 275-278
 para la distribución de Poisson, 278
 precisión de la, 278
 Correlación, 140-142, 475-489
 e independencia, 142
 muestral, 477
 no implica causalidad, 483-484
 poblacional, 140, 487
 Correlación de variables, 478
 Correlación poblacional, 140, 487
 Covarianza, 136-140
 e independencia, 141-142
 C_p , véase Mallows C_p , véase Índice de capacidad del
 proceso
 C_{pk} , véase Índice de capacidad del proceso
 Calidad seis-sigma, 758
 Cuartil, 18-19
 primer cuartil, 19
 segundo cuartil, 19
 tercer cuartil, 19
 Cuarto de réplica, 704

D

Datos
 categóricos, 11
 cualitativos, 11
 cuantitativos, 11
 numéricos, 11
 Datos atípicos, 16
 datos extremos, 35
 distribución lognormal, 247
 distribución normal, 240
 eliminación de, 16
 en los diagramas de caja, 35-36
 y el coeficiente de correlación, 482-483
 y el uso de la media recortada, 17
 y la distribución t de Student, 324
 y la regresión lineal simple, 536-538
 Datos bivariados, 39
 Datos multivariados, 39-40
 Desviación estándar
 de una variable aleatoria continua, 104
 de una variable aleatoria discreta, 96
 Desviación estándar muestral, 13, 14
 muestral, véase Desviación estándar muestral
 poblacional, véase Desviación estándar poblacional
 Desviación estándar poblacional, 232
 Diagrama cuantil-cuantil, 268
 Diagrama CUSUM, 751-753
 Diagrama de caja, 35-39
 comparativos, 37-39
 representando datos atípicos en un, 35-36
 Diagrama de control de suma acumulativa, véase
 Diagrama CUSUM
 Diagrama de dispersión, 39-40
 Diagrama de puntos, 26-27
 Diagrama de tallo y hojas, 25-26
 Diagrama- c , 748-750
 límites de control para el, 749
 Diagrama-CC, 268
 Diagrama- p , 746-748
 límites de control para, 747
 Diagrama- R
 comparación con el diagrama- S , 741-742
 límites de control para el, 729
 pasos para el uso del, 732
 Diagrama- S
 comparación con el diagrama- R , 741-742
 límites de control para el, 739
 Diagrama- X
 límites de control para el, 731, 740
 pasos para usar el, 732

- Diagramas de control
 diagrama-*R*, véase Diagrama-*R*
 diagrama-*c*, véase Diagrama-*c*
 diagrama CUSUM, véase Diagrama CUSUM
 diagrama-*p*, véase Diagrama-*p*
 diagrama-*S*, véase Diagrama-*S*
 diagrama-*X*, véase Diagrama-*X*
 para atributos de datos, véase Diagrama-*p*
 para conteo de datos, véase Diagrama-*c*
 para datos binarios, véase Diagrama-*p*
 para datos variables, véase Diagrama-*X*
- Diagramas de probabilidad, 265-269
 interpretando los, 269
 para detectar efectos en experimentos factoriales, 703-704
- Diagramas de Venn, 52
- Diferencia honestamente significativa, 651
- Diferencia significativa mínima, 647
- Diferencia significativa mínima de Fisher, 647
- Diseño balanceado, 635, 659
- Diseño completo, 659, 685
- Diseño factorial completo, 659
 diseños de bloques completamente aleatorios, 689
 método de Tukey en los, 689
- Diseños de bloques completamente aleatorios, 683-689
 bidireccional ANOVA, 675
- Distribución
 análisis de la varianza bidireccional, 667
 análisis de la varianza unidireccional, 630
 en experimento factorial 23, 696
 para coeficientes en la regresión múltiple, 563-564
 para la diferencia entre medias, 399, 412
 para la diferencia entre proporciones, 404
 para la media, 375, 393
 para la pendiente y el intercepto, 513-515
 para la proporción, 388
 Prueba de Mann-Whitney, véase Prueba de la suma del rango de Wilcoxon para el problema de pruebas múltiples, 458-460
 prueba de nivel fijo, 443-447
 prueba *F*, véase prueba *F*
 Pruebas no paramétricas, véase Pruebas libres de
- Distribución de una cola, 374
 datos apareados, 417
 dos colas, 374
 error tipo I, 446-447
 error tipo II, 446-447
 nivel de significancia, 443-444
 pasos en el desarrollo, 372
 potencia, véase Potencia
- prueba del rango con signo, véase Prueba del rango con signo de Wilcoxon
- prueba del rango de la suma, véase Prueba de la suma del rango de Wilcoxon
- prueba *t*, véase Prueba *t* de Student
- P*-valor, véase *P*-valor
- rechazando la hipótesis nula, 370, 377-378, 444
- región de rechazo, 444
- relación con intervalos de confianza, 381-382
- simultáneas, véase Comparaciones múltiples
- Distribución binomial, 195-204
 aproximación de la normal a una, 274
 función de masa de probabilidad de una, 198
 media, 201
 suma de variables aleatorias de Bernoulli, 200
 varianza, 201
- Distribución binomial negativa, 224-227
 función de masa de probabilidad de la, 225
 media, 226
 suma geométrica de variables aleatorias, 226
 varianza, 226
- Distribución de Bernoulli, 192-194
 función de masa de probabilidad de una, 192-193
 media, 194
 varianza, 194
- Distribución de Erlang, 262
- Distribución de Poisson, 206-217
 aproximación a la binomial, 206-207, 208, 209
 aproximación a la normal de la, 278
 función de masa de probabilidad de la, 209
 media, 209
 varianza, 209
- Distribución de probabilidad, 91, 99
 Bernoulli, véase Distribución de Bernoulli
 binomial, véase Distribución binomial
 binomial negativa, véase Distribución binomial negativa
 exponencial, véase Distribución exponencial
F, véase Distribución *F*
 gamma, véase Distribución gamma
 gaussiana, véase Distribución normal
 geométrica, véase Distribución geométrica
 hipergeométrica, véase Distribución hipergeométrica
 Ji-cuadrado, véase Distribución Ji-cuadrada
 lognormal, véase Distribución lognormal
 multinomial, véase Distribución multinomial
 normal, véase Distribución normal
 Poisson, véase Distribución de Poisson
t, véase Distribución *t* de Student
 Weibull, véase Distribución de Weibull
- Distribución de rango studentizado, 651

Distribución exponencial, 250-257
 función de densidad de probabilidad, 250
 función de distribución acumulativa, 251
 media, 251
 propiedad de la falta de memoria de la, 252-253
 varianza, 251
 y su relación con un proceso de Poisson, 252

Distribución F , 439-440
 grados de libertad de la, 439-440

Distribución gamma, 259-262

Distribución gaussiana, *véase* Distribución normal
 función de masa de probabilidad de una, 224
 varianza, 224

Distribución geométrica, 223-224
 media, 224

Distribución hipergeométrica, 220-223
 función de masa de probabilidad de la, 221
 media, 222
 varianza, 222

Distribución J_i cuadrada, 262, 430, 433
 caso especial de la distribución gamma, 262

Distribución lognormal, 244-248
 datos atípicos y, 247
 función de densidad de probabilidad, 244
 media, 245
 relación con la normal, 244
 uso de la tabla z con la, 245-246
 varianza, 245

Distribución multinomial, 227-229

Distribución normal, 231-241
 datos atípicos, 240
 desviación estándar de la, 231
 función de densidad de probabilidad de la, 231
 media de la, 231
 mediana, 231
 población normal estándar, 232
 puntaje z , 232
 unidades estándar, 231
 varianza de la, 231

Distribución t de Student, 321-324
 en la prueba de hipótesis, *véase* Prueba t de Student
 grados de libertad para la, 321, 322, 341
 intervalos de confianza usando la, 324-325, 341, 343
 y datos atípicos, 324-325

Distribución Weibull, 262-264
 función de densidad de probabilidad de la, 262
 función de distribución acumulativa, 263
 media de la, 263
 varianza de la, 263

E

Efecto columna, 661
 cálculo del, 663

Efecto de la suma de cuadrados
 en el experimento factorial 2^p , 700
 en el experimento factorial 2^3 , 696

Efecto renglón, 661
 cálculo del, 663

Efectos principales, 661
 interpretación de los, 669-672

El kilogramo, 161

Eliminación hacia atrás, 594

Ensayo de Bernoulli, 192

Ensayo multinomial, 227, 429

Error
 aleatorio, 158
 sistemático, 158

Error aleatorio, 158

Error cuadrático medio
 análisis de la varianza bidireccional del, 667
 análisis de la varianza unidireccional del, 629

Error de especificación, 577

Error estándar de la media, 21

Error sistemático, 158

Error tipo I, 446-447

Error tipo II, 446-447

Errores
 análisis de la varianza bidireccional, 662
 análisis de la varianza unidireccional, 636
 en la regresión lineal simple, 492

Espacio muestral, 50
 con resultados equiprobables, 57

Esperanza, *véase* Media poblacional

Esperanza condicional, 133-134

Estadística, 23

Estadística bayesiana, 379

Estadística descriptiva, 2, 20

Estadística inferencial, 2

Estadística J_i cuadrada, 430, 433

Estimación bootstrap, 291-293
 e intervalos de confianza, 356-358
 estimando el sesgo con, 291-292
 no paramétrica, 292-293
 paramétrica, 292-293
 y prueba de hipótesis, 463-464

Estimación puntual, 300

Evento(s), 51
 complemento de, 52
 exhaustivos, 76
 independientes, 73

intersección de, 52
 mutuamente excluyentes, 53
 regla de multiplicación de, 74
 unión de, 52
 Eventos exhaustivos, 76
 Eventos independientes, 73
 regla de la multiplicación para, 74
 Eventos mutuamente excluyentes, 53
 Experimento completamente aleatorio, 625
 Experimento con una muestra, 10
 Experimento de múltiples muestras, 10
 Experimento factorial,
 2^p diseño, *véase* Experimento factorial 2^p
 2^3 diseño, *véase* Experimento factorial 2^3
 fraccionado, *véase* Experimento factorial fraccionado
 Experimento factorial fraccionado, 704-707
 cuarto de réplica en el, 704
 fracción principal en el, 705
 media réplica en el, 704
 seudónimos en, 706
 Experimentos de dos factores, 659-677
 Experimentos de un factor, 623-642, 646-655
F
 Factor columna, 659
 Factor renglón, 659
 Falla a detectar, 733
 Falsa alarma, 733
 Fórmula de propagación del error
 multivariados, 179
 los resultados sólo aproximan, 173, 179
 univariado, 173
 Fracción principal, 705
 Frecuencias, 22
 Frecuencias relativas, 22, 27
 Función de densidad de probabilidad, 99-101
 condicional, 132
 conjunta, 123-124
 marginal, 125
 Función de densidad de probabilidad condicional, 132
 Función de densidad de probabilidad conjunta, 123-124
 Función de densidad de probabilidad marginal, 125
 celda, 662
 columna, 663
 condicional, *véase* Esperanza condicional
 de una combinación lineal de variables aleatorias,
 114, 142
 de una función de variables aleatorias, 129-131
 de una variable aleatoria continua, 104
 de una variable aleatoria discreta, 95

gran, *véase* Gran media
 muestral, *véase* Media muestral
 poblacional, *véase* Media poblacional
 recortada, 17
 renglón, 663
 Función de distribución acumulativa
 continua, 102
 discreta, 93
 Función de masa de la probabilidad condicional, 131
 Función de masa de probabilidad, 91, 93
 condicional, 131
 conjunta, 123
 marginal, 123
 Función de masa de probabilidad conjunta, 123
 Función de masa de probabilidad de la marginal, 123
 Función gamma, 159
 Función riesgo, 264
G
 Gosset, William Sealy (estudiante), 210, 321
 Gráfica de los residuales, 527
 en la regresión múltiple, 564-565
 interpretación del, 528, 533, 539
 tendencia en el, 539
 Gran media
 muestra, 626, 662
 población, 637, 660
H
 Heterocedasticidad, 528
 Hipótesis alternativa, 369
 Hipótesis nula, 369
 análisis de la varianza bidireccional, 664
 análisis de la varianza unidireccional, 626
 elección de la, 379-380
 haciendo un ensayo, 369
 rechazando la, 370, 377-378, 444
 Histograma, 27-35
 bimodal, 32-33
 intervalos de clase, 27
 sesgado, 31-32
 simétrico, 31-32
 unimodal, 32-33
 Histograma de probabilidad, 97
 Homocedasticidad, 528
I
 i.i.d., *véase* Independientes e idénticamente distribuidos

- Identidad del análisis de varianza
 para el análisis de la varianza bidireccional, 666
 para el análisis de la varianza unidireccional, 636
 para la regresión lineal simple, 503
 para la regresión múltiple, 558
- Incertidumbre
 de una función de una medición, 173-177
 de una función de varias mediciones, 179-185
 en la media muestral, 165
 límite superior para una combinación lineal de mediciones dependientes, 169
 límite superior para una función no lineal de mediciones dependientes, 182
 relativa, *véase* Incertidumbre relativa y desviación estándar, 159
- Incertidumbre aleatoria, *véase* Incertidumbre
- Incertidumbre relativa
 de una función de una medición, 174-176
 de una función de varias mediciones, 182-184
 dos métodos para el cálculo de la, 175, 183
- Independientes e idénticamente distribuidos, 117
- Índice de capacidad del proceso
 C_p , 757
 C_{pk} , 756
 C_{pl} , 759
 C_{pu} , 759
- Interacción
 análisis de la varianza bidireccional, 661
 cálculo de, 663
 e interpretación de efectos principales, 669-672
 en la regresión múltiple, 557
 media cuadrática, 667
 suma de cuadrados, 665
- Interacción media cuadrática, 667
- Interacción suma de cuadrados, 665
- Intercepto
 intervalo de confianza para el, 513
 prueba de hipótesis para el, 513
- Intervalo de predicción, 519
- Intervalos de confianza
 basado en muestras aleatorias, 311
 con la distribución t de Student, 325, 341, 343
 de un lado, 310-311, 317, 328
 determinación del tamaño muestral, 308-309, 318
 estimación bootstrap e, 356-358
 grado de confianza, *véase* Grado de confianza
 para datos apareados, 348
 para diferencia de proporciones, 336
 para la diferencia de medias, 333, 341, 343
 para la media, 304, 325, 329
 para la pendiente y el intercepto, 513
 para la proporción, 316, 319
 para la respuesta media, 516
 para los coeficientes en regresión múltiple, 563, 564
 simultáneo, *véase* Comparaciones múltiples en muestras pequeñas, 325, 329, 341, 343
 y su relación con pruebas de hipótesis, 381-382
- Intervalos de confianza simultáneos, *véase* Comparaciones múltiples
- L**
 Ley de la probabilidad total, 77
 Ley física, 541
 Límite de confianza, 310-311, 317, 328
 Límite de especificación, 755
 Límites de control
 para el diagrama- c , 749
 para el diagrama- p , 747
 para el diagrama- R , 729
 para el diagrama- S , 739
 para el diagrama- X , 731, 740
 Linealizando el problema, 176, 184
- M**
 Mallows C_p , 593
 Mann-Whitney prueba de, *véase* Prueba de la suma del rango de Wilcoxon
 Media cuadrática
 para el error, *véase* Error medio cuadrático
 para la interacción, *véase* Interacción media cuadrática
 para tratamiento, *véase* Tratamiento de la media cuadrática
 Media de la celda, 662
 Media de la columna, 663
 Media del renglón, 663
 Media muestral, 13
 desviación estándar de la, 118, 144
 incertidumbre en la, 165
 media de la, 118, 144
 teorema del límite central, 270
 varianza de la, 118, 144
 y centro de masa de histograma, 30-31
- Media poblacional
 de una variable aleatoria continua, 104
 de una variable aleatoria discreta, 95
 intervalo de confianza para la, *véase* Intervalo de confianza
 prueba de hipótesis para la, *véase* Prueba de hipótesis
- Mediana
 muestral, 16-17
 poblacional, 105-107

Mediana muestral, 16-17
 Medidas dependientes
 límite superior para la incertidumbre de una combinación lineal de, 169
 límite superior para la incertidumbre en una función no lineal de, 182
 Mejores subconjuntos en regresión, 591-594
 Método de Bonferroni, 459-460, 650
 Método de Tukey
 unidireccional ANOVA, *véase* Método de Tukey-Kramer
 Método de Tukey-Kramer, 652
 Mitad de réplica, 704
 Moda, 18
 de una histograma, 32
 Modelo aditivo, 662
 Modelo de efectos aleatorios, 624, 639-640, 676
 Modelo de regresión cuadrática, 557
 Modelo de regresión polinomial, 556
 Modelo empírico, 541
 Modelo lineal, 492
 Modelo mixto, 677
 Modelos de efectos fijos, 624, 639, 676
 Muestra, 3
 agrupada, 10
 aleatoria estratificada, 10
 aleatoria simple, 3
 de conveniencia, 4
 ponderada, 10
 Muestra aleatoria, *véase* Muestra aleatoria simple
 Muestra aleatoria simple, 3
 Muestreo con reemplazo, 9

N

Navaja de Occam, 584
 Nivel de confianza
 interpretación del, 306, 307
 y probabilidad, 306
 Nivel de significancia, 443-444
 Nivel de significancia observado, 372
 Nivel de una prueba de hipótesis, *véase* Nivel de significancia
 Niveles de un factor, 624
 No correlacionados, 142, 478
 No sesgado, 158

P

Parámetro, 23
 Parsimonia, 584

Pendiente
 intervalo de confianza para la, 513
 prueba de hipótesis para la, 514-515
 Percentil
 de una muestra, 19
 de una población, 105-107
 Permutación, 63-65
 Población, 3
 conceptual, 6
 tangible, 6
 Población normal estándar, 232
 Potencia, 449-455
 análisis de la varianza unidireccional, 637-639
 depende de la hipótesis alternativa, 449, 451, 452
 determinada por el tamaño de la muestra, 452
 pasos en el cálculo de la, 449
 Precisión, 158
 Principio de la parsimonia, 584
 excepciones al, 584
 Principio fundamental de conteo, 63
 Probabilidad
 axiomas de la, 55
 condicional, 71
 diagramas de árbol, 79
 frecuentista, 379
 incondicional, 69
 regla de la multiplicación, 74
 regla de la suma, 58-59
 subjética, 379
 Probabilidad condicional, 71
 diagramas de árbol para la, 79
 Probabilidad de éxito, *véase* Proporción poblacional
 Probabilidad frecuentista, 379
 Probabilidad incondicional, 69
 Probabilidad subjética, 379
 Proceso de Poisson
 estimación de la razón, 213
 incertidumbre en la razón estimada, 214
 la razón estimada es insesgada, 214
 relación con la distribución exponencial, 252
 Promedio, 13
 Promedio del tiempo de la corrida, 734
 Propiedad de la falta de memoria, 252-253
 Proporción
 cálculo de, *véase* Proporción poblacional
 intervalo de confianza para la, *véase* Intervalo de confianza
 muestral, *véase* Proporción muestral
 poblacional, *véase* Proporción poblacional
 prueba de hipótesis para, *véase* Prueba de hipótesis

Proporción muestral, 22, 201
 como estimación de la proporción poblacional, 201
 incertidumbre en la, 202
 no sesgada, 202

Proporción poblacional
 cálculo de, 201
 intervalo de confianza para la, *véase* Intervalo de confianza
 prueba de hipótesis para la, *véase* Prueba de hipótesis

Prueba aleatorizada, 464-466

Prueba de hipótesis
 eligiendo la hipótesis nula, 379-380
 estimación bootstrap y, 463-464
 método de Bonferroni, 459-460, 650
 prueba Ji cuadrada, *véase* Ji cuadrada, prueba
 pruebas libres de distribución, *véase* Pruebas libres de punto crítico, 444

Prueba de la suma del rango, 425-427

Prueba de la suma del rango de Wilcoxon, 425-427

Prueba de nivel fijo, 443-447

Prueba de permutación, 465

Prueba de significancia, *véase* Prueba de hipótesis

Prueba del rango con signo, 420-425

Prueba del rango con signo de Wilcoxon, 420-425

Prueba estadística, 372

Prueba *F*
 análisis de la varianza bidireccional, 667
 análisis de la varianza unidireccional, 630
 en el experimento factorial 2^p , 700
 en el experimento factorial 2^3 , 696-697
 en la regresión múltiple, 587
 para la igualdad de varianzas, 439-442

Prueba Ji cuadrada, 429-436
 grados de libertad para, 430, 433
 para la homogeneidad, 431, 435
 para la independencia, 435, 436
 para probabilidades específicas, 429-431

Prueba *t* de Student
 dos muestras, 410, 412
 una muestra, 393

Prueba *z*, 371, 386, 393

Pruebas de hipótesis, *véase* Prueba de hipótesis

Pruebas de hipótesis simultáneas, *véase* Comparaciones múltiples

Pruebas de rango, *véase* Pruebas libres de distribución

Pruebas libres de distribución, 420-427
 prueba del rango con signo, 420-425
 prueba del rango de la suma, 425-427

Pruebas no paramétricas, *véase* Pruebas libres de distribución

Punto crítico, 444

Punto influyente, 537

p-valor, 369, 372
 interpretando el, 370-371, 372
 no es la probabilidad de que H_0 sea verdadera, 379

R

R^2 ajustado, 593

R^2 , *véase* Coeficiente de determinación

Rango, 18

Rango de Wilcoxon

Rango intercuartil, 35

Recta de mínimos cuadrados, 492-504
 bondad del ajuste de la, 500-502
 cálculo de la, 494-496
 no extrapolar, 498
 no se use cuando los datos no son lineales, 498-499

Región de rechazo, 444

Regla de la multiplicación para probabilidades, 74

Regla de la suma de probabilidades, 58-59

Reglas eléctricas de occidente, 737

Regresión lineal simple
 gráfica de los residuales contra tiempo, 540
 supuestos en la, 508, 527
 tabla del análisis de la varianza, 522
 transformación de variables en la, 531-536
 y datos atípicos, 536-538

Regresión múltiple
 coeficientes de mínimos cuadrados, 557
 modelo de regresión múltiple, 556
 prueba *F*, 587
 puede detectar error de especificación, 574-577
 selección del modelo, *véase* Selección del modelo
 sumas de cuadrados, 558
 supuestos en la, 558
 tabla del análisis de la varianza, 562
 y colinealidad, 577-579

Residual
 análisis de la varianza unidireccional, 628
 en la regresión lineal simple, 493-494

Respuesta media, 515-516
 intervalo de confianza para la, 516

Resumen estadístico, 13-23

S

Selección del modelo, 583-599
 es arte no una ciencia, 599
 excepción en el principio de la parsimonia, 584

navaja de Occam, 584
 principio de la parsimonia, 584
 Selección hacia delante, 594
 Sesgo, 158, 159
 de funciones no lineales, 173, 179
 Seudonónimos, 706
 Significancia estadística, 378-379, 444
 no es lo mismo significancia práctica, 380-381
 Simulación
 comparación con la propagación de errores, 286
 estimación bootstrap, *véase* Estimación bootstrap
 usada en el análisis de confiabilidad, 288-290
 usada en el cálculo de medias y varianzas, 285-286
 usada en el cálculo de probabilidades, 282-285
 usada en el cálculo de sesgo, 290-292
 usada en el desarrollo de pruebas de hipótesis,
 462-466
 usada en la construcción de intervalos de confianza,
 351-358
 usada en la estimación de la potencia, 466-467
 usada en la prueba de normalidad, 286-288
 Subgrupos racionales, 724
 Suma acumulativa, 751
 Suma de cuadrados
 para error, *véase* Error de la suma de cuadrados
 para interacción, *véase* Interacción de la suma de
 cuadrados
 para tratamiento, *véase* Tratamiento suma de cuadrados
 por columnas, *véase* Suma de cuadrados por columna
 por renglones, *véase* Suma de cuadrados por renglón
 total, *véase* Suma total de cuadrados
 Suma de cuadrados de la regresión
 en la regresión lineal simple, 501, 522
 en la regresión múltiple, 558, 562
 Suma de cuadrados por renglón, 665
 Suma de los cuadrados de los errores
 análisis de la varianza bidireccional, 665
 análisis de la varianza unidireccional, 627
 en el experimento factorial 2^n , 700
 en el experimento factorial 2^3 , 696
 en la regresión lineal simple, 501, 522
 en regresión múltiple, 558, 562
 Suma de los cuadrados por columna, 665
 Suma de variables aleatorias
 media de la, 114, 142
 teorema del límite central para, 270
 varianza de la, 116, 143
 Suma total de cuadrados
 análisis de la varianza bidireccional, 665
 análisis de la varianza unidireccional, 632

en múltiple regresión, 558, 562
 en regresión lineal simple, 501, 522

T

t distribución, *véase* Distribución *t* Student.
t prueba, *véase* Prueba *t* Student.
 Tabla de contingencia, 431
 Tabla de frecuencias, 27
 Tabla de signos, 694, 701
 Tabla del análisis de la varianza
 para el análisis de la varianza bidireccional, 666
 para el análisis de la varianza unidireccional, 632
 para el experimento factorial, 23, 697
 para la regresión lineal simple, 522
 para la regresión múltiple, 562
 Teorema de Bayes, 79
 Teorema del límite central, 270-279
 para la distribución binomial, 274
 para la distribución de Poisson, 278
 para la media de una muestra, 270
 para la suma, 270
 tamaño muestral necesario para la validación, 271, 272
 Tiempo de espera, 250
 Transformación de variables, 531-536
 Transformación potencia, 531
 Tratamiento de la media cuadrática, 629
 Tratamiento de suma de cuadrados, 627
 Tratamiento del efecto, 637
 Tratamientos, 624, 659

U

Unidades estándares, 231

V

Valor ajustado, 493
 Valor esperado, *véase* Media poblacional
 Variable aleatoria
 combinación lineal de, *véase* Combinación lineal de
 variables aleatorias
 conjuntamente continua, 123-128
 conjuntamente discreta, 121-123
 continua, *véase* Variable aleatoria continua
 discreta, *véase* Variable aleatoria discreta
 independiente, 115, 135
 suma de, *véase* Suma de variables aleatorias
 Variable aleatoria continua, 89, 99-107
 desviación estándar de una, 104

- función de distribución acumulativa, 102
- función de la densidad de probabilidad de una, 99-101
- media de una, 104
- varianza de una, 104
- Variable aleatoria discreta, 89, 91-98
 - desviación estándar de una, 96
 - función de distribución acumulativa de una, 93
 - función de masa de probabilidad de una, 91, 93
 - media de una, 95
 - varianza de una, 96
- Variable de respuesta, 624
- Variable dependiente, 492
- Variable independiente, 492
- Variables aleatorias conjuntamente continuas, 123-128
- Variables aleatorias conjuntamente discretas, 121-123
- Variables aleatorias independientes, 115, 135
- Variación del muestreo, 5
- Varianza
 - bidireccional, *véase* bidireccional, análisis de la varianza
 - combinada, *véase* Varianza combinada
 - de una combinación lineal de variables aleatorias, 116, 142-143
 - de una variable aleatoria continua, 104
 - de una variable aleatoria discreta, 96
 - muestral, *véase* Varianza muestral
 - poblacional, *véase* Varianza poblacional
- Varianza de una combinación
 - análisis de la varianza unidireccional, 631
 - y la distribución *t* de Student, 343, 411, 412
- Varianza muestral, 14
 - y el momento de inercia del histograma, 31
- Varianza poblacional
 - de una variable aleatoria continua, 104
 - de una variable aleatoria discreta, 96

