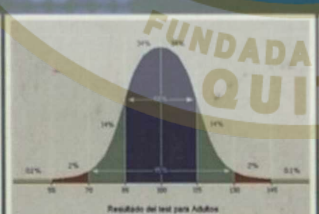
Germán González Bahamonde

MÉTODOS ESTADÍSTICOS

y Principios de Diseño Experimental

Tercera Edición



Premio Universidad Central del Ecuador 1976

Sistema Integrado de Bibliotecas
Biblioteca digital UCE
Calle Jerónimo Leyton y Gato Sobral – Cdla. Universitaria
+593 2230757 ext: 808 biblioteca.general@uce.edu.ec

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	Sistema integrado de bibliotecas	**PISO**	
		ESTANTE	
		BANDEJA	
CONVENIO INTERINSTITUCIONAL		**UNIVERSIDAD**	

MÉTODOS ESTADÍSTICOS Y PRINCIPIOS DE DISEÑO EXPERIMENTAL


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

519.5
G643
E6

Germán González Bahamonde, I.A., MSc.
Ex Profesor Principal de Estadística de la Facultad de Ciencias Agrícolas de la Universidad Central del Ecuador

MÉTODOS ESTADÍSTICOS Y PRINCIPIOS DE DISEÑO EXPERIMENTAL

Tercera Edición

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

PREMIO UNIVERSIDAD CENTRAL 1976

Universidad Central del Ecuador – Quito, 2010

EDITORIAL UNIVERSITARIA

229.11-Adquisición Dra. LBC-Dir. Bib. Ed. Uce. 2011-06-23-$ 10.00
56d/83811 ID = 952

© 1985. Reservados los derechos del autor.
© 2010. id. (IEPI No. 033705)
ISBN-978-9978-92-883-7

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST SAPIENTIA OMNIUM

INVENTARIO
FUNDADA EN 1620
QUITO
Compra ..... Donación ..X..... Canje .....
Proveedor Correspondencia U.C.E
Factura ..... Fecha 2010 Precio 10 ¢
Nº Inf. Recp ..... Fecha Inf .....
Nº Alta 5118606
Biblioteca General

CONTENIDO

Págs

Prólogo 13

PRIMERA PARTE

Capítulo 1 Introducción 19

1.1 Función de la Estadística y del Diseño Experimental 19
1.2 Definición 20
1.3 Campos en los que se aplica la Estadística 21
1.3.1 Genética 21
1.3.2 Nutrición 22
1.3.3 Comercialización 22
1.3.4 Industria 22
1.3.5 Medicina 23
1.3.6 Agronomía 23
1.3.7 Información general 23
1.4 Algunos datos históricos 23
1.5 La Estadística y el Método Científico 24

Capítulo 2. Variación - Variables 27

2.1 Variación ambiental y variación hereditaria 28
2.1.1 La variación provocada por causas ambientales 28
2.1.2 Las variaciones hereditarias 28
2.2 Variables 29
2.2.1 Variable cualitativa 29
2.2.2 Variable cuantitativa 29
2.3 Observaciones, hechos 29
2.4 Población y muestra 30
2.5 Distribuciones 32
2.6 Presentación de datos 34

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO
BIBLIOTECA

Capítulo 3. Muestreo 39

3.1 Poblaciones 39
3.2 Muestras 40
3.3 Estimación 40
3.4 Teorema del límite central 41
3.5 Tipos de muestreo 41
3.5.1 Muestreo aleatorio simple 42
3.5.2 Muestreo probabilístico 43
3.5.3 Muestreo estratificado 44
3.5.4 Muestreo sistemático 45
3.5.5 Muestreo por áreas 45
3.5.6 Muestreo dirigido 46
3.5.7 Muestreo por cuotas 46

Capítulo 4. Medidas de tendencia central y dispersión 47

4.1 Simbología matemática 47
4.2 Funciones 51
4.3 Redondeo de cifras 51
4.4 Parámetros y estadísticas 51
4.5 La media y otras medidas de tendencia central 51
4.5.1 La media aritmética 53
4.5.2 La mediana 55
4.5.3 El modo 55
4.6 La desviación típica y otras medidas de dispersión 55
4.6.1 El rango 55
4.6.2 La variancia 56
4.6.3 La desviación típica 59
4.7 Cambios en las observaciones y sus influencias en μ y σ^2 60
4.7.1 Si se suma o resta 61
4.7.2 Si se multiplica 61
4.8 Desviación típica de las medias 62
4.9 Coeficiente de variación (C.V.) 63
4.10 Modelo lineal aditivo 64
4.11 Intervalos de confianza 64

Capítulo 5. Cuadros y curvas de frecuencia – Histogramas 69

5.1 Cuadros de frecuencia 69
5.2 Histogramas y polígonos de frecuencia 71

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST QUAM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 6. Probabilidad 77

6.1 Sucesos independientes. 79
6.2 Sucesos dependientes. 81
6.3 Sucesos mutuamente excluyentes 82
6.4 Análisis combinatorio. 84
6.5 Factorial N 84
6.6 Permutaciones. 85
6.7 Combinaciones 86

Capítulo 7. Distribuciones teóricas de frecuencias 89

7.1 La distribución binomial 89
7.1.1 La expansión binomial 94
7.1.2 Algunas propiedades de la distribución binomial. 95
7.2 La distribución normal 96

Capítulo 8. Pruebas de hipótesis. 105

8.1 La hipótesis nula, H_0 107
8.2 Prueba de hipótesis y nivel de significación 108
8.3 Errores tipo I y tipo II 110
8.4 Potencia de la prueba. 110
8.5 Pruebas de una y de dos colas 110

Capítulo 9. La distribución de X^2 115

9.1 Cálculo de X^2 116
9.2 Cuadros con dos criterios de clasificación. 120
9.3 Corección de Yates 123

Capítulo 10. La distribución de t 125

10.1 Prueba de t para observaciones no pareadas 126
10.2 Prueba de t para observaciones pareadas. 128

Capítulo 11. Correlación 133

11.1 Coeficiente de correlación 135
11.2 Propiedades del coeficiente de correlación 136

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST SUMMUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 12. Regresión 139

12.1 Ecuación de la línea recta 141
12.2 El método de los cuadrados mínimos 143
12.3 Cálculo del coeficiente de regresión y de la ecuación de regresión 143
12.4 Valores ajustados 145
12.5 Fuentes de variación en regresión 146
12.6 Desviación y límites de confianza 147
12.7 Propiedades y suposiciones en regresión lineal 148

Capítulo 13. Análisis de covariancia (ANACOVA) 149

13.1 Usos del ANACOVA 151
13.2 Suposiciones en el ANACOVA y el modelo lineal aditivo 152
13.3 Modelo matemático 153
13.4 Ilustración 155
13.5 Ajuste de medias 158
13.6 Ganancia en precisión por covariancia 159

SEGUNDA PARTE

Capítulo 14. Experimentación 163

14.1 Introducción 163
14.2 Objetivos 165
14.3 Fases de la experimentación 166
14.4 Clases de experimentos 167
14.5 Tratamientos 169
14.6 Selección de tratamientos 170
14.7 Unidad experimental 171
14.8 Repeticiones 173
14.9 Disposición de los bloques 176
14.10 Factores que determinan el número de repeticiones 177
14.11 Error experimental 177
14.12 Control del error experimental 178
14.13 Selección de sitio y del material experimental 179
14.14 Precisión relativa del experimento 180
14.15 Randomización 182
14.16 Selección del diseño experimental 182
14.17 Inferencia estadística 183

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 15. Análisis de variancia: ordenación unilateral 185

15.1 El análisis de variancia 185
15.2 Grados de libertad 187
15.3 Diseño completamente al azar **(DCA)** 188
15.3.1 **DCA** con repetición desigual 189
15.3.2 **DCA** con igual repetición 191
15.4 Comparaciones entre tratamientos 193
15.4.1 Diferencia mínima significativa 193
15.4.2 Prueba de rango múltiple de Duncan 195
15.4.3 Prueba de Tukey 197
15.4.4 Prueba de Scheffé 198
15.5 Modelo lineal 199
15.6 Análisis de variancia con submuestras 200

Capítulo 16. Análisis de variancia: ordenación múltiple 205

16.1 Diseño de bloques completos al azar (BCA) 206
16.1.1 Modelo lineal aditivo 210
16.1.2 Cálculo de parcelas perdidas 211
16.1.3 Grado de eficiencia de BCA vs. DCA 213
16.2 Diseño de bloques completos al azar con submuestras 214
16.3 Cuadradro latino 217
16.3.1 Modelo matemático para cuadrado latino 218
16.3.2 Parcelas perdidas 219
16.3.3 Eficacia relativa 220
16.3.4 Modelo lineal 221
16.4 Transformaciones 221
16.4.1 Transformación a raíz cuadrada 222
16.4.2 Transformación logarítmica 222
16.4.3 Transformación arco-seno 223

Capítulo 17. Experimentos factoriales 225

17.1 Factores, niveles, factoriales 225
17.2 Modelo lineal 232
17.3 Comparaciones de grados de libertad individuales 233
17.4 Respuesta superficial; grados de libertad individuales: tratamientos igualmente espaciados 236
17.4.1 Respuesta superficial; "diseño" de diamante doble 241
17.5 Experimentos en confundido 242

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO
BIBLIOTECA GENERAL

Capítulo 18. Parcelas divididas 251

18.1 Valores perdidos en parcelas divididas 255
18.2 Modelo lineal aditivo 256

Capítulo 19. Cálculo de aptitud combinatoria general y específica, en cruzamientos dialélicos 257

19.1 Modelo matemático 259
19.2 El experimento 260
19.3 Método de cruzamiento 261
19.4 Semillero y proceso de infección 262
19.5 Toma de notas 262
19.6 Análisis estadístico 265
19.6.1 Cálculo de la S.C. para a.c.g. 267
19.6.2 Cálculo de la S.C. para a.c.e. 267
19.6.3 Cálculo de la S.C. para recíprocos (28 cruzamientos simples) 269
19.6.4 Resumen del cálculo de las S.C. de los componentes 270
19.6.5 Cálculo de a.c.g. para cada combinación ($\widehat{g}_i$) 272
19.6.6 Cálculo de a.c.e. para cada combinación ($\widehat{e}_{ij}$) 272
19.7 Conclusiones 274

Referencias 275
Apéndice 276
Tablas 277

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM STUDIORUM SAPIENTIA
FUNDADA EN 1620
QUITO

PRÓLOGO

Esta tercera edición, al igual que las dos primeras, tiende a presentar los Métodos Estadísticos como base de la experimentación. Se ha tratado de despertar la inquietud del estudiante hacia los problemas de la investigación, así como en lo referente a los Métodos Estadísticos, en los que aquella se basa. Para el estudioso de la Estadística y por insinuación de algunos colegas, se ha añadido la Tabla No XIII del Apéndice, la que incluye 312 referencias adicionales especializadas, referentes a temas Estadísticos y de Diseño Experimental.

La decisión de reunir notas de clase y llevar a cabo la primera edición de este manual, nació de la inquietud presentada por estudiantes y profesionales jóvenes —a través de varios años de dictar la cátedra de Estadística— en el sentido de poder contar con un texto en español, de fácil lectura y simbología simple. Si se ha satisfecho este deseo, se habría cumplido el propósito para el que se preparó este material.

De acuerdo con la intención de esta publicación, se ha dejado a un lado las bases matemáticas de los Métodos Estadísticos, las que, por otro lado, pueden ser encontradas en varias obras de Estadística matemática.

Al preparar esta revisión, el principal objetivo fue mejorar la primera y segunda ediciones corrigiendo algunos errores deslizados en las mismas, añadiendo ejemplos usa-

dos en clase y ampliando tópicos sobre los que —a nivel de tesis de grado y experimentación de campo— se ha recibido consultas en forma repetida. Así, por ejemplo, en la Facultad se ha hecho popular el uso de los diseños completamente al azar y bloques completos al azar con submuestras, por cuya razón se incluye los ejemplos respectivos, mostrando todo el procedimiento de cálculo.

En igual forma, investigaciones de tipo sociológico, industrial, económico, etc., etc., y que se realizan mediante entrevistas o encuestas, sugirieron la necesidad de incluir un capítulo completo sobre muestreo.

Los trece capítulos de la primera parte, pueden ser usados en la secuencia dada, para cualquier curso de Métodos Estadísticos. De hecho, la división en capítulos fue conformada de acuerdo con la didáctica usada por el autor, en el Curso que se ofreció en la Facultad de Ciencias Agrícolas de la Universidad Central del Ecuador, antes que siguiendo una ordenación estrictamente estadística.

La finalidad de esta ordenación ha sido permitir al estudiante una comprensión clara de cada tema, en forma separada, con el objeto de que los conocimientos puedan ser aplicados en la práctica, directamente de cada capítulo, sin tener que revisar previamente capítulos anteriores.

En diferentes textos se presenta, por ejemplo, las distribuciones de t y de X^2, juntas en varios capítulos. Con el objeto de que cada una de ellas pueda ser aplicada directamente en la resolución de problemas específicos, el autor prefirió dedicar un capítulo a t y otro a X^2.

En este sentido, prácticamente todos los capítulos son independientes unos de otros; si bien, como es natural, existen algunos en los que se sienta las bases para una mejor comprensión de capítulos posteriores.

En el capítulo 1, se dan algunas definiciones básicas y se hace énfasis sobre la universalidad del uso de la Estadística y del Diseño Experimental. A este respecto, si bien la aplicación de los principios enunciados es general para todos los campos de investigación, se hace hincapié en su utilización dentro de la experimentación agropecuaria.

En el capítulo 2 se introduce la idea sobre variabilidad y, en la forma más simple, sobre la teoría de muestreo, ilustrando esta última con la descripción de una práctica muy usada por los agricultores —aunque no bien comprendida por muchos de ellos— y que se refiere a las "experiencias en papas".

En el capítulo 4 se presenta la simbología que se usa en los capítulos posteriores. En las secciones 4.6.2 y 4.6.3, se ilustra el cálculo de la variancia y de la desviación típica, el que es fundamental para los análisis de variancia, covariancia, correlación y regresión.

En el capítulo 6 se ha tratado de presentar todo lo relacionado con probabilidad —tópico que generalmente causa confusión al estudiante— en la forma más simple.

Los capítulos 14, 15, 16, 17, y 18, dentro de la Segunda Parte, pueden ser usados para cualquier Curso de Diseño Experimental, en el orden dado.

El capítulo 14 presenta los diferentes aspectos que contempla la experimentación, los que deben ser tomados en cuenta por el estudiante que va a planificar su tesis de grado, así como por el profesional que va a iniciar un ensayo experimental. Al escribir este capítulo, el autor ha considerado las condiciones ambientales ecuatorianas, con especial referencia a los cambios topográficos y microclimáticos que prevalecen en el país.

En algunos casos se ha usado datos supuestos, empleando números pequeños, enteros, con el objeto de facilitar los cálculos y permitir que el estudiante pueda reproducirlos, sin pérdida de tiempo; lo que interesa es el procedimiento.

Si bien el capítulo 19 constituye un problema específico de fitomejoramiento el que, quizás, debería ir en el Apéndice, ha sido añadido como un capítulo más dentro del presente trabajo, por tratarse de aspectos de cálculo estadístico, aplicados al mejoramiento. Para ilustrar el procedimiento de cálculo de aptitud combinatoria general y específica, se utilizaron datos provenientes de cruzamientos dialélicos en alfalfa, analizados por el autor. Sin embargo, los métodos presentados pueden ser aplicados a cualquier cultivo y para mejoramiento animal, en todos los casos que incluyan genética cuantitativa.

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

El autor deja constancia de su reconocimiento a todas las personas que —habiendo hecho valiosos comentarios y sugerencias sobre las dos primeras ediciones, han insinuado, desde hace varios años, la conveniencia de publicar la tercera, por haberse agotado las dos anteriores. Así mismo, el autor agradece al profesor E. S. Pearson, editor de Biometrika y al Dr. P. G. Hoel, profesor de Matemáticas de la Universidad de California, por su gentil autorización para reproducir las Tablas IV, V y I, II, respectivamente, del Apéndice; a la Facultad de Ingeniería Agronómica y Medicina Veterinaria de la Universidad Central, por el permiso concedido para reproducir las Tablas Nos. VI, VII, VIII, IX, X y XI, compiladas por el Dr. Galo Romero; así como al Dr. R. J. Walby, editor del Australian Journal of Biological Sciences y al Dr. Bruce Griffing, director del Departamento de Genética de la Universidad de Ohio, por el consentimiento otorgado para usar uno de los métodos de Griffing (1956), para el cálculo de aptitud combinatoria, en el capítulo 19. Finalmente, el autor agradece al licenciado Antonio Negrete Puebla, jefe del Almacén Universitario y al director de la Editorial Universitaria, MSc. Édison Benavides por la nítida impresión de este trabajo.

GERMÁN GONZÁLEZ BAHAMONDE.

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM SCIENTIA
FUNDADA EN 1620
QUITO



Primera Parte

MÉTODOS ESTADÍSTICOS


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 1

INTRODUCCIÓN

1.1 Función de la Estadística y del Diseño Experimental

El progreso en el campo agropecuario —así como en otros a los que el hombre dedica su atención— tiene que basarse en la investigación o experimentación.

Esta verdad, que ha sido reconocida en los últimos años, ha permitido el incremento del trabajo experimental en muchos países y ha estimulado a grupos cada vez más numerosos de profesionales jóvenes, para que se adiestren y dediquen su tiempo y esfuerzo a la investigación, en los diferentes campos.

La experimentación, en su forma moderna, tiene que planearse y ejecutarse sobre bases científicas para, de esta manera, poder llegar a conclusiones válidas y confiables, que más tarde se traducirán en recomendaciones de tipo práctico para el público y que significan incremento de rendimientos, calidad, eficacia, etc.

El conocimiento de los Métodos Estadísticos permite una clara compresión del **Diseño Experimental,** a base del que es posible identificar el problema, plantear una hipótesis de trabajo, conducir el ensayo e interpretar sus resultados en forma correcta.

1.2 Definición

La palabra Estadística es de uso generalizado. Desafortunadamente —aún para muchos hombres cultos— es una simple recolección de datos. Así se oye hablar, por ejemplo, de estadísticas de consumo, de accidentes, censos de población, etc., que —si bien son colecciones importantes de datos, a base de las que se puede hacer un verdadero trabajo estadístico posterior— no representan sino una pequeña porción del alcance de la Estadística.

La simple colección y ordenación de datos, se llama **Estadística Descriptiva,** mientras que la obtención sistemática de una o más conclusiones a partir de los datos, se llama **Inferencia Estadística.** Así vemos que los datos obtenidos por la Estadística Descriptiva, constituyen un fin para ésta, pero tan solo un medio para llegar a la Inferencia Estadística.

Steel y Torrie (1960), definen a la Estadística como "la ciencia pura y aplicada que crea, desarrolla y aplica procedimientos, en tal forma, que se pueda evaluar la certeza de la inferencia inductiva".

De acuerdo con Spiegel (1961), "la Estadística es la ciencia que usa métodos para reunir, organizar, resumir y analizar datos, así como para obtener conclusiones válidas y tomar decisiones razonables, a base de tales análisis".

Para la mayoría de científicos, la Estadística es lógica o sentido común, en combinación con procedimientos aritméticos. La lógica daría el método para la toma de datos y la aritmética proporciona el material sobre el que ha de basarse la inferencia. La aritmética constituye generalmente la parte rutinaria del procedimiento.

Finalmente, se dice que la Estadística es el estudio científico del análisis de datos numéricos.

De acuerdo con lo expuesto, se usa en forma común la palabra **Estadística,** por lo que, en realidad, corresponde a **Métodos Estadísticos.**

1.3 Campos en los que se aplica la Estadística

En los últimos años se ha podido apreciar cambios notables en la Estadística. Nuevas teorías y prácticas han estado surgiendo continuamente. Seguidamente, se describen algunas aplicaciones de la Estadística.

1.3.1 Genética. Ciencia de comunicación biológica entre generaciones.

Lo que se transmite de padre a hijo (tercer principio dilucidado por Johan Gregor Mendel, en 1865).

Mendel tenía una mente brillante —física, matemática, estadística—. Su capacidad como experimentador para seleccionar individuos que diferían en ciertas características cualitativas y su meticulosidad para realizar, tabular y analizar contajes de las progenies resultantes de sus cruzamientos, permitió llegar, más tarde, a una clara comprensión de la genética de poblaciones (Srb y colaboradores, 1965).

Algunos atributos importantes de plantas, animales y del hombre, se ubican en escalas continuas de medida y se expresan mejor como kilogramos, centímetros, cuocientes de inteligencia, etc. Estas variaciones son de naturaleza cuantitativa y se llaman caracteres cuantitativos. La descripción y análisis de éstos, requiere de métodos especiales, dados por la rama de las matemáticas, llamada Estadística.

El conocimiento de las leyes de azar **(probabilidad),** es básico para comprender la transmisión de factores heredables. Así podemos preguntar, "en familias que tienen dos hijos, que proporción de ellas tendrán dos varones?; varón y mujer?; dos mujeres?; ojos claros y cafés? La **expansión binomial** daría respuesta a estas preguntas.

La prueba de X^2 indicaría si el número y clase de progenie obtenida en cruzamientos de alfalfa, está de acuerdo a la hipótesis mendeliana. Tal sería el caso del cálculo de recombinaciones de genes, por sobrecruzamiento, cuando los genes se hallan ligados en cromosomas homólogos.

UNIVERSIDAD CENTRAL DEL ECUADOR
BIBLIOTECA

Dentro del trabajo de creación de nuevas variedades, la determinación de la **aptitud combinatoria general** y **especifica,** utiliza métodos y modelos estadísticos como los presentados por Sprague y Tatum (1942), Griffing (1956), Gates y Wilcox (1964), entre otros, que son considerados clásicos en el campo de fitomejoramiento.

Por considerar de interés para los fitomejoradores, hemos añadido el capítulo 19, en el que se demuestra el uso de uno de los métodos de Griffing (1956), para él cálculo de aptitud combinatoria general y específica. Para el efecto, se utilizaron datos provenientes de cruzamientos dialélicos en alfalfa, analizados por el que escribe.

1.3.2 Nutrición. Tanto en nutrición humana como animal, el control de calidad de alimentos es utilizado extensamente para mantener la uniformidad de productos elaborados. Continuamente se está probando la bondad de nuevas dietas, que se traduzcan en mayores ganancias de peso en aves, porcinos, etc. Un diseño experimental que permita evaluar a nuevos balanceados producidos por casas comerciales, frente a un testigo, deja oportunidad al investigador para seleccionar el mejor y recomendarlo al público.

1.3.3 Comercialización. A las entidades encargadas de controlar la distribución de productos agrícolas, interesa conocer zonas de cultivo, época de siembra y cosecha, costos de producción, canales por los que se distribuye el producto. Una encuesta bien diseñada y un muestreo adecuado, pueden dar resultados económicos satisfactorios tanto para el agricultor como para el consumidor.

1.3.4 Industria. Los fabricantes de piezas de repuesto, necesitan conocer el porcentaje de unidades defectuosas que produce una máquina. En esta forma será posible determinar si la máquina se halla funcionando en forma eficiente y económica, o si debe hacerse algún cambio para conseguir tal propósito. Por ejemplo, suponiendo que los diámetros de arandelas de presión, permiten una tolerancia establecida, usando una prueba de **z** es posible determinar el porcentaje de arandelas defectuosas. Así mismo, usando la **distribución normal,** se puede determinar la probabilidad de que, en una muestra de 10 tuercas sacadas al azar, dos de ellas sean defectuosas.

1.3.5 Medicina. La prensa informa diariamente sobre nuevos productos para controlar o prevenir tal enfermedad; o sobre el resultado de varios métodos para curar el cáncer. Para conseguir esos nuevos productos o métodos, el investigador médico puede usar conejos o ratas, a los que se inocula el organismo causante de la enfermedad, quizás en varias dosis, frente a un testigo. Después de tomar cuidadosamente los datos y del análisis estadístico respectivo, es posible multiplicar nuevas vacunas, sueros, etc.

En el mismo campo médico, por medio de un simple muestreo, el profesional obtiene inferencias sobre la sangre de sus pacientes, a través del examen de una sola gota.

Así mismo, se puede calcular la probabilidad de que el tercer hijo de una familia que padece de hemofilia, herede este carácter. O se puede establecer quién es el padre de un niño, cuya madre tiene grupo sanguíneo O.

1.3.6 Agronomía. El ensayo más sencillo sería la prueba de adaptación de algunas variedades de trigo, en cierta localidad. Si las condiciones de fertilidad, riego, cuidados culturales, etc., se las mantiene constantes, el único factor en estudio estará formado por las variedades y así se podrá establecer cuál es la de mayor rendimiento. Si se quiere obtener mayor información, es posible usar dosis de fertilizante y/o sembrar las variedades en varias localidades, con lo que se amplía el alcance de las inferencias.

1.3.7 Información general. Un muestreo bien diseñado puede predecir el resultado de las próximas elecciones, o informar sobre las preferencias del consumidor.

1.4 Algunos datos históricos

La Estadística moderna es considerada como una herramienta de trabajo para el investigador. Producto del siglo XX, para el biólogo cobró especial auge en 1925, cuando Sir Ronald Fisher publicó sus Métodos Estadísticos para Investigadores.

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

El término Estadística es muy antiguo y quiere decir aritmética del Estado. Con el objeto de conseguir que todos los súbditos paguen impuestos, o contribuciones para librar guerras, el gobernante antiguo pedía que aquellos se acerquen al "estadístico" más cercano a cumplir con sus obligaciones.

Posteriormente encontramos una aplicación empírica del cálculo de probabilidades en el seguro de barcos de que disponían los flamencos en el siglo XIV. El sistema pudo haber sido especulación o juego de azar, pero ha devenido modernamente en el lucrativo negocio de seguros.

Los juegos de azar condujeron a la teoría de probabilidades, originada por Pascal y Fermat, a mediados del siglo XVII.

Charles Darwin, en el siglo XIX, basó su trabajo biológico en métodos estadísticos. Sin embargo, los problemas planteados por su Teoría de la Evolución, hicieron evidente la necesidad de usar métodos más refinados y fue Karl Pearson (1936), inicialmente un físico-matemático, quien aplicó sus conocimientos para mejorar la toma de datos y las evaluaciones respectivas.

Un alumno de Pearson, WilliamS. Gosset –científico y cervecero–, desarrolló en forma empírica lo que hoy se conoce como la distribución de **"t de Student"**, que fue el seudónimo que empleaba Gosset en sus publicaciones en Biometrika.

Sir Ronald Fisher ha hecho numerosas y valiosas contribuciones en el campo de la Estadística, particularmente con el análisis de variancia, que es el procedimiento más usado por investigadores, prácticamente en todo tipo de experimento. Igual es el caso de los análisis de covariancia y regresión.

1.5 La Estadística y el Método Científico

Es conocido que los investigadores usan métodos científicos dentro de su trabajo diario. Resulta difícil definir lo que es el método científico, puesto que, en investigación, puede usarse cualquier sistema ideado por una persona, para llegar a obtener un resultado. Sin embargo, la mayoría de los procedimientos tienen algunas características en común:

a) Una revisión de hechos, teorías, con miras a

b) Formular una hipótesis lógica, sujeta a ser probada por métodos experimentales.

c) Diseño y conducción del experimento.

d) Evaluación objetiva de la hipótesis, en base a los resultados experimentales.

e) Publicación o divulgación de los resultados.

La ciencia es una rama de estudio que tiene que ver con la observación y clasificación de hechos. El investigador debe estar en capacidad de observar un evento, como resultado de un plan o diseño. Este es el experimento, la sustancia del método científico.

La evaluación objetiva de una hipótesis, presenta un problema: no es posible observar todos los eventos posibles en la naturaleza y, por cuanto las leyes exactas de causa y efecto, son generalmente desconocidas, va a haber variación entre las observaciones. El científico debe siempre partir de casos particulares, para llegar a generalizaciones **(proceso inductivo).** Por cuanto es imposible, o poco práctico, estudiar todo el **universo** (población), por razones de costo, tiempo y esfuerzo, el investigador debe contentarse con estudiar la **muestra** y, a partir de ella, obtener conclusiones sobre la población (inferencia estadística). El proceso inverso, como nos enseña la lógica, es la deducción, por la que, de una norma o ley generales, se llega a casos particulares.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 2

VARIACIÓN – VARIABLES

Los seres vivos –plantas, animales y el hombre– difieren, aun dentro de la misma especie, en muchos caracteres. Puede generalizarse que, inclusive en el caso de gemelos homocigóticos, no existen dos seres vivos idénticos. Trabajos experimentales realizados en Australia, demostraron que dos terneras gemelas homocigóticas, llevadas la una a un hato en el que el manejo y alimentación fueron adecuados; y la otra a un hato en el cual estos aspectos fueron deficientes, se comportaron de acuerdo con el ambiente en que crecieron, siendo este último factor, responsable en 85% de la precocidad y ganancia de peso obtenidos.

Si observamos un potrero de raigrás –cultivo de fecundación cruzada– a primera vista puede impresionarnos la uniformidad de las plantas: altura y desarrollo de los tallos; forma, tamaño y disposición de las hojas; forma y tamaño de la espiga, etc. Pero, si comparamos dos plantas de raigrás, arrancadas al azar, en detalle y se mide cuidadosamente caracteres cuantitativos de diferentes partes de la planta, encontraremos que los individuos difieren en varios aspectos. Habrá notables diferencias en vigor de plántulas, días a floración, maduración, producción de forraje, producción de semilla, etc. De este tipo de variación es justamente de lo que se sirve el fitomejorador para la selección y creación de nuevas variedades.

2.1 Variación ambiental y variación hereditaria

Las variaciones dentro de una especie se deben a: 1) variaciones debidas al ambiente en que desarrolla una especie y 2) variaciones debidas a la herencia.

2.1.1 La variación provocada por causas ambientales, puede probarse sembrando plantas de igual **genotipo** en diferentes localidades. Esta situación da lugar a la formación de agrotipos y ecotipos como en el caso de la alfalfa "Nacional". La semilla de alfalfa, introducida originalmente del Perú y sembrada en diferentes regiones de la Sierra ecuatoriana, ha dado origen a ecotipos cuyos individuos varían en caracteres morfológicos y constitución fisiológica que resultan en mayor o menor altura de planta, resistencia al frío, a la altura, a las enfermedades, etc.

Dos plantas de trigo de la misma variedad, tendrán desarrollo y rendimiento diferentes, si la una es atacada por el "polvillo", mientras a la otra se protege de la infección.

Estas variaciones provienen del ambiente en el que desarrollaron las plantas y pueden o no pueden presentarse en las respectivas progenies.

2.1.2 Las variaciones hereditarias son resultado de la diferente constitución genética (genotipo), de las plantas o animales. Generalmente, estas variaciones pueden probarse sembrando diferentes especies o variedades, en condiciones ambientales similares.

Las variaciones hereditarias pueden ser fácilmente observadas: plantas de trigo enanas o altas; diferencia en el color de la planta, flor o semilla; cantidad de pubescencia en tallos y hojas; presencia o ausencia de aristas, etc. Existen, desde luego, caracteres más difíciles de identificar como vigor de plántula, capacidad de macollamiento, resistencia a enfermedades, etc.

Por cuanto estas variaciones son heredables, ellas se manifiestan en la progenie, aun cuando la intensidad con que se expresen, varíe con el ambiente. Todo lo expuesto sugiere que genotipo y ambiente no son dos entidades in-

dependientes. Más bien, se sabe que el individuo es el resultado de los dos factores, más la **interacción** correspondiente.

2.2 Variables

Asertos como "Juan es moreno" o "él pesa sobre las 150 libras", son comunes e informativas. Se refieren a características que no son constantes y que varían de un individuo a otro.

Variable es, entonces, la cantidad o carácter que pueden ser medidos y se hallan, en consecuencia, sujetos a variación: edad, peso, altura, temperatura, etc. Las variables pueden ser cualitativas o cuantitativas.

2.2.1 Variable cualitativa. Es aquella en que cada individuo pertenece a una de varias categorías mutuamente excluyentes, generalmente no numéricas: color, sabor, nacionalidad. Así podríamos decir que el color blanco, excluye automáticamente al negro; el sabor agrio, excluye al dulce. Es claro que el color puede calificarse de acuerdo a una escala relativa numérica, en cuyo caso podrían analizarse los datos como si fueran cuantitativos.

2.2.2 Variable cuantitativa. Se refiere a datos numéricos: contajes, medidas, pesos, etc.

Variable cuantitativa discreta. Es aquella en que los valores son clasificados en categorías específicas: número de plantas por surco, número de sépalos en una flor.

Variable cuantitativa continua. Es aquella en que es posible tener todos los valores dentro de una escala o rango, en forma continua: altura, peso, temperatura.

2.3 Observaciones, hechos

Son elementos o atributos de información: altura de una planta o persona; peso de novillos.

Las observaciones son la materia prima con que trabaja el investigador. Para poder analizarlas, esas observaciones tienen que estar en forma de números. En Agronomía, las cifras pueden ser rendimientos por parcela o por hectárea; en investigación médica o veterinaria, tiempo de duración de una vacuna, número de pústulas de acuerdo a varias dosis de un producto dado; en industria, número de piezas defectuosas en lotes producidos en serie; etc., etc.

Los números constituyen los "datos" y su característica común es la variación o variabilidad.

2.4 Población y Muestra.

La primera preocupación del investigador frente a un grupo de datos, es saber si ellos constituyen la totalidad de individuos u observaciones, o si forman parte de un grupo mayor. Aun cuando la diferencia entre los dos casos parezca trivial, es en realidad de gran importancia.

Población. La población o universo, consiste de todos los valores posibles dentro de una variable: todos los estudiantes de un curso; el número de caras en 100 lanzamientos de una moneda; los pesos al nacer de una camada de cerdos.

Población finita. Lo que se puede contar: el número de estudiantes del curso.

Población infinita. Las bacterias Azotobacter en el suelo. Asociada con procesos repetitivos como lanzar una moneda.

Muestra. Es una parte de la población; es una selección de individuos tomados del universo o población.

La muestra debe ser representativa si vamos a llegar a una inferencia válida sobre la población. Tal sería el caso de la toma de muestras de suelos, encuestas políticas, encuestas sobre preferencias del público, etc.

Muestra al azar. Es una porción del todo, en la que cada individuo tiene igual oportunidad de ser incluido (verdadera muestra al azar). No es posible tomar

una muestra al azar, si no es por métodos mecánicos. La influencia subjetiva de una persona, impediría que esa muestra sea al azar.

Los mecanismos para extraer una muestra al azar se refieren al uso de papeles numerados (caso de la lotería) o de tablas de números al azar. Si se va a nombrar un jurado imparcial, dentro de una clase, se podría seleccionar a los alumnos de acuerdo con los números impares de la lista. El caso típico está dado por "las experiencias de papas": ni al agricultor ni al comprador les interesa efectuar la compra-venta "al ojo". Si el primero estima el rendimiento de una hectárea en 400 quintales, por ejemplo, y el comprador obtiene 500, aquél se habría perjudicado; en caso contrario, si el cálculo es de 400 quintales y el comprador cosecha 300, éste habría perdido.

Con el objeto de que la compra-venta sea equitativa para las dos partes, se procede a realizar las "experiencias", que no son sino un muestreo al azar, que consiste en lo siguiente: se cosecha de tres a cinco surcos separadamente, dentro de una "tabla", y el producto se amontona al extremo de los surcos y se pesa; se obtiene el promedio de los tres surcos y se cuenta el número de surcos de la "tabla". Se multiplica luego el peso promedio por surco por el número de surcos y así se puede estimar el rendimiento por hectárea.

Otro sistema más preciso, aun cuando requiere de más tiempo, se refiere a "cavar" de tres a cinco matas por surco, de un total de tres a cinco surcos de la "tabla". Una vez obtenido el promedio por mata, se cuenta el número de matas por surco y el número de surcos, con lo que se puede llegar a estimar el rendimiento por hectárea. De acuerdo con la superficie ocupada por la sementera, se optará por uno u otro sistema de muestreo.

La Tabla N° 1 del Apéndice presenta un grupo de números al azar.

Estadística. Escrita con minúscula, es una cantidad que se refiere o describe a la muestra.

Parámetro. Es una cantidad que se refiere a la población.

2.5 Distribuciones

Los valores de una variable sirven para describir o clasificar individuos o para distinguirlos unos de otros. Muchos de nosotros hacemos algo más que describir o clasificar datos, porque tenemos cierta idea sobre la frecuencia relativa de los valores de una variable. No daríamos crédito si nos hablan de personas de tres metros de estatura; un niño de ocho libras de peso al nacimiento, sería algo común, excepto para sus padres. Generalmente asociamos una medida cualquiera con el valor de una variable, una medida de que tan común o raro es ese valor. En Estadística se dice que la variable tiene una distribución, una distribución de frecuencias o de probabilidades. Para una moneda, por ejemplo, la probabilidad de obtener sello o cara es de 0.5.

En la **distribución binomial,** por ejemplo, los elementos pueden ser clasificados en una de dos clases o categorías: cara o sello; macho o hembra; defectuoso o no defectuoso; inoculado o no inoculado, etc. Generalmente se considera a las dos clases como **éxito y fracaso,** tal como se verá en la Sec. 7.1.

Otra distribución a la que se adaptan datos correspondientes a variables cuantitativas discretas, es la **distribución de Poisson.** Este resulta un modelo apropiado para la distribución de un número de elementos por unidad de tiempo o espacio, cuando el número promedio de elementos por unidad es relativamente pequeño. Tal sería el caso que resulta del contaje de huevos de insectos o larvas por determinada área foliar en papa, por ejemplo: el número de focos que deben ser reemplazados en los semáforos de la ciudad, cada mes; el número de radios defectuosos producidos en una semana, etc., etc. Como se verá más adelante, este tipo de datos no se distribuyen en forma "normal", es decir, no se agrupan bajo la campana simétrica que caracteriza a la **distribución normal,** la misma que se describe ligeramente a continuación y en forma amplia en la Sec. 7.2.

Si registramos y tabulamos datos como el cociente intelectual o pesos de un número de individuos, vamos a conformar cuadros de frecuencia tal como se verá en la Sec. 5.1. Es decir, se determina las clases o categorías (el cociente intelectual o el peso en libras) y seguidamente la frecuencia o número de

personas dentro de cada clase. A partir del cuadro de frecuencias, es posible ubicar los datos dentro de un sistema de coordenadas, en el que las clases se localizan en el eje de las abscisas (X) y la frecuencia en el de las ordenadas (Y). Para la Fig. 2.1, podemos apreciar que la media o promedio del cociente intelectual de un número dado de personas, es 100, registrándose el mayor número de ellas alrededor de este valor; a los extremos de la curva (los cocientes más altos y más bajos) y en menor número, se encuentran los individuos de mayor y menor cociente intelectual. En igual forma, en la Fig. 2.2 se ve que el peso promedio de 30 personas es de 150 libras; aquellas de peso más alto y más bajo se ubican, en menor número en las colas derecha e izquierda, respectivamente, de la distribución, encontrándose el mayor número alrededor del valor promedio.

Cuando se grafican datos como altura, edad, peso de personas, temperatura y muchas otras variables de naturaleza continua, veremos que los datos dan una curva parecida a las Figs. 2.1 y 2.2.

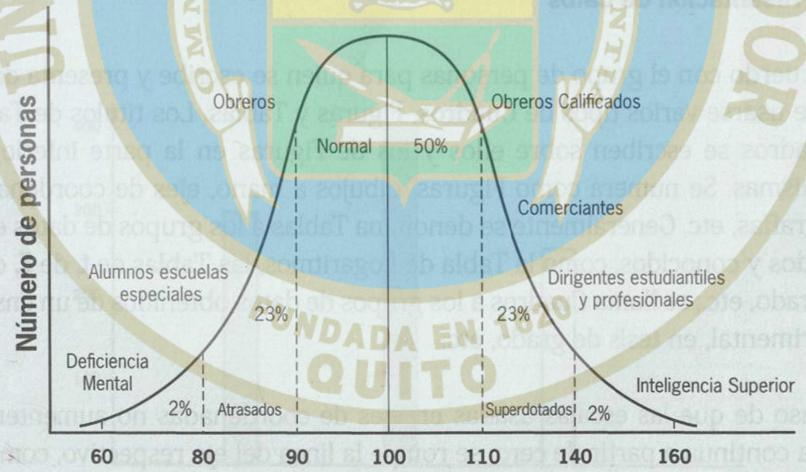


Fig. 2.1 Distribución de frecuencias del coeficiente intelectual.

UNIVERSIDAD CENTRAL DEL ECUADOR
SIB - UCE

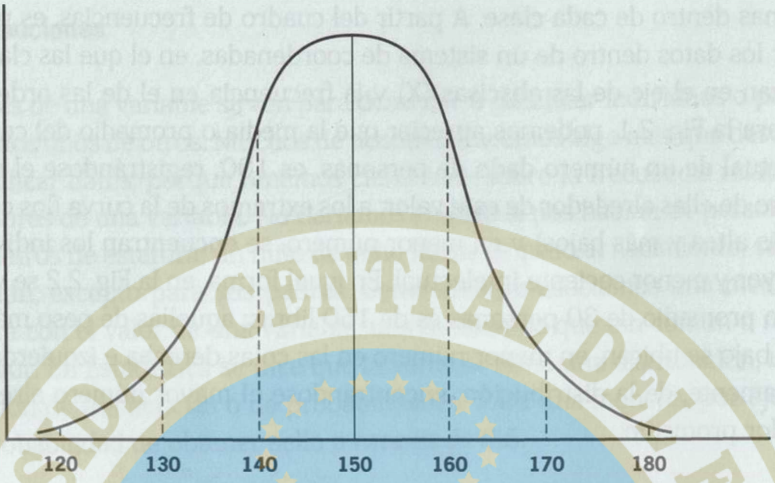


Fig. 2.2 Peso en libras de un grupo de personas.

2.6 Presentación de datos

De acuerdo con el grupo de personas para quien se escribe y presenta datos, puede usarse varios tipos de Cuadros, Figuras y Tablas. Los títulos de Tablas y Cuadros se escriben sobre ellos y los de Figuras en la parte inferior de las mismas. Se numera como Figuras, dibujos a mano, ejes de coordenadas, fotografías, etc. Generalmente se denomina Tablas a los grupos de datos establecidos y conocidos, como la Tabla de Logaritmos, las Tablas de **t**, de **F,** de **Ji** cuadrado, etc.; se llama Cuadros a los grupos de datos obtenidos de un ensayo experimental, en tesis de grado, etc.

En caso de que las escalas usadas en ejes de coordenadas no aumenten en forma continua a partir de cero, se rompe la línea del eje respectivo, como se indica en la Fig. 2.6.

Seguidamente, se presenta algunos tipos de Figuras, las que van de la más simple a la más complicada, debiendo usarse cualesquiera de ellas, de acuerdo con el grupo de lectores para quienes va dirigida la publicación.

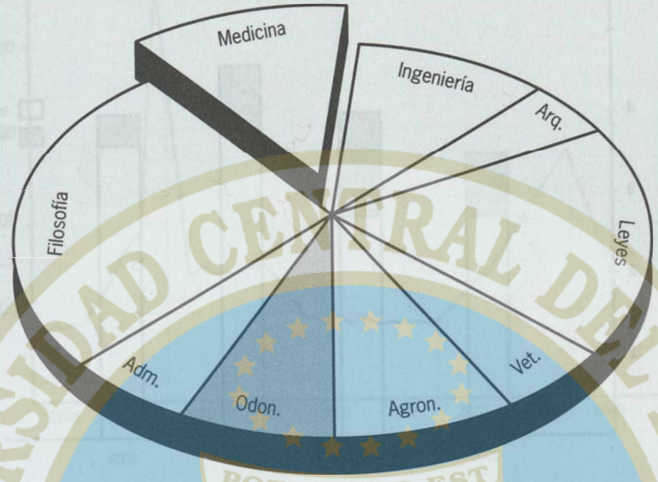


Fig. 2.3 Distribución de estudiantes por Facultades.

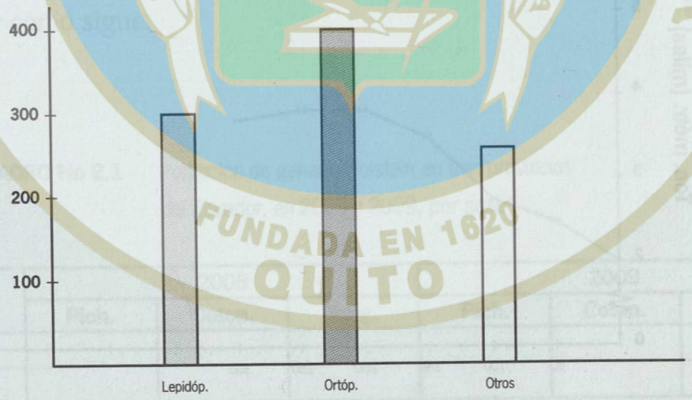


Fig. 2.4 Número de insectos atrapados.

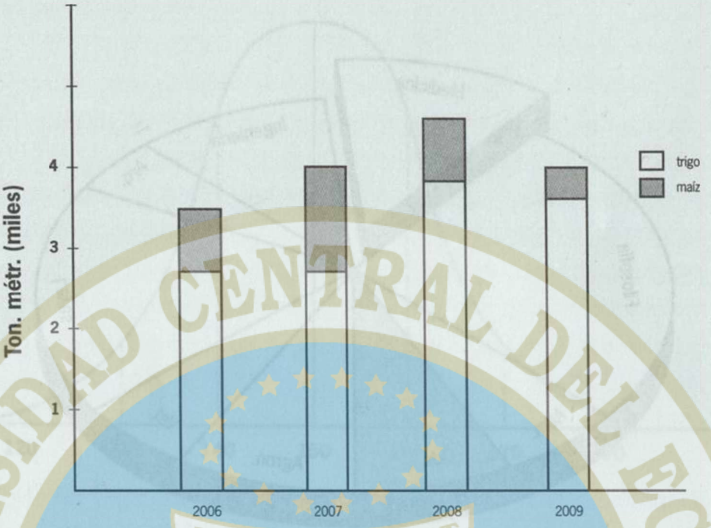


Fig. 2.5 Sistema de "barras" con dos variables.

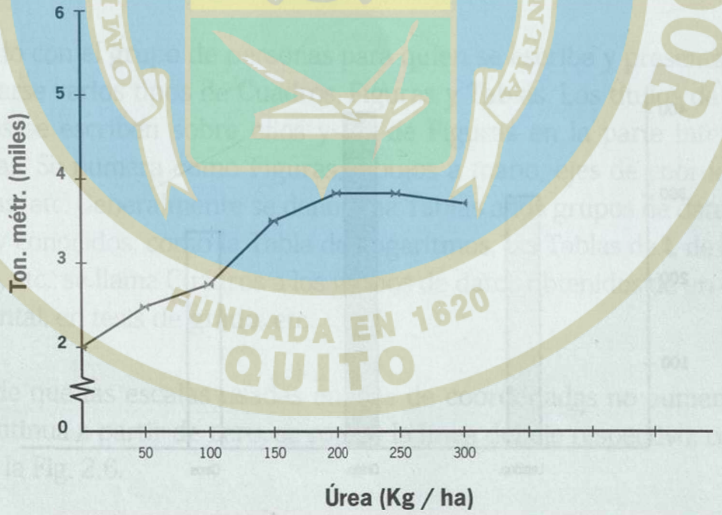


Fig. 2.6 Sistema de línea continua.

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO

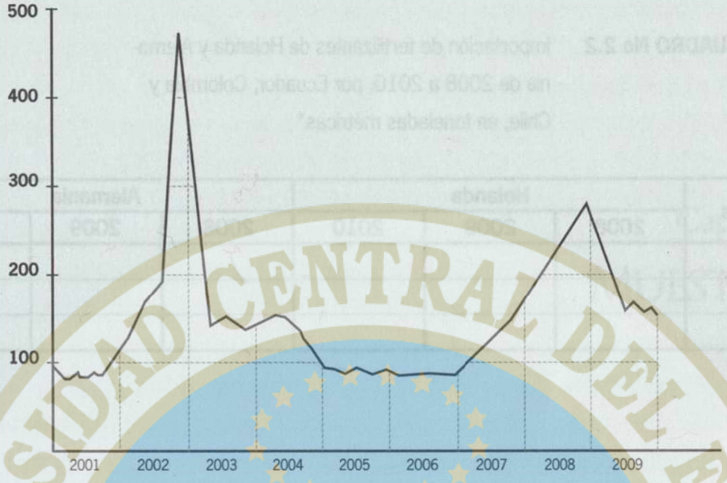


Fig. 2.7 Precio mundial del azúcar en libras esterlinas por tonelada.
Fuente; Intern. Agric. Develop. Vol. I (9), octubre, 2008.

La forma de presentar cifras en cuadros (bloques compactos de números), podrá ser como sigue:

CUADRO No 2.1 Población de ganado Holstein en tres provincias del Ecuador, en 2008 y 2009, por sexo*

Sexo	2008			2009		
	Pich.	Cotop.	Tung.	Pich.	Cotop.	Tung.
Machos						
Hembras						

**** Fuente:***

CUADRO No 2.2 Importación de fertilizantes de Holanda y Alemania de 2008 a 2010, por Ecuador, Colombia y Chile, en toneladas métricas*

Países	Holanda			Alemania		
	2008	2009	2010	2008	2009	2010
Ecuador						
Colombia						
Chile						

Fuente:

CUADRO No 2.3 Matrícula estudiantil en la Universidad Central, por sexo, de 2006 a 2009*

	Agronomía		Medicina		Ingeniería		Odontología		Economía	
Año	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.
2006										
2007										
2008										
2009										

Fuente:

Capítulo 3

MUESTREO

3.1 Poblaciones. De acuerdo con la definición común, la población está formada por la totalidad de habitantes dentro de un área geográfica determinada. Desde el punto de vista estadístico, una **población** o **universo**, está constituida por la totalidad de elementos que poseen una o más características en común. Así hablamos de la población de radios marca "X", producida por una fábrica durante un año; de la población de animales vacunos que posee una hacienda; de la población de plantas de abacá dentro de un lote; del número de estudiantes del curso, etc., etc.

La definición estadística de población, no incluye solamente a poblaciones de individuos u objetos, sino que considera poblaciones de valores numéricos obtenidos al medir una o más características de personas u objetos como la altura o peso de individuos, el diámetro de tuercas producidas por una fábrica; o pueden consistir de todos los valores posibles que se puede conseguir cuando se hacen observaciones repetidas, como sería el caso de registrar el número de caras o sellos al lanzar una moneda mil veces.

El principal objetivo de la inferencia estadística es sacar conclusiones a partir de la muestra, sobre poblaciones. En consecuencia, es de fundamental importancia especificar o definir la población con la que estamos trabajando.

Muchas veces, la población para la que se han obtenido ciertas inferencias, difiere en algún aspecto importante de la población que produjo los datos en los que se basó la inferencia.

Siempre existe riesgo cuando se generalizan conclusiones para una población amplia, partiendo de una población limitada o que no es representativa de aquella. Cuando se realizan pruebas de adaptación de una nueva variedad en varias localidades o haciendas, dentro de un cantón o provincia, tendremos seguridad de recomendarla –si los resultados fueron satisfactorios– para las localidades o haciendas en las que se hizo el ensayo. Sería peligroso generalizar las conclusiones y afirmar que la nueva variedad puede sembrarse indistintamente en todo el cantón o provincia. Esto es así por cuanto conocemos la variabilidad de suelos, clima, prácticas culturales que existen dentro de un mismo cantón o provincia.

Para estudios de tipo sociológico, industrial, económico, etc., nos interesa obtener una muestra representativa y suficiente –que refleje lo más exactamente que sea posible las características de la población para la que se sacará conclusiones– si vamos a tener confianza en las predicciones. Cuando se trata de establecer, por ejemplo, la edad promedio de los estudiantes de la Universidad, o los ingresos de sus padres para adoptar una decisión de carácter económico, debemos realizar un muestreo estratificado formando clases o grupos con características similares dentro de cada uno.

3.2 Muestras. Cuando queremos obtener datos sobre una población cualquiera, sería ideal si pudiéramos analizar cada elemento de la población. En la mayoría de los casos, tal procedimiento sería descartado por razones de tiempo y costo, principalmente. Por otro lado, no se justificaría debido a que, en la práctica, pueden obtenerse resultados precisos, en forma rápida y con menor costo, estudiando solo una parte de la población. Esta parte de la población constituye la muestra. En la Sec. 3.5 veremos los diferentes tipos de muestreos que pueden utilizarse con diversas poblaciones.

3.3 Estimación. En general, debido a que difícilmente se puede estudiar y analizar una población íntegra, desconocemos el valor real de los **parámetros** de la distribución teórica de la que suponemos que los datos tomados por noso-

tros constituyen una muestra. Aspiramos entonces, a que la muestra tomada sea un **estimado insesgado** (no viciado, confiable) de los valores poblacionales o paramétricos. Por ejemplo, cuando se cumple con lo que dice el **teorema del límite central** (Sec. 3.4), la media de la muestra $\bar{X}$ es un estimador de la media poblacional μ; la desviación típica de la muestra s, es un estimador de la desviación típica de la población σ.

3.4 Teorema del límite central. Uno de los teoremas más importantes en lo que se refiere a distribuciones de muestreo, para series de variables tomadas al azar, independientes, idénticamente distribuidas, es el teorema del límite central. Se dice que un juego de observaciones al azar es idénticamente distribuida cuando todos los miembros o elementos del juego de observaciones han sido seleccionados de la misma distribución (población).

El teorema, en su forma más simple, dice: **si se seleccionan muestras al azar de una población determinada, conforme aumenta el tamaño de la muestra, la distribución se aproxima a la distribución normal.** (Ver Secs. 2.5 y 7.2). Es decir que podemos hacer uso de la teoría sobre distribuciones normales y de distribuciones de muestreo derivadas de poblaciones normales, con el objeto de obtener inferencias acerca de la población involucrada, sin tomar en cuenta la forma o tipo de ésta, con tal de que el tamaño de la muestra sea suficientemente grande. El tamaño de la muestra que puede considerarse suficiente en una situación dada, depende del grado en que la población involucrada se desvía de la normalidad.

3.5 Tipos de muestreo. Sabemos que, a partir de la muestra, se obtienen generalizaciones sobre la población. La exactitud de aquellas depende del cuidado con que se diseña y ejecuta el análisis muestral.

Trabajar con la muestra no solo reduce el tiempo sino el costo de obtener la información. Por ser imposible o impráctico trabajar con la población total, se han creado diseños muestrales, algunos de los que se trata seguidamente. El procedimiento o diseño del muestreo que ha de utilizarse depende:

1) **Del tipo de población de que se trata;**

2) **De la información deseada; y,**

3) **De los fondos y tiempo disponibles.**

El diseño de muestreo es un plan que especifica cómo se seleccionará la muestra que se ha de extraer de una población dada.

En general, las poblaciones pueden ser finitas o infinitas. Las casas de un barrio, los agricultores de una parroquia, el número de cabezas de ganado de una hacienda, son ejemplos de poblaciones finitas.

Poblaciones infinitas se asocian con algún proceso repetitivo como el de lanzar una moneda, que puede dar caras o sellos, en forma indefinida; también constituyen poblaciones infinitas, los individuos u observaciones cuyo contaje o mensuración sería difícil o imposible de establecer: las bacterias **Azotobacter** del suelo.

3.5.1 Muestreo aleatorio simple. Si bien el término aleatorio (al azar) puede implicar que las observaciones son seleccionadas de manera fortuita o casual, las verdaderas muestras aleatorias requieren de un cuidadoso diseño y ejecución a fin de asegurar la independencia de dichas observaciones. Si la muestra no es aleatoria, no se puede tener confianza sobre las conclusiones derivadas de ella.

La selección de muestras aleatorias de una población finita, puede ser una tarea larga y tediosa. Supongamos que se desea estimar la edad promedio de los estudiantes de la Universidad Central; una persona podría instalarse en la puerta principal y preguntar la edad de cada 100 estudiantes que entren. Esta no sería una muestra aleatoria porque la probabilidad de usar esa puerta es mayor para cierto grupo de estudiantes que para otros.

También se podría telefonear, una determinada noche, al 10% de alumnos que habitan en Residencia Universitaria; esta tampoco sería una muestra aleatoria, porque es posible que estudiantes de ciertos cursos tengan mayor acceso a la Residencia o porque determinados grupos de estudiantes no permanezcan en ella las noches.

Una manera de lograr una muestra aleatoria, sería conseguir una lista completa de todos los estudiantes matriculados, asignar un número a cada uno, anotarlos en pedazos de papel y, luego de colocarlos en una caja, seleccionar los números, reponiéndolos cada vez y mezclando los papeles antes de cada extracción. Esta sería una verdadera muestra al azar, para la que estaríamos trabajando con todo el universo o población. El proceso, claro está, duraría varios días y sería impráctico y tedioso dado el alto número de estudiantes que conforman la población universitaria. Este tipo de muestreo sería utilizable cuando se trata de poblaciones pequeñas, bien definidas; las rifas o sorteos, como el juego de lotería de Guayaquil emplean muestreos aleatorios simples, en los que la población está constituida por los números cero a nueve, para cada una de las cinco fichas que dan los diferentes premios.

3.5.2 Muestreo probabilístico. En este caso, lo primero es fijar el tamaño de la muestra, a fin de obtener un determinado grado de precisión a la estimación de un parámetro.

Supongamos que el rector de una escuela quiere determinar el coeficiente intelectual de los estudiantes, a partir de una muestra aleatoria. Al efecto, decide usar la media aritmética y una muestra al azar formada por los datos contenidos en 50 tarjetas de archivo; el total de alumnos es de 1.000. En este caso, el rector estaría trabajando con 5% del estudiantado. Por cuanto sería impráctico emplear pedazos de papel en los que se halle el número de cada uno de los alumnos, de acuerdo con el listado de archivo, resulta más fácil recurrir a una tabla de números randomizados, como la Tabla No 1 del Apéndice. Dicha Tabla está formada por dígitos obtenidos al azar, mecánicamente, de tal modo, que cuando seleccionamos una serie de aquellos, podemos tener confianza de que al hacerlo estamos obteniendo una verdadera muestra al azar, sin que intervenga el criterio o preferencia de la persona al realizar el sorteo.

Supongamos que se elige como punto de partida la primera columna de la Tabla respectiva; por cuanto el listado de archivo irá de 001 a 1.000 (la población total con la que estamos trabajando), vamos a escoger 50 series de tres números cada una. El primer número que aparece en la primera columna es el 039; por tanto se incluirá en la muestra la tarjeta No 39; la segunda será

la No 385, luego la 175 y así sucesivamente, hasta completar las 50 tarjetas. Este tipo de muestras fue muy simple: la población era relativamente pequeña y estaba bien definida y fue fácil asegurarse de que cada elemento de la población tuviera igual oportunidad de ser escogido. Sin embargo, no siempre es posible obtener una lista completa de todos los elementos de una población. Por ejemplo, no existe una guía actualizada de todos los productores de maíz en el Ecuador y, si quisiéramos determinar costos de producción o el grado de mecanización que aquellos usan en sus propiedades, tendríamos que recurrir a un tipo de muestreo diferente que el probabilístico.

3.5.3 Muestreo estratificado. El muestreo proporcional de una población estratificada da una estimación más precisa que la media poblacional global. La idea fundamental de este tipo de muestreo es la de dividir a la población total en subpoblaciones, con la condición de que existan diferencias entre las medias de las subpoblaciones y si, dentro de éstas, existe relativa homogeneidad. Supongamos que el directorio de una compañía quiere estimar el monto de ingresos o calcular el valor de seguros de vida de sus empleados. Se considera que los ejecutivos y empleados administrativos tienen mayores ingresos que los obreros de planta. Por tanto el directorio decide clasificarlos en la siguiente forma:

Estrato	Clasificación	Número	Proporción por estrato
1	Obreros	800	0,80
2	Administrativos	150	0,15
3	Ejecutivos	50	0,05
	Total	1.000	1.00

Si seleccionamos una muestra de 80 empleados (8%), por muestreo proporcional, obtendríamos:

80 X 0,8 = 64 obreros
80 X 0,15 = 12 administrativos
80 X 0,05 = 4 ejecutivos
80 empleados, los que serán seleccionados al azar, dentro de cada estrato o categoría.

En esta forma daríamos cumplimiento a lo establecido por este tipo de muestreo, puesto que a la población total, se habría dividido en tres subpoblaciones cuyos promedios de ingresos son diferentes unos de otros, existiendo relativa homogeneidad dentro de cada estrato o subpoblación.

3.5.4 Muestreo sistemático. Para este tipo de muestreo, se incluye cada k-ésimo elemento de una población ordenada, por ejemplo, en forma alfabética. El punto de partida se elige al azar entre los primeros k elementos y el muestreo se continúa, de acuerdo con el intervalo decidido, hasta completar los n objetos u observaciones. Se podría, por ejemplo, hacer una encuesta telefónica, llamando a cada quincuagésimo nombre.

El principal inconveniente radica en que las muestras sistemáticas no siempre son aleatorias. Si la población está ordenada en forma sistemática con respecto a la característica de interés, es posible que se incluyan más elementos de una clase que de otra. Por ejemplo, si se trata de averiguar el ingreso de empleados administrativos y obreros de una fábrica de gran tamaño, posiblemente van a aparecer en la lista de pagos, más obreros que empleados administrativos, sesgando el ingreso medio general.

En algunos casos, puede modificarse el sistema cambiando ocasionalmente el punto de partida y el intervalo entre muestra y muestra.

3.5.5 Muestreo por áreas. O por conglomerados, en el que las subdivisiones pueden ser barrios, distritos, parroquias, cantones, etc. Se usa este tipo de muestreo, especialmente en encuestas sociológicas, en las que se asigna al encuestador una porción (conglomerado), de los distintos tipos de personas que han de ser entrevistadas (grupos sociales, religiosos, profesionales, económicos, etc.). El encuestador hace la selección aleatoria dentro de su conglomerado.

El muestreo por áreas elimina el tener que hacer un listado completo de los elementos de una población finita. Es adecuado para encuestas a nivel nacional y se puede aumentar el tamaño de la muestra sin mayor problema.

UNIVERSIDAD CENTRAL DEL ECUADOR
SIB - UCE

3.5.6 Muestreo dirigido. En este tipo de muestreo, el criterio desempeña un papel importante en la selección. Puede ser de utilidad si el investigador está bien familiarizado con la población y puede elegir elementos representativos para la integración de la muestra. El analista debe aplicar su propio criterio para decidir si una muestra es "buena" o "mala". La muestra debe ser representativa si vamos a llegar a obtener una inferencia válida sobre la población. Tal sería el caso de la toma de muestras de suelo, encuestas políticas, encuestas sobre preferencias del público, etc. Para la primera situación, el profesional divide el campo de acuerdo con la homogeneidad aparente de la vegetación y la topografía para, dentro de cada sector o lote relativamente uniforme, tomar al azar las muestras que serán analizadas en el laboratorio. Las recomendaciones sobre el uso de fertilizantes, por ejemplo, se harán en forma separada para los lotes de los que se extrajo las muestras. Si este trabajo hubiera sido hecho en forma indiscriminada, para lotes obviamente diferentes, no tendría valor la recomendación del laboratorio.

En el caso de encuestas de tipo político sobre preferencias del público, debe precederse de manera similar. Diferentes grupos de personas (aspectos raciales, religiosos, económicos, educativos, etc.), van a tener diversas preferencias y, en consecuencia, conviene separarlas en grupos a fin de obtener inferencias más precisas sobre las poblaciones involucradas. Como vemos, este tipo de muéstreos son parte dirigidos y parte aleatorios.

3.5.7 Muestreo por cuotas. Es un tipo de muestreo dirigido en el que se fijan cuotas de acuerdo con ciertos criterios, para, dentro de ellas realizar la selección en base al criterio personal del entrevistador. Supongamos que se va a investigar sobre el uso de un nuevo detergente y se decide entrevistar a 500 personas en esta forma: 50% amas de casa; 20% ganaderos y 30% lavanderías comerciales. Ajustándose a las cuotas señaladas, el entrevistador puede seleccionar libremente dentro de cada grupo. Ahora bien, el entrevistador puede tener cierta preferencia por ciertos sectores de la ciudad o del campo; o de trabajar ciertas horas del día o ciertos días de la semana, en los que puede o no puede encontrar a la persona indicada. Para resolver este problema, la persona podría hacer la encuesta a la mitad de las amas de casa, ganaderos y lavanderías pasadas las seis de la tarde o durante los fines de semana.

Capítulo 4

MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN

4.1 Simbología matemática. En Estadística se usan algunos símbolos para representar observaciones. La utilización de dichos símbolos, es lo que constituye la **notación** de una expresión matemática.

Como se ha indicado en la Sec. 2.2, la base del trabajo en Estadística está formada por variables, las que se designan generalmente con las letras mayúsculas **X, Y, Z.** Las primeras letras del alfabeto, en minúsculas, sirven para designar constantes.

Uno de los signos de mayor uso en Estadística, es la letra griega mayúscula Σ **(Sigma),** que significa "la suma de" o "sumatoria de".

Una observación cualquiera se representa por X_i y, si existen varias observaciones, ellas serán $X_1, X_2, X_3, \ldots, X_n$. Es decir que la letra i en X_i, llamada **subscrito** o **índice**, representa cualesquiera de los valores 1, 2, 3, ... , n. Si tenemos los números 6, 8, 4, 12, la suma de ellos está dada por:

$$\sum_{i=1}^{4} X_i = X_1 + X_2 + X_3 + X_4$$

$$= 6 + 8 + 4 + 12 = 30$$

En general, $\sum_{i=1}^{n} X_i = X_1 + X_2 + X_3 + \ldots + X_n$, que dice la suma $i=1$ de las X_i, desde $i = 1$ hasta n (la enésima o última observación), es igual a la suma total de las observaciones. Algunos autores representan a ΣX_i como ΣX.

Para ciertos modelos matemáticos usados en Diseño Experimental, se emplea expresiones como:

$$\sum_{i=2}^{n} X_i = X_2 + X_3 + \ldots + X_n$$

en las que no se toma en cuenta X_1 y se suma desde X_2 hasta X_n.

Otras dos expresiones de uso común son ΣX_i^2 y $(\Sigma X_i)^2$. La primera representa "la suma de los cuadrados de las X", es decir

$$\sum_{i=1}^{n} X_i^2 = X_1^2 + X_2^2 + X_3^2 + \ldots + X_n^2$$

En el ejemplo anterior,

$$\sum_{i=1}^{4} X_i^2 = 6^2 + 8^2 + 4^2 + 12^2 = 260$$

La segunda expresión $(\Sigma X_i)^2$ representa "el cuadrado de la suma de las X", es decir

$$\left(\sum_{i=1}^{n} X_i\right)^2 = (X_1 + X_2 + X_3 + X_n)^2$$

En el ejemplo,

$$\left(\sum_{i=1}^{4} X_i\right)^2 = (6 + 8 + 4 + 12)^2 = 30^2 = 900$$

Seguidamente, se da algunas notaciones y su desarrollo:

a)

$$\sum_{j=1}^{n} X_j Y_j = X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + \ldots + X_n Y_n$$

b) $\sum_{i=1}^{5}(X_i+2)=(X_1+2)+(X_2+2)+(X_3+2)+(X_4+2)+(X_5+2)$

$$=X_1+X_2+X_3+X_4+X_5+10$$

c) $\sum_{i=2}^{5}(4X_iY_i)=4X_2Y_2+4X_3Y_3+4X_4Y_4+4X_5Y_5$

d) $\sum_{j=1}^{4}F_jX_j^3=F_1X_1^3+F_2X_2^3+F_3X_3^3+F_4X_4^3$

e) Si $X_1=2, X_2=-5, X_3=4, X_4=-8$

$Y_1=-3, Y_2=-8, Y_3=10, Y_4=6$. calcular:

$\sum X_i, \sum Y_i, \sum X_iY_i, \sum X_i^2, \sum Y_i^2, (\sum X_i)(\sum Y_i), \sum X_iY_i^2$

$\sum X_i=(2)+(-5)+(4)+(-8)=-7$

$\sum Y_i=(-3)+(-8)+(10)+(6)=+5$

$\sum X_iY_i=[(2)(-3)+(-5)(-8)+(4)(10)+(-8)(6)]=+26$

$\sum X_i^2=(2)^2+(-5)^2+(4)^2+(-8)^2=109$

$\sum Y_i^2=(-3)^2+(-8)^2+(10)^2+(6)^2=209$

$(\sum X_i)(\sum y_i)=(-7)(5)=-35$

$\sum X_iY_i^2=[(2)(-3)^2+(-5)(-8)^2+(4)(10)^2+(-8)(6)^2]=-190$

Cuando se trata de una observación formada por dos o tres componentes, se usa el símbolo $\boldsymbol{X_i}$ y $\boldsymbol{X_{ijk}}$, respectivamente, en cuyo caso la suma total de las $\boldsymbol{X}$ se representa así:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}X_{ij} \quad y \quad \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}X_{ijk}$$

o simplemente,

$$\sum_i\sum_j X_{ij} \quad y \quad \sum_i\sum_j\sum_k X_{ijk}$$

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

Si, para un cálculo dado, los componentes **j** son mayores que los **i**, se escribe:

$$\sum_{i<j}\sum X_{ij} \quad \text{ó} \quad \sum_{i\neq j}\sum X_{ij}$$

Como puede apreciarse en el Cap. 19, existen combinaciones más complejas que las expuestas al comienzo del presente Capítulo. En esos casos es necesario comprender claramente las notaciones, a fin de evitar errores.

Si, para cruzamientos dialélicos, por ejemplo, el número de progenitores **p=3**, tendríamos el siguiente cuadro de **3x3**:

(Progenitor i) \ (Progenitor j)				Σ
	X_{11}	X_{12}	X_{13}	$X_{1.}$
	X_{21}	X_{22}	X_{23}	$X_{2.}$
	X_{31}	X_{32}	X_{33}	$X_{3.}$
Σ	$X_{.1}$	$X_{.2}$	$X_{.3}$	$X_{..}$

El cruzamiento, u observación X_{23}, corresponde a X_{ij}.

El cuadro presentado correspondería también a cualquiera combinación de dos factores, cada uno de los cuales tiene tres niveles.

Haciendo nuevamente referencia el Cap. 19, las notaciones para los diferentes cruzamientos, serían:

$$X_{i.} = \sum_{j} X_{ij} = X_{i1} + X_{i2} + X_{i3}$$

$$X_{.j} = \sum_{i} X_{ij} = X_{1j} + X_{2j} + X_{3j}$$

$$X_{..} = \sum_{i}\sum_{j} X_{ij} = X_{11} + X_{12} + \ldots + X_{33} \quad \text{(todas las nueve observaciones)}$$

$$X_{..} = \sum_{i\leq}\sum_{j} X_{ij} = X_{11} + X_{12} + X_{13} + X_{22} + X_{23} + X_{33}$$

$$X_{i.} = \sum_{i\neq j} X_{ij}\,; \qquad \text{p. ej: } X_{2.} = X_{21} + X_{23}$$

$$X_{.j} = \sum_{i\neq j} X_{ij}\,; \qquad \text{p. ej: } X_{.2} = X_{12} + X_{32}$$

$$X_{..} = \sum_{i\neq}\sum_{j} X_{ij} = X_{12} + X_{13} + X_{21} + X_{23} + X_{31} + X_{32}$$

4.2 Funciones

Decimos que Y es función de X, cuando a cada valor de X, corresponde un valor de Y; es decir, existe dependencia funcional de Y en X, la que se representa como $Y=f(X)$.

Si suponemos que el rendimiento R, es función de Fertilizante F, escribimos: $R=f(F)$.

Si $Y=f(X)$, se acostumbra a indicar "Y es función de X, cuando X vale tanto". Por ejemplo, si $X=5$, $Y=3(X) - 3$, sería igual a:

$$Y=3(5) - 3 = 12$$

4.3 Redondeo de cifras

Cuando se trabaja con decimales, y se necesita cierto grado de precisión, el investigador debe ser consistente en el redondeo de números. Así se evita acumular errores por redondeo, cuando se trabaja con un número crecido de cifras.

En general, la idea es aproximar el número a la cifra que se encuentra más cerca, superior o inferior. En el caso de 103.7, el redondeo es hacia 104.0, puesto que 103.7 está más cerca a 104.0 que a 103.0. Así mismo, 34.2246, se redondearía, si se quiere sólo dos decimales, a 34.22 y no a 34.23. En el caso de 34.465, la cifra quedaría como 34.46 y no 34.47, de acuerdo con el criterio generalizado de redondear al centésimo más bajo, cuando este es un número par y el milésimo es 5 o mayor; si el centésimo fuera un número impar, el redondeo sería al centésimo más alto: 34.475 quedaría como 34.48.

4.4 Parámetros y estadísticas

Como se ha indicado en la Sec. 2.4, los parámetros son cantidades fijas, que describen o definen a la población, en tanto que las estadísticas son cantidades variables que definen a la muestra. En el último caso, la estadística es

una cifra y no la ciencia que nos ocupa. Son de uso común en Estadística, los símbolos que se indican seguidamente, para representar parámetros y estadísticas:

	Parámetro (Población)	Estadística (Muestra)
Media	μ (miu)	$\overline{X}$
Variancia	σ^2	s^2
Desviación típica	σ	s
Desviación típica de las medias		$s_{\overline{x}}$
Desviación típica de la diferencia		$s_{\overline{d}}$
		z
		t
		X^2

Por cuanto el investigador trabaja generalmente con muestras, para, a base de ellas, sacar conclusiones sobre la población, se hace constar sólo como estadísticas las cantidades dadas por z, t y X^2, así como $s_{\overline{x}}$ y $s_{\overline{d}}$.

4.5 La media y otras medidas de tendencia central

Cuando se presenta datos, generalmente se hace referencia a un valor central ($\overline{X}$), y a otro que denote el grado de dispersión o variación en un juego de datos. Esta última cifra nos indicaría qué valores se hallan alrededor de la media (del centro) y cuáles se alejan hacia los extremos.

Expresiones como "Juan es de altura mediana", son vagas, aun cuando dan alguna información general. Se refieren a que, si la altura promedio de la población ecuatoriana, por ejemplo, es de 1.65, la estatura de Juan se halla alrededor de esa cifra.

Sin embargo, cuando el investigador colecta datos que implican gasto de tiempo, dinero y esfuerzo, él no puede darse el lujo de tener informaciones vagas. En una investigación de cualquier tipo, es necesaria una medida precisa de tendencia central, la que nos da un resumen parcial de los datos.

A pesar de que estas medidas de tendencia central son útiles, no dan información alguna sobre la variación de los datos, es decir, qué valores se hallan al centro y cuáles se ubican hacia los extremos. Tal sería el caso al determinar el sueldo promedio mensual de una institución: si el jefe gana $ 15.000,00; ocho funcionarios intermedios ganan $ 8.000,00 y 10 empleados inferiores ganan $ 2.500,00, el sueldo promedio mensual para la institución, sería de $ 5.473,00, valor que no refleja la remuneración del jefe, ni la de los empleados inferiores.

Seguidamente, vamos a tratar sobre las medidas más comunes de tendencia central, que son la media aritmética o promedio, la mediana y el modo.

4.5.1 La media aritmética

La media de un grupo formado por n números, $X_1, X_2, X_3, ..., X_n$, se denota por $\overline{X}$ (equis barra) y se define como,

$$\overline{X} = \frac{X_1 + X_2 + X_3 + ... + X_n}{n} = \sum_{i=1}^{n} \frac{X_i}{n} = \sum \frac{X}{n}$$

Por ejemplo la media de los números 6, 8, 10, 12 y 14, es

$$\overline{X} = \frac{6 + 8 + 10 + 12 + 14}{5} = \frac{50}{5} = 10$$

Si los números $X_1, X_2, X_3, ..., X_n$, aparecen con frecuencias $f_1, f_2, f_3, ..., f_k$, la media será,

$$\overline{X} = \frac{f_1X_1 + f_2X_2 + f_3X_3 + ... + f_kX_n}{f_1 + f_2 + f_3 + ... + f_k} = \frac{\sum_{i=1}^{n} f_iX_i}{\sum_{i=1}^{k} f_i} = \frac{\sum fX}{n}$$

donde $n=\sum f$, llamada frecuencia total.

Por ejemplo, si tenemos los números 4, 8, 12 y 14, que aparecen con frecuencias 2, 4, 3 y 1, respectivamente, la media será,

$$\overline{X} = \frac{(2)(4)+(4)(8)+(3)(12)+(1)(14)}{2+4+3+1}$$

$$= \frac{8+32+36+14}{10} = \frac{90}{10} = 9$$

Media ponderada. A veces se asocia a los números $X_1, X_2, X_3, ..., X_n$, con ciertos factores llamados de ponderación, $W_1, W_2, W_3, ... W_k$; en cuyo caso,

$$\overline{X} = \frac{W_1X_1 + W_2X_2 + W_3X_3 + ... + W_kX_n}{W_1 + W_2 + W_3 + ... + W_k} = \frac{\sum_{i=1}^{n} W_iX_i}{\sum_{i=1}^{n} W_i}$$

Propiedades de la media aritmética

a) La suma algébrica de las desviaciones de un grupo de números, desde su media, es igual a cero. Por ejemplo, la media de los números 4, 8, 12, 14 y 16, es $\overline{X} = \frac{54}{5} = 10.8$. Las desviaciones de cada número, con respecto a la media son:

X_i	$\overline{x}$	$X_i - \overline{x}$	
4	10.8	- 6.8	
			= - 9.6
8		- 2.8	
12		+ 1.2	$= 0; \sum_{i=1}^{n}(X_i - \overline{X}) = 0$
14		+ 3.2	
			= + 9.6
16		+ 5.2	

b) La suma de los cuadrados de las desviaciones de un grupo de números X_i, desde un número cualquiera **a** es un mínimo, si y solo si **a** = $\overline{X}$.

c) Si f_1 números tienen media m_1, f_2 números tienen media m_2,..., f_k números tienen media m_k, la media de todos los números es,

$$\overline{X} = \frac{f_1 m_1 + f_2 m_2 + f_3 m_3 + ... + f_k m_k}{f_1 + f_2 + f_3 + ... + f_k},$$

es decir que $\overline{X}$ es una media ponderada de todas las medias.

4.5.2 La mediana

La mediana de un grupo de números arreglados en orden de magnitud, es el valor medio (en posición), o la media aritmética de los dos valores centrales. Por ejemplo, en las series,

4, 8, 8, 10, **12,** 14, 18, 18, 20, la mediana es 12

4, 6, 8, **8, 12,** 13, 13, 16, la mediana es $\frac{8 + 12}{2} = 10$

4.5.3 El modo

El modo de un grupo de números, es el valor que aparece con mayor frecuencia. Es decir que puede no existir modo, o éste puede no ser único.

4.6 La desviación típica y otras medidas de dispersión

Como se indicó en la Sección precedente, se llama variación o dispersión de los datos, al grado en que éstos se ubican hacia los extremos de un valor promedio. Las dos medidas más comunes de dispersión son el rango y la variancia.

4.6.1 El rango

Es una medida muy fácil de establecer y se refiere a la diferencia entre el valor más alto y el más bajo, dentro de un grupo de datos. Por ejemplo, en los números,

4, 6, 7, 8, 12 y 23, el rango es 23 – 4 = 19

Como puede verse, el rango nos da una idea sobre el grado de variación de los datos.

4.6.2 La variancia

Llamada también **varianza** o **cuadrado medio**, es el cuadrado de la desviación típica (Sec. 4.6.3) y todas las propiedades de ésta, se aplican a aquélla. Se representa a la variancia de la población por σ^2 y a la de la muestra por s^2.

Estadísticamente, se define a la variancia como la suma de cuadrados de las desviaciones de un grupo de números con respecto a su media, dividida por el número de desviaciones menos uno (para el caso de la muestra). Para poblaciones finitas, se divide la suma de cuadrados de las desviaciones para N (el total de observaciones). Este concepto tiene que ver con el de **grados de libertad**, el cual se discute ampliamente en la Sec. 15.2.

El concepto de variancia puede ser algo difuso para el estudiante. Para aclararlo, vamos a servirnos de un ejemplo que es familiar para todos. En el caso de la clasificación de huevos para la venta al público, encontramos que al momento de recoger los huevos, en el criadero avícola, van a haber huevos muy grandes y otros muy pequeños, en cuyo caso su variancia va a ser grande. Después de clasificarlos por tamaño, para ponerlos en los cartones conocidos, los huevos serán más o menos uniformes, en cuyo caso la variancia será pequeña. Si todos los huevos fueran idénticos, su variancia sería cero.

La fórmula de difinición de la variancia está dada por,

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} = \frac{\sum X^2}{N}, \text{ para la población.}$$

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{\sum X^2}{n-1}, \text{ para la muestra.}$$

Si $X_1, X_2, X_3, ..., X_n$ observaciones ocurren con frecuencias $f_1, f_2, f_3, ..., f_k$, la fórmula sería,

$$\sigma^2 = \frac{\sum_{i=1}^{n} f_i (X_i + \mu)^2}{N} = \frac{\sum fX^2}{N}$$, para la población y

$$s^2 = \frac{\sum_{i=1}^{n} f_i (X_i - \overline{X})^2}{n-1} = \frac{\sum fX^2}{n-1}$$, para la muestra

Las fórmulas de definición pueden usarse cuando el número de observaciones es pequeño. Para el caso de análisis de datos provenientes de ensayos experimentales, en los que el número de datos es crecido, se usa la fórmula de trabajo, que da idénticos resultados:

$$\sigma^2 = \frac{\sum_{i=1}^{N} X_i^2 - \frac{(\sum X_i)^2}{N}}{N}$$, para la población y

$$s^2 = \frac{\sum_{i=1}^{n} X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}$$, para la muestra

Como se verá en el Cap. 15, $\sum_{i=1}^{n} X_i^2$ es lo que se llama la suma de cuadrados no corregidos, y $\frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}$, el factor de corrección y $n - 1$, los **grados de libertad**. Los dos términos del numerador constituyen la suma de cuadrados y el cociente que resulta de dividir la suma de cuadrados por los grados de libertad, se denomina la variancia o **cuadrado medio**.

Para demostrar el uso de las dos fórmulas para el cálculo de la variancia –la fórmula de definición así como la de trabajo– vamos a utilizar la muestra constituida por los siguientes datos. Por definición, el resultado debe ser idéntico.

3, 5, 6, 2, 4, 6, 2.

La fórmula de definición para la muestra está dada por:

$$s^2 = \frac{\sum_{i=1}^{7}(X_i - \overline{X})^2}{n-1}$$

Paso 1: Encontrar la suma de cuadrados (el numerador):

$$\sum X_i = 3 + 5 + 6 + 2 + 4 + 6 + 2 = 28$$

$$\overline{X} = \frac{\sum_{i=1}^{7} X_i}{n} = \frac{28}{7} = 4$$

$X_i - \overline{X}$	-1	1	+2	-2	0	+2	-2	
$(X_i - \overline{X})^2$	1	1	4	4	0	4	4	= 18

$$\sum_{i=1}^{7}(X_i - \overline{X})^2 = 18$$

Paso 2: Para obtener la variancia, dividir la suma de cuadrados (el numerador), para los grados de libertad *(n - 1)*.

$$s^2 = \frac{\sum_{i=1}^{7}(X_i - \overline{X})^2}{n-1} = \frac{18}{6} = 3.00$$

Paso 3: Para calcular la desviación típica, extraemos la raíz cuadrada de la variancia:

$$s = \sqrt{s^2} = \sqrt{3.00} = 1.73$$

Si queremos usar la fórmula de trabajo, para el mismo juego de datos, procederíamos así:

Paso 1: Calcular la suma de cuadrados no corregidos (el primer término del numerador):

$$s^2 = \frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1}$$

$$\sum X_i^2 = 3^2 + 5^2 + 6^2 + 2^2 + 4^2 + 6^2 + 2^2$$
$$= 9 + 25 + 36 + 4 + 16 + 36 + 4 = 130$$

Paso 2: Calcular el factor de corrección (segundo término del numerador):

$$\frac{(\sum X_i)^2}{n} = \frac{(28)^2}{7} = 112$$

Paso 3: Calcular la suma de cuadrados, corregidos:

$$\sum X_i^2 - \frac{(\sum X_i)^2}{n} = 130 - 112 = 18$$

Paso 4: Dividir la suma de cuadrados corregidos para los grados de libertad:

$$s^2 = \frac{18}{6} = 3.00$$

y la desviación típica: $s = \sqrt{3.00} = 1.73$

A pesar de que las dos fórmulas dan igual resultado, es preferible usar la fórmula de trabajo, la que elimina o reduce errores por redondeo, los que se producen al usar la fórmula de definición.

4.6.3 La desviación típica

Como se anotó en la Sección precedente, la desviación típica no es sino la raíz cuadrada de la variancia,

$$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(X_i - \mu)^2}{N}} = \sqrt{\frac{\sum X_i^2 - \frac{(\sum X_i)^2}{N}}{N}}$$, para la población y

$$s = \sqrt{\sum_{i=1} \frac{(X_i - \overline{X})^2}{n-1}} = \sqrt{\frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1}}$$, para la muestra

Para demostrar la relación entre el rango y la variancia, vamos a comparar cinco grupos de datos (poblaciones):

	Grupo	Σ	μ	Rango	σ^2	σ
a)	6,6, 6	18	6	0	0	0
b)	4,6, 8	18	6	4	2.66	1.63
c)	3,7, 8	18	6	5	4.66	2.15
d)	2,6,10	18	6	8	10.66	3.26
e)	0,1,17	18	6	17	60.66	7.78

a) $\sigma^2 = \dfrac{(6-6)^2+(6-6)^2+(6-6)^2}{3} = 0$

b) $\sigma^2 = \dfrac{(4-6)^2+(6-6)^2+(8-6)^2}{3} = \dfrac{8}{3} = 2.66$

c) $\sigma^2 = \dfrac{(3-6)^2+(7-6)^2+(8-6)^2}{3} = \dfrac{14}{3} = 4.66$

d) $\sigma^2 = \dfrac{(2-6)^2+(6-6)^2+(10-6)^2}{3} = \dfrac{32}{3} = 10.66$

e) $\sigma^2 = \dfrac{(0-6)^2+(1-6)^2+(17-6)^2}{3} = \dfrac{182}{3} = 60.66$

Como vemos, la media de los cinco grupos es la misma, pero el rango, σ^2 y σ van creciendo, conforme aumenta el grado de dispersión de los datos, alrededor de $\mu=6$; esta es la variación que se pretendió ilustrar con el ejemplo de los huevos.

4.7 Cambios en las observaciones y su influencia en μ y σ^2

Es de esperar que al cambiar las observaciones, la media y la variancia cambiarán. Al respecto, se puede establecer dos teoremas.

4.7.1 Si se suma o resta un valor fijo a cada una de las observaciones, la media aumentará o disminuirá en tal valor, pero la variancia y la desviación típica no se alteran.

En el juego de datos, 12, 13 y 14, su media es $\mu = \frac{39}{3} = 13$.

Si sumamos 10 a cada observación, los nuevos valores son, 22, 23 y 24, y su media $\mu = \frac{69}{3} = 23$, que es justamente 10 más que la media original.

Si restamos 10 de cada observación, los nuevos valores son 2, 3 y 4 y su media $\mu = \frac{9}{3} = 3$, que es 10 menos que la media original. La variancia, en cambio, no se altera;

$$\sigma^2 = \frac{(12-13)^2 + (13-13)^2 + (14-13)^2}{3} = \frac{1+0+1}{3} = \frac{2}{3} = 0.66$$

Si añadimos 10 a cada observación,

$$\sigma^2 = \frac{(22-13)^2 + (23-23)^2 + (24-23)^2}{3} = \frac{1+0+1}{3} = \frac{2}{3} = 0.66$$

4.7.2 Si se multiplica cada observación por un valor fijo m, la nueva media es m veces la media original; la nueva variancia es m^2 veces la variancia original y la nueva desviación típica es m veces la desviación típica original.

En el ejemplo anterior, 12, 13 y 14, sabemos que $\mu = 13$.

La variancia original es 0.66. La nueva variancia será:

$$\sigma^2 = \frac{2}{3}; \sigma = \sqrt{\frac{2}{3}} = \sqrt{0.66} = 0.81$$

Si multiplicamos las observaciones por m=10, los nuevos valores serán 120, 130 y 140; su media μ = 130, que es justamente m veces, o sea 10 veces mayor que la media original 13.

$$\sigma^2 = \frac{(120-130)^2 + (130-130)^2 + (140-130)^2}{3}$$

$$= \frac{100 + 0 + 100}{3} = \frac{200}{3} = 66.66$$

que es justamente m^2 veces la variancia original.

La nueva desviación típica será:

$\sigma = \sqrt{\frac{200}{3}} = \sqrt{66.66} = 8.1$, que es m veces la σ original.

Estos mismos principios se aplican a la división, puesto que no es sino una multiplicación fraccionaria.

4.8 Desviación típica de las medias

De una población de observaciones, es posible sacar submuestras, las que tendrán su propia $\overline{X}$ y s. Dichas estadísticas constituyen una población de $\overline{X}$ y s y, naturalmente, se hallan sujetas a variación. Se espera que esta variación va a ser menor que la existente entre las observaciones individuales, a las cuales representan.

Para la población,

$$\sigma_{\mu}^2 = \frac{\sigma}{N}; \sigma_{\mu} = \sqrt{\frac{\sigma^2}{N}} = \frac{\sigma}{\sqrt{N}}$$

Para la muestra,

$$s_{\overline{x}}^2 = \frac{s^2}{n}; s_{\overline{x}} = \sqrt{\frac{s^2}{n}} = \frac{s}{\sqrt{n}}$$

Supongamos que tenemos la muestra formada por los números 4, 6, 8 y 10; para los que $s^2 = 6.66$ y $s = 2.58$.

$$s_{\overline{x}} = \frac{s}{\sqrt{n}} = \frac{2.58}{\sqrt{4}} = \frac{2.58}{2} = 1.29$$

O también,

$$s_{\overline{x}} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{6.66}{4}} = \sqrt{1.64} = 1.29$$

4.9 Coeficiente de Variación (C. V.)

Es una medida relativa de variabilidad, independiente de las unidades de medida, en contraste con la desviación típica, la que se expresa en las mismas unidades de medida que las observaciones cuyo grado de homogeneidad establece.

El coeficiente de variación para peso de estudiantes será el mismo si los datos se registran en libras, kilogramos, toneladas, etc. Se puede comparar el coeficiente de variación para pesos de caballos con el de pollos. De hecho, si queremos comparar el grado de dispersión de dos juegos de datos que tienen diferentes medias, es más interesante comparar los coeficientes de variación que las desviaciones típicas correspondientes. Si un tipo de vacuna da una protección promedio de 12 meses con $s = 30$ días, y otro tipo de vacuna con una duración de 24 meses con $s = 60$ días, parecería que el segundo tipo es el más variable. Sin embargo, si observamos que los coeficientes de variación son de 25 y 14%, respectivamente, concluiremos que, con relación a su duración promedio, la primera vacuna es más variable.

El coeficiente de variación es generalmente mayor en ensayos con animales que con plantas y en el campo que en ensayos de invernadero, donde las condiciones ambientales son controladas.

El investigador usa el **C.V.** para evaluar resultados de diferentes experimentos que tratan sobre el mismo factor y que han sido llevados a cabo por diferentes personas. El **C.V.** mide también el grado de precisión del diseño y de la conducción del experimento. Por ejemplo, si en la siembra de las parcelas experimentales, o en la fertilización, se emplea jornaleros, podemos estar seguros de que el **C.V.** puede acercarse al doble de lo que sería, en caso de usar maquinaria o equipos calibrados. Esto es especialmente cierto en situaciones en las que se utiliza varios jornaleros indistintamente para efectuar cierta labor, la misma que debiera realizarse uniformemente para todas las parcelas dentro del **bloque** o **repetición**.

Para decidir si un **C.V.** es grande o pequeño, es necesario tener experiencia sobre datos similares.

Se define al **C.V.** como la desviación típica de la muestra, expresada como porcentaje de la muestra,

$$C.V. = \frac{s}{\overline{X}}(100)$$

En la Sec. 4.11 se ilustra el uso del **C.V.**

4.10 Modelo lineal aditivo

El desdoblamiento de los componentes que forman una observación cualquiera X_i, da lugar a un modelo que, en su forma más simple, consiste de la media más un error

$$X_i = \mu + \varepsilon_i$$

Se supone que los ε_i son al azar y provienen de una población de ε con media cero, es decir que la población de X_i ha sido muestreada al azar.

El tamaño del error será menor conforme aumenta el tamaño de la muestra y, en ese caso, podemos decir que $\overline{X}$ es un buen estimador de μ.

En el caso de un diseño de bloques completos al azar (ver Sec. 16.2), el modelo lineal aditivo está dado por:

$$X_i = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

en el que, a la media de una observación X_{ij}, se suman los efectos de tratamientos, bloques y del error experimental.

4.11 Intervalos de confianza

A partir de los datos de la muestra, es posible inferir sobre la ubicación de μ. Este es un proceso inductivo, en el que se razona de la parte al todo. Puesto que $\overline{X}$ es una variable (población de $\overline{X}$), dudaríamos de afirmar que μ esta en $\overline{X}$. Sería preferible llegar a un intervalo, quizás con $\overline{X}$ al centro, y afirmar que confiamos en que μ se halla dentro de ese intervalo.

En la forma más simple, esta afirmación podría escribirse así,

$$P\ (3 < \mu < 6) = .95,$$

es decir, que la probabilidad de que μ se encuentre comprendida entre tres, límite inferior, y seis, límite superior, es de 95%. En otras palabras, hay cinco probabilidades entre 100 o una entre 20, de que esta afirmación sea equivocada.

A fin de tener mayor grado de confianza en la predicción, se utiliza la estadística **t. de Student,** la que se presenta con amplitud en el Cap. 10.

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}}$$

fórmula en que la única incógnita es μ. Para el cálculo estadístico de los intervalos de confianza, buscamos el valor tabular de t el que es parte de la fórmula

$$P = [\overline{X} - (t_{.05})(s_{\overline{X}}) < \mu < \overline{X} + (t_{.05})(s_{\overline{X}})] = 0.95$$

donde t es el valor tabular al 95% de probabilidades, de acuerdo con el número de grados de libertad $(n - 1)$. En la Tabla V del Apéndice encontramos, por ejemplo, que el valor de t al 95% (0.05), para seis grados de libertad, es 2.447.

Como ilustración, supongamos que se ha probado siete muestras de una nueva variedad de maíz y queremos predecir el rendimiento de la variedad (población), en una localidad dada. Los siguientes son los rendimientos obtenidos:

Localidad	Rendimiento
1	80
2	62
3	75
4	48
5	76
6	82
7	60
	483

$$\overline{X} = \frac{483}{7} = 69$$

$$s^2 = \frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1}$$

$$= \frac{34.273 - \frac{233.289}{7}}{7-1} = \frac{34.273 - 33.327}{6}$$

$$= \frac{946}{6} = 158; \quad s = \sqrt{158} = 12.56\ qq.$$

$$s_{\overline{X}} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{158}{7}} = \sqrt{22.57} = 4.75$$

$$s_{\overline{X}} = \frac{12.56}{\sqrt{7}} = 4.75$$

$$C.V. = \frac{s}{\overline{X}}(100) = \frac{12.56}{69}(100) = 18\%$$

$$t_{.05}(6\ g.\ de\ l.) = 2.447$$

Intervalo de confianza, al 95%,

$$P = [\overline{X} - (t_{.05})(s_{\overline{X}}) < \mu < \overline{X} + (t_{.05})(s_{\overline{X}})] = 0.95$$

$$69 - (2.447)(4.75) < \mu < 69 + (2.447)(4.75)$$

$$57.38 < \mu < 80.62$$

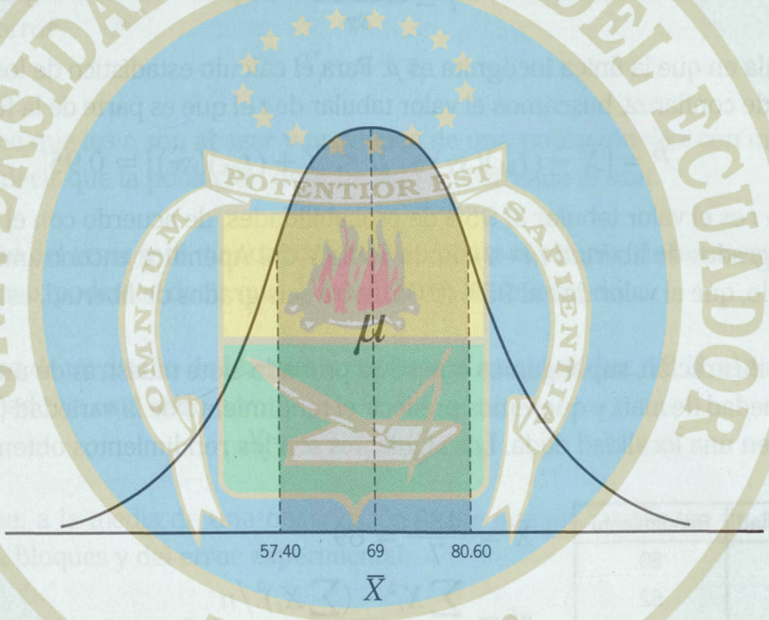


Fig. 3.1 Ubicación de μ e intervalos de confianza

La media de la muestra $\overline{X}$ = 69 y la desviación típica s = 12.56, nos dan los mejores estimados de los parámetros correspondientes μ y σ, respectivamente. Puesto que nuestro interés está en el rendimiento de la población (variedad de maíz), cuya muestra hemos probado, se ha calculado un intervalo y afirmamos que μ se encuentra entre 57.40 y 80.60, con una probabilidad de 95% de estar en lo cierto.

Para aclarar de otro modo el significado del intervalo de confianza, digamos que queremos determinar si la utilización de un balanceado *X* es económica. El costo del producto es tal, que su uso sería económico sólo en el caso de que deje una ganancia de peso, en novillos de engorde, de 1,5 libras diarias, en promedio. Si después de realizar el experimento respectivo, encontramos que los límites de confianza al 95%, son de 1,8 y 2,0 libras diarias, estaríamos seguros de que conviene usar el balanceado. Pero si los límites de confianza son de 1.0 y 1,2 libras, se descartaría el empleo del producto. Habría duda, en caso de que los límites fueran 1,0 y 2,0, libras, porque puede haber una pequeña ganancia o una pequeña pérdida; en todo caso habría riesgo al hacer la recomendación. Lo adecuado sería repetir el ensayo.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 5

CUADROS Y CURVAS DE FRECUENCIA – HISTOGRAMAS

5.1 Cuadros de frecuencia

En la Sec. 2.6 se indicó la forma de presentar datos numéricos en forma gráfica. A veces, este sistema llama la atención del lector no especializado, más que si los datos estuvieran incluidos en un cuadro numérico. De todos modos, cuando se dispone de una gran cantidad de cifras, es útil para el investigador, tabularlas antes de confeccionar cuadros, figuras o realizar cualquier tipo de análisis. Por otro lado, para cierto tipo de revistas científicas, es importante presentar en un cuadro compacto, el resumen de los datos, incluyendo la frecuencia de las observaciones dentro de cada clase o categoría. A este resumen se llama justamente cuadro de frecuencias. Los datos que se encuentran desordenados en el libro de campo pueden, en esa forma, ser manejados más fácilmente. Finalmente, de esos cuadros de frecuencia, se puede calcular $\overline{X}$, s^2 y s, en forma más fácil que a partir de los datos originales.

Supongamos que tenemos los pesos de 10 cerdos en libras,

80, 85, 80, 70, 75, 80, 75, 80, 70 y 75

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST SAPIENTIA
FUNDADA EN 1826
QUITO
UNIVERSIDAD CENTRAL DEL ECUADOR
BIBLIOTECA

Para determinar la frecuencia de clase, es útil confeccionar un cuadro de control, como el Cuadro No 5.1.

CUADRO No 5.1 Control de datos

Clase	Control	Recuento
70	//	2
75	///	3
80	////	4
85	/	1
	Total	10

Del Cuadro puede apreciarse que existen dos cerdos de 70 libras, tres de 75, cuatro de 80 y uno de 85 libras. Es decir que, del grupo desordenado de datos, se ha formado cuatro subgrupos o clases.

El número de observaciones o individuos dentro de cada clase, se llama frecuencia de clase. El número total de observaciones de todas las clases, se llama frecuencia total. Una vez establecidas las clases, se puede determinar la frecuencia relativa y acumulada de las observaciones, como se indica en el Cuadro No 5.2.

CUADRO No 5.2 Frecuencias relativas y acumuladas de clase

Peso (X)	Frecuencia (f)	Frecuencia relativa (f.r.)	Frecuencia acumulada (f.a.)
70	2	20%	20%
75	3	30%	50%
80	4	40%	90%
85	1	10%	100%
Total:	10	100%	

La frecuencia relativa indica el porcentaje de individuos dentro de cada clase, con relación a la frecuencia total; la frecuencia acumulada, indica el porcentaje de individuos que se encuentran hasta un peso dado. Por ejemplo, vemos

que el 50% de los cerdos llegan hasta las 75 libras, el 90% hasta las 80 libras y, obviamente, el 100% de los cerdos alcanzan las 85 libras.

El Cuadro No 5.3 demuestra la forma de calcular $\overline{X}$, s^2 y s, a partir de un cuadro de frecuencias.

CUADRO No 5.3 Cálculo de $\overline{X}$, s^2 y s, de un cuadro de frecuencias

X	f	fX	$(X - \overline{X})$	$(X - \overline{X})^2$	$f(X - \overline{X})^2$
70	2	140	- 7.0	49	98
75	3	225	- 2.0	4	12
80	4	320	3.0	9	36
85	1	85	8.0	64	64
Σ	10	770	$X = 770 / 10 = 77,0$		210

Puede llamar la atención del lector, el hecho de que $\Sigma(X - \overline{X})$, no sea igual a cero, como lo establece una propiedad de la media aritmética. Esto se debe a que en el Cuadro No 5.3, constan solamente los cuatro pesos (clases), y no los 10 pesos considerados. Si se resta la media de los 10 pesos, la suma de las desviaciones dará cero.

$$\overline{X} = \frac{770}{10} = 77$$

$$s^2 = \frac{\sum f(X - \overline{X})^2}{n - 1} = \frac{210}{9} = 23.33$$

$$s = \sqrt{23.33} = 4.83$$

5.2 Histogramas y polígonos de frecuencia

Se llama histograma al gráfico construido a partir del cuadro de frecuencias. Consta de un conjunto de rectángulos, cuya altura es proporcional a la frecuencia de clase.

Para nuestro ejemplo, la Fig. 5.1 representa el número de cerdos dentro de cada clase y el porcentaje con respecto al total.

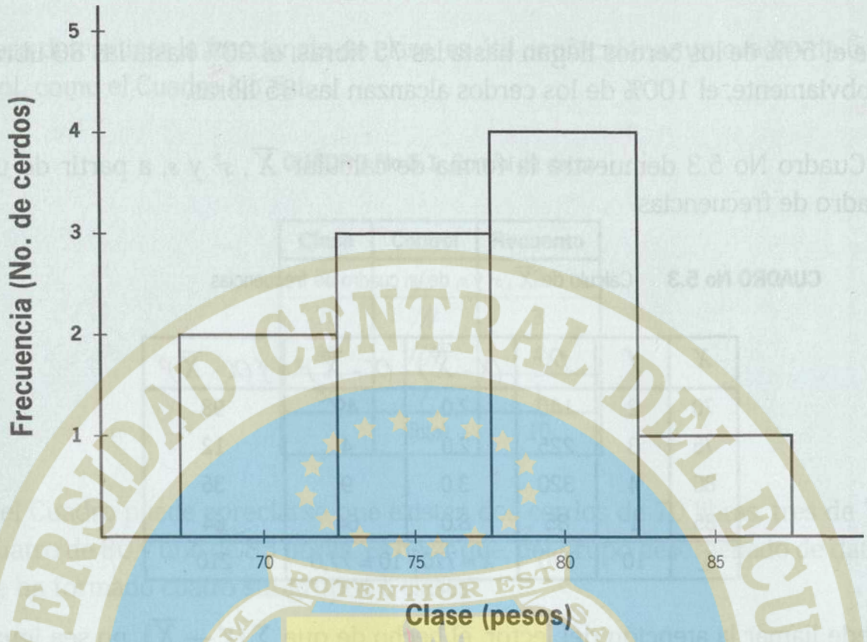


Fig. 5.1 Número de cerdos y porcentaje dentro de cada clase.

Por medio del histograma, se puede observar de un vistazo, la naturaleza de la distribución de los datos. Si se desea comparar una distribución teórica con la real, se puede graficar la una sobre la otra y averiguar la razón de la diferencia.

El polígono de frecuencias, se prepara uniendo los puntos medios de cada clase, por medio de líneas rectas. El polígono de frecuencias tiende a sugerir el tipo de curva de la población de la que se extrajo la muestra.

El histograma y el polígono de frecuencias son formas más objetivas que el cuadro de frecuencias, para presentar datos.

El siguiente ejemplo permitirá una clara comprensión sobre la forma en que se construye un histograma y el polígono de frecuencias respectivo.

Se ha registrado la altura de 100 plantas tomadas al azar, de una parcela de pasto azul. Los datos constan en el Cuadro No 5.4

CUADRO No 5.4 Altura de 100 plantas de pasto azul

Límites verdaderos de clase	Intervalo de clase (Altura en cm.)	Frecuencia (No. plantas)	Marca de clase
59,5 – 62,5	60 – 62	5	61
> 62,5 – 65,5	63 – 65	18	64
> 65,5 – 68,5	66 – 68	42	67
> 68,5 – 71,5	69 – 71	27	70
> 71,5 – 74,5	72 – 74	8	73
		$\sum = 100$	

Como puede apreciarse, la menor altura encontrada fue de 60 cm. Por redondeo aritmético, todas las plantas que midan entre 59,5 y 62,5, pertenecen al primer intervalo. Una planta que mida 62,6 cm., entraría al segundo intervalo. En la práctica, se obtiene los verdaderos límites de clase, sumando el límite superior de una clase, al límite inferior de la siguiente y dividiendo para dos:

$$\frac{62 + 63}{2} = 62.5$$

Por conveniencia, se ha establecido que el tamaño de clase sea tres; al restar 62,5 – 3,0 = 59,5, en cuyo caso tenemos el límite verdadero inferior, que es el punto de partida. La marca de clase se obtiene sumando los límites verdaderos inferior y superior de una clase y dividiendo para dos:

$$\frac{59.5 + 62.5}{2} = 61$$

O sumando los límites inferior y superior del intervalo de clase y dividiendo para dos:

$$\frac{60 + 62}{2} = 61$$

De todos modos, la diferencia entre los límites verdaderos de clase, así como entre las marcas de clase, debe ser tres.

Con los datos del Cuadro N° 5.4, estaríamos listos para construir el histograma y el polígono de frecuencias, que se vería como en la Fig. 5.2

Como reglas generales para la construcción de histogramas, se pueden mencionar:

a) Determinar el rango.

b) Dividir el rango en un número conveniente de intervalos de clase del mismo tamaño. Este número va generalmente de 5 a 20, de acuerdo a los datos.

c) Determinar el número de observaciones dentro de cada intervalo de clase.

d) A partir de estos datos, pueden construirse histogramas y polígonos de frecuencia.

En el ejemplo propuesto,

a) Rango: 74 - 60 = 14

b) Para intervalos de clase tamaño 3, $\frac{14}{3} = 4,66 \cong 5$ clases.

Para determinar el número de clases, se puede también usar la relación empírica

$$k = 1 + 3.3 \log n$$

donde k es el número de clases, n es el número de valores u observaciones y $\log n$ es el logaritmo común (base 10), de n.

Para los 10 pesos de cerdos del Cuadro No 5.1. tenemos:

$$k = 1 + 3.3 \log (10)$$
$$= 1 + 3.3 (1) = 4.3 \simeq \text{a 4–5 clases}$$

No es conveniente tener muchos intervalos, por cuanto el trabajo aumenta, sin mayor ganancia de información; ni muy pocos intervalos, en cuyo caso se sacrificaría información.

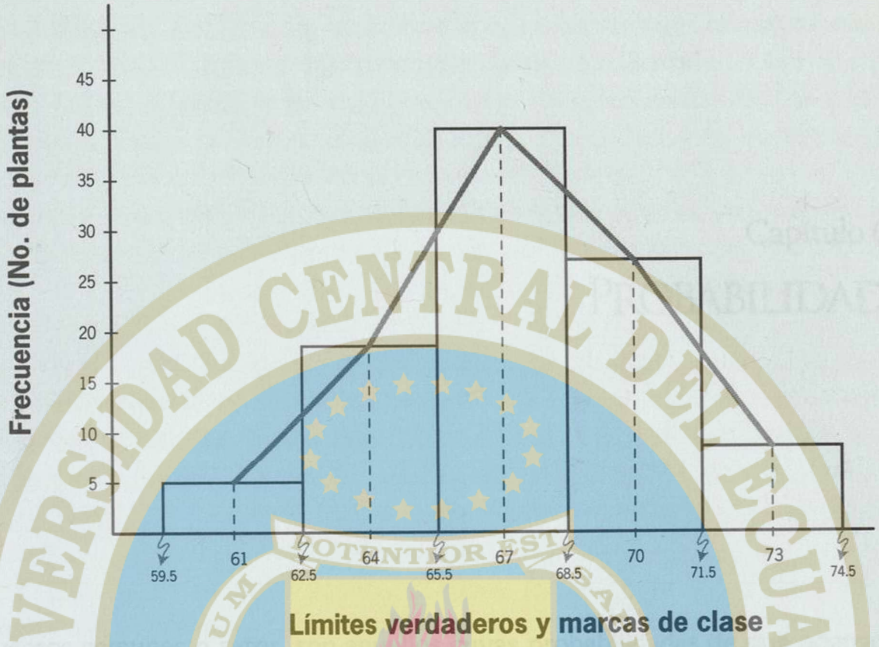


Fig. 5.2 Histograma y polígono de frecuencias de la altura de 100 plantas de pasto azul.

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM
SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 6

PROBABILIDAD

Sucesos comunes o raros, son aquellos cuyas probabilidades de que ocurran son grandes o pequeñas, respectivamente. En la vida diaria, usamos todo el rango de probabilidades, de una u otra manera. Así decimos "la muerte de *X* pudo deberse a que manejaba descuidadamente", cuando no estamos muy seguros de la causa del fallecimiento; o, "la muerte de *X* seguramente se debió a que se le rompieron los frenos del automóvil", cuando tenemos mayor seguridad sobre la causa del accidente.

En Estadística, se reemplaza las palabras "pudo" y "seguramente", por números que varían de cero a uno. Es decir, se indica con precisión que tan probable o improbable puede ser un suceso o evento.

El cálculo de probabilidades tiene que ver con variables discretas, involucradas en problemas de muestreo, como encuestas, estudio de caracteres genéticos, etc. Lógicamente, cuando se habla de muestreos, no es posible disponer de toda la población. Se ha anotado en el Cap. 2 que se razona a partir de la muestra, para luego sacar inferencias sobre la población; en consecuencia, no es posible obtener todas las predicciones correctas. En el proceso de predicción, el azar juega un papel predominante.

Sin embargo, la Estadística dispone de métodos que permiten afirmar con que frecuencia las afirmaciones serán correctas, en promedio.

El uso de un número fraccionario o relación, es común en Estadística. Así, se dice que el equipo "Nacional" tiene cinco chances contra tres (5 : 3), de ganar al Emelec. Esto se interpretaría como que el equipo de casa puede ganar por un margen estrecho; o que, si los dos equipos jugaran varias veces, el "Nacional" ganaría 5/8 o el 62% de los juegos:

$$5/(5 + 3) = 5/8 = 0{,}62$$

Los chances de perder, o de que gane el Emelec, serían de 3/8 = 0,38 = 38%. Como se ve, las probabilidades de ganar o perder, son valores que pueden estar entre 0 y 1, como se anotó al comienzo del Capítulo.

Supongamos que un evento *Ei* puede ocurrir en *r* formas de un total de *N* formas, igualmente posibles. La probabilidad *P* de que el evento ocurra **(probabilidad de éxito)**, está dada por:

$$P(E_i) = \frac{p}{N}, \text{ es decir } P(E_i) = \frac{\text{número de éxitos}}{\text{número total de eventos}}$$

El número total de eventos está dado por el total de **éxitos** más **fracasos**, llamándose fracaso a la probabilidad de que un evento no ocurra.

La probabilidad de fracaso *(q)* se calcula así,

$$P(No\ E_i) = q = \frac{N - p}{N} = 1 - p = 1 - P(E_i)$$

De lo expuesto, vemos que $p + q = 1$, o $P(E_i) + P(No\ E_i) = 1$

Se ha anotado antes que el cálculo de probabilidades se aplica a muchos campos de orden práctico. Experimentos simples de orden repetitivo, como el lanzamiento de una moneda, pueden ilustrar como funciona el cálculo de probabilidades. Supongamos que tiramos una moneda dos veces o lo que es lo mismo, tiramos dos monedas una vez. En los dos casos existen cuatro po-

sibilidades: CC, CS, SC, SS; la combinación SC, en el tercer caso, significaría que se ha obtenido sello en el primer lanzamiento y cara en el segundo. Si en lugar de dos, tiramos tres monedas, pueden producirse cualesquiera de ocho resultados:

CCC, CCS, CSC. SCC. CSS, SCS, SSC o SSS

Para experimentos de este tipo, es necesario conocer la totalidad de los resultados posibles y hacer una lista de aquellos que son de nuestro interés. Tal sería el caso al averiguar la probabilidad de sacar siete al tirar dos dados. Como cada dado puede dar seis resultados, las combinaciones posibles son 36, siendo de nuestro interés, las siguientes:

(3, 4), (2, 5), (1, 6) y las recíprocas
(4, 3), (5, 2), (6, 1)

es decir que la probabilidad de sacar siete, en este caso, sería = $\frac{6}{36} = \frac{1}{6}$

6.1 Sucesos independientes

Si $E_1, E_2, E_3, \ldots, E_n$, corresponden a n eventos independientes, los cuales tienen probabilidades $p_1, p_2, p_3, \ldots, p_n$ la probabilidad de que dichos eventos se cumplan, será $p_1 p_2 p_3 \ldots p_n$; es decir que la ocurrencia de E_1 no afecta a que E_2 o E_3 se produzcan.

En este caso, tenemos lo que se llama la regla de la multiplicación y que se expresa así,

$$P(E_1, E_2, E_3) = P(E_1)\, P(E_2)\, P(E_3)$$

Cuando se produce un evento E_1 y este no afecta la ocurrencia de E_2, la expresión sería:

$$P(E_1, E_2) = P(E_1)\, P(E_2 / E_1),$$

o lo que es igual, para sucesos independientes

$$P(E_1, E_2) = P(E_1)\,P(E_2)$$

Ejemplos:

a) Calcular la probabilidad de que, al sacar dos cartas de una baraja ordinaria de 52 cartas, con reemplazo, la primera sea el as de diamante y la segunda el as de trébol. Si llamamos E_1 al evento "as de diamante" y E_2 al evento "as de trébol"; tenemos que,

$$P(E_1) = \frac{1}{52}\,;\, P(E_2) = \frac{1}{52}$$

$$P(E_1, E_2) = \frac{1}{52} \times \frac{1}{52} = \frac{1}{2704}$$

b) Si se lanza un dado dos veces, encontrar la probabilidad de sacar 2, 3 ó 5 en el primer lanzamiento y 1, 4, 5 ó 6 en el segundo.

Si llamamos E_1 al evento "2, 3, 5" y E_2 al evento "1, 4, 5 ó 6", vemos que las seis formas en que puede caer el primer dado, están asociadas con las seis formas en que puede caer el segundo; es decir, existen 6 x 6 = 36 combinaciones. Las tres formas en que puede caer el primer dado tiene relación con las cuatro formas en que puede caer el segundo, para dar 3X4=12 formas en que E_1 y E_2 pueden ocurrir; entonces,

$$P(E_1, E_2) = P(E_1)P(E_2) = \frac{3}{6} \times \frac{4}{6} = \frac{12}{36} = \frac{1}{3}$$

c) Si lanzamos una moneda cuatro veces, la probabilidad de obtener cuatro caras en forma sucesiva es:

$$P(4\ caras) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

d) Si una urna contiene cinco bolas blancas y tres negras y otra urna tiene tres bolas blancas y dos negras y extraemos una de cada urna, la probabilidad de que las dos sean blancas es

$$P(B,B) = P(B)P(B) = \frac{5}{8} \times \frac{3}{5} = \frac{3}{8}$$

6.2 Sucesos dependientes

La probabilidad de un suceso dependiente de otro, se llama probabilidad condicional, en la que la probabilidad de que ocurra E_2, depende de la ocurrencia de E_1, así,

$$P(E_1, E_2) = P(E_1)P\left(\frac{E_2}{E_1}\right)$$

Ejemplos:

a) Cuál es la probabilidad de sacar un as de diamante, de una baraja ordinaria de 52 cartas, después de haber sacado, sin reemplazo, el As de trébol?

Si designamos al evento as de diamante E_2 y al evento as de trébol E_1, tendremos:

$$P(E_1) = \frac{1}{52}; \quad P\left(\frac{E_2}{E_1}\right) = \frac{1}{51}$$

$$P(E_1, E_2) = P(E_1)P\left(\frac{E_2}{E_1}\right) = \frac{1}{52} \times \frac{1}{51} = \frac{1}{2652}$$

Por cuanto se ha extraído una carta, sin reemplazo, quedan 51 cartas en la baraja, para la segunda extracción.

b) En un ánfora hay 25 papeletas, cinco de las cuales son premiadas. Cuál es la probabilidad de sacar dos premios al extraer las dos primeras papeletas?

$$P(E_1, E_2) = \frac{5}{25} \times \frac{4}{24} = \frac{1}{30}$$

c) Una urna contiene cuatro bolas blancas y cinco negras; sacamos dos en forma sucesiva sin reemplazo. ¿Cuál es la probabilidad de que las dos bolas sean blancas? La probabilidad de que la primera sea blanca es 4/9. Si la primera bola fue blanca, la probabilidad de que la segunda también lo sea es 3/8, en cuyo caso:

$$P(E_1, E_2) = P(E_1)P(E_2)$$
$$= \frac{4}{9} \times \frac{3}{8} = \frac{1}{6}$$

6.3 Sucesos mutuamente excluyentes

Son aquellos en los que la ocurrencia de uno de ellos excluye la ocurrencia de otros eventos. Es decir, que si E_1 y E_2 son dos eventos mutuamente excluyentes, $P(E_1E_2) = 0$; los dos eventos no pueden ocurrir, como en el caso del nacimiento de un hijo en que E_1 sería "varón" y E_2 "mujer".

En estos casos se aplica la regla de la adición,

$$P(E_1 + E_2) = P(E_1) + P(E_2)$$

Si dos eventos presentan el caso de que E_1 o E_2 o los dos ocurran,

$$P(E_1 + E_2) = P(E_1) + P(E_2) - P(E_1E_2),$$

como en el siguiente ejemplo,

a) Calcular la probabilidad de que la primera carta sacada de una baraja, sea jota (E_1) o corazones (E_2),

$$P(E_1 + E_2) = P(E_1) + P(E_2) - P(E_1E_2)$$
$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

en este caso, estamos restando la posibilidad de que la carta sacada sea "la jota de corazones", que tiene una probabilidad de 1/52.

b) Encontrar la probabilidad de que una K, el 8 de trébol o el 5 de picas, aparezca en una extracción de una de 52 barajas. El evento puede ocurrir en una de seis formas:

$$\frac{4}{52} + \frac{1}{52} + \frac{1}{52} = \frac{6}{52} = \frac{3}{26}$$

c) Cuál es la probabilidad de que, en el próximo lanzamiento de un dado salgan el tres o el cuatro? Hay seis formas en que puede caer el dado: 1, 2, 3, 4, 5 ó 6. Si E_1 puede ocurrir en dos de las seis formas:

$$P(E_1) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}; \quad P(No\ E_1) = 1 - \frac{1}{3} = \frac{2}{3}$$

d) Queremos comprar tres tomates de un camión que va a descargarlos en el mercado; elegimos uno al azar y lo calificamos como bueno (B) y lo dejamos en su lugar. Repetimos la operación con otros dos tomates. En esta muestra de tres hemos encontrado dos buenos y uno malo (M). Cuál es la probabilidad de este resultado si sabemos que, por un muestreo previo, existe 10% de tomates en mal estado en el camión?

X (variable aleatoria) = cantidad de tomates en mal estado. Por tanto, X puede valer 0, 1, 2 ó 3.

El resultado fue $X = 1$.

La probabilidad de elegir un tomate malo en una determinada extracción es $p = 0.10$; la probabilidad de que sea bueno es $(1 - p) = 0.90$

Hay tres formas en las que un tomate malo puede salir en una muestra de tres:

M, B, B
B, M, B
B, B, M

formas que son mutuamente excluyentes, en cuyo caso sumamos la probabilidad total de cada extracción. Dentro de cada extracción, se trata de eventos independientes, porque el resultado de la primera extracción no afecta a la segunda ni a la tercera. Por tanto, tenemos:

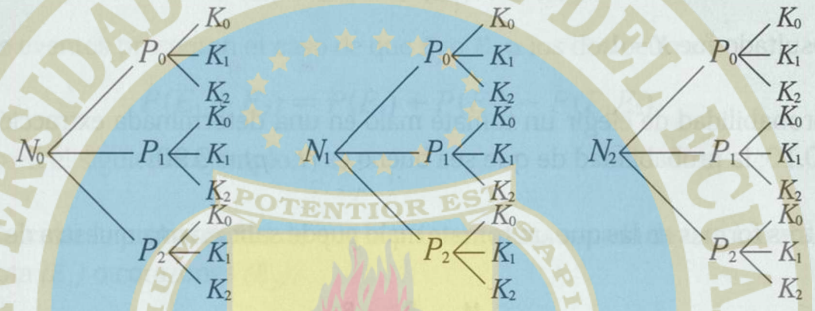
$$P(M, B, B) = P(1-p)(1-p) = (0.1)(0.9)(0.9) = 0.081$$
$$P(B, M, B) = (1-p)(p)(1-p) = (0.9)(0.1)(0.9) = 0.081$$
$$P(B, B, M) = (1-p)(1-p)(p) = (0.9)(0.9)(0.1) = 0.081$$
$$P(X = 1) = P(M, B, B) + P(B, M, B) + P(B, B, M) = 0.243$$

$P(X = 1) = 0.243 = 1/4$, la probabilidad de sacar en una extracción cualquiera, dos tomates buenos y uno malo.

6.4 Análisis combinatorio

Para resolver probabilidades de eventos complejos, como algunos juegos de azar, es conveniente disponer de fórmulas que ahorran tiempo y trabajo, el que, de otro modo, sería difícil y tedioso.

El análisis combinatorio utiliza dichas fórmulas, como se indica en esta Sección. Supongamos que deseamos aplicar tres niveles de nitrógeno, tres de fósforo y tres de potasio. Todas las combinaciones posibles serían:



Con un total de 27 tratamientos o niveles. Se identifica a las dosis cero como niveles; tal es el caso de *No, Po* y *Ko*.

6.5 Factorial N

Se define a **N!** (N factorial), en la siguiente forma:

$$N! = N(N-1)(N-2)(N-3) \ldots 1$$

Por ejemplo, 4! = 4 x 3 x 2 x 1 = 24

$$7!3! = (7\times6\times5\times4\times3\times2\times1)(3\times2\times1) = 30.240$$

$$\frac{7!3!}{4!} = \frac{(7\times6\times5\times4\times3\times2\times1)(3\times2\times1)}{(4\times3\times2\times1)} = 1.200$$

$$0! = 1$$

Se puede simplificar valores iguales del numerador y del denominador, para ahorrar trabajo.

6.6 Permutaciones

La permutación es una selección de N objetos, tomados r a la vez, respetando el orden de arreglo,

$$N\text{Pr} = (N-1)(N-2)...(N-r+1) = \frac{N!}{(N-r)!}$$

Al símbolo $N\text{P}r$, se llama la permutación de N objetos tomados r a la vez. Así, hay N formas de seleccionar el objeto para el primer lugar; luego, $N-1$ formas para el segundo lugar, hasta llegar a la *r-ésima* forma, para el último lugar, para el que quedan $N-r+1$ formas posibles de seleccionar el objeto dado.

Supongamos que se quiere formar todos los grupos posibles de dos letras cada uno, a partir de las letras a, b, c y d, sin que ninguna de ellas se repita. Podemos imaginar que las letras se hallan pegadas en dados y que se juntaran dos dados a la vez. Aplicando la fórmula respectiva tenemos,

$$4\text{P}2 = \frac{\text{N!}}{(\text{N}-\text{r})!} = \frac{4!}{(4-2)!} = \frac{4!}{2!} = \frac{4\times3\times2}{2}$$

= 12 grupos, que estarían formados así:

ab, ba, ac, ca, ad, da, be, cb, bd, db, de, cd

Ejemplos:

a) En cuántas formas puede acomodarse a ocho personas, si existen tres asientos disponibles?

$$8\text{P}3 = \frac{8!}{(8-3)!} = \frac{8\times7\times6\times5\times4\times3\times2}{5\times4\times3\times2} = 336 \textit{ formas}$$

b) Si en una carrera de caballos, participan siete, ¿en cuántas formas pueden tres de ellos terminar primero, segundo y tercero?

$$7\text{P}3 = \frac{7!}{(7-3)!} = \frac{7!}{4!} = 7\times6\times5 = 210$$

c) La Facultad ha contratado a seis nuevos jornaleros para Rumipamba. El Administrador decide ubicar a tres de ellos en tres lugares claves. ¿De cuántas formas puede conseguirse este objeto? O sea, ¿cuántas permutaciones de seis hombres existen, tomados de tres en tres?

$$6P3 = \frac{6!}{(6-3)!} = \frac{6!}{3!} = 6 \times 5 \times 4 = 120$$

d) Evaluar: 1) $5P5$, 2) $8P3$, 3) $10P8$

$$1)\ 5P5 = \frac{5!}{(5-5)!} = 120$$

$$2)\ 8P3 = \frac{8!}{(8-3)!} = 336$$

$$3)\ 10P8 = \frac{10!}{(10-8)!} = 181.320$$

6.7 Combinaciones

La combinación es una selección de r de los N objetos u observaciones, sin tomar en cuenta el orden de arreglo.

$$NCr = \frac{N(N-1)(N-2)...(N-r+1)}{r(r-1)...1} = \frac{N!}{r!(N-r)!}$$

En el caso de las letras a, b, c, d, todas las combinaciones posibles tomadas de dos en dos estarían dadas por,

$$4C2 = \frac{N!}{r!(N-r)!} = \frac{4!}{2!(4-2)!} = \frac{4!}{2!2!}$$

$$= \frac{4 \times 3 \times 2}{2 \times 2} = \frac{24}{4} = 6$$

y serían: ab, ac, ad, bc, bd, cd. Nótese que ab es la misma combinación que ba, pero no la misma permutación.

Ejemplos:

a) ¿En cuántas formas puede formarse un jurado de 12 hombres y mujeres, de un total de 18 personas?

$$18C12 = \frac{18!}{12!(18-12)!} = \frac{18!}{12!(6)!}$$

$$= \frac{18\times17\times16\times15\times14\times13}{6\times5\times4\times3\times2} = 18.564 \textit{ formas}$$

Nótese que se ha simplificado 12! del denominador con la cifra equivalente del numerador.

b) Evaluar: 1) 8C8, 2) 4C2, 3) 7C3

1) $8C8 = \frac{8!}{8!(8-8)!} = \frac{8!}{8!(0)!} = 1$, *por cuanto* $0! = 1$

2) $4C2 = \frac{4!}{2!(4-2)!} = \frac{4\times3\times2}{2\times2} = 6$

3) $7C3 = \frac{7!}{3!(7-3)!} = \frac{7\times6\times5}{3\times2} = 35$

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM CONSULTA SAPIENTIA
FUNDADA EN 1620
QUITO


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 7

DISTRIBUCIONES TEÓRICAS DE FRECUENCIAS

En el Capítulo 5 se trató sobre cuadros y polígonos de frecuencia. En el presente Capítulo se presenta las distribuciones de frecuencia y sus propiedades.

La distribución de frecuencias de una muestra sugiere el tipo de distribución de la población a la que aquella representa. Si el tamaño de la muestra es suficientemente grande, es de esperar que la distribución de frecuencias de la muestra sea una buena aproximación de la distribución de frecuencias de la población. Por ejemplo, si en una encuesta sobre preferencias de consumo de una población de 10.000 habitantes, se entrevista a 50 personas, no se puede esperar que las dos distribuciones de frecuencias sean similares; si, en cambio, el tamaño de la muestra es de 3.000 individuos, las distribuciones de frecuencias, serán similares:

El cálculo de probabilidades en relación con los juegos de azar, ofrece ejemplos de distribuciones de frecuencias teóricas, para ciertas variables discretas. Tal es el caso de las Figs. 7.1 y 7.2.

7.1 La distribución binomial

En un evento binomial –lanzamiento de una moneda, posibilidad de que un hijo sea varón o mujer– existen sólo dos alternativas, en cuyo caso solo una

de ellas puede tener lugar: cara o sello; y varón o mujer, en los dos ejemplos propuestos.

En Estadística, estos eventos se conocen como "éxito" o "fracaso", tal como vimos en el Cap. 6. El procedimiento usual consiste en observar un número dado de eventos y registrar la frecuencia de éxitos.

El muestreo de una distribución binomial es similar a sacar papeles numerados de un ánfora, con reemplazo. Es decir que la probabilidad asociada con un evento, como el de sacar una ficha roja de una bolsa, es constante para cada extracción, en cuyo caso se dice que los eventos son independientes.

Si p es la probabilidad de que un suceso tenga lugar en una sola prueba (probabilidad de éxito) y $q = 1 - p$, la probabilidad de que no suceda (probabilidad de fracaso), la probabilidad de que el evento tenga lugar exactamente X veces en N pruebas (o sea X éxitos y $N - X$ fracasos), está dada por,

$$p(X) = NC \times p^X q^{N-X} = \frac{N!}{X!(N-X)!} p^X q^{N-X}$$

Por ejemplo, preguntamos, ¿cuál es la probabilidad de que en una familia con seis hijos, haya seis mujeres? o, lo que es igual, ¿seis varones? Se supone que la probabilidad de varón o mujer antes del nacimiento es 1/2 y que se trata de eventos independientes, o sea que la probabilidad de varón o mujer permanece constante para cada nacimiento. La probabilidad es entonces,

$$6C6\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^0 = \frac{6!}{6!0!}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^0$$

$$= \frac{1}{64} = 0.0156$$

Es decir que la probabilidad de que una familia tenga seis hijos del mismo sexo es muy baja; en otras palabras, de una a dos familias de cada 100 que tienen seis hijos, podrán contar con seis hijos del mismo sexo.

Ejemplo: La probabilidad de que los padres con ojos azul-cafés, tengan un hijo con ojos azules es de 1/4 (carácter recesivo). Si existen seis hijos en la familia,

cuál es la probabilidad de que por lo menos la mitad de ellos tenga ojos azules? Ya sabemos que se trata de eventos independientes, para los que la probabilidad de éxito, en un solo intento, es de 1/4 y la probabilidad de fracaso es de 3/4; entonces $N = 6$ y $p = 1/4$. Es necesario calcular p (3), p (4), p (5) y p (6), y sumar las probabilidades correspondientes, debido a que ellas equivalen a las maneras mutuamente excluyentes en que pueden ocurrir el evento.

$$p(3) = \frac{6!}{3!3!}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^3 = \frac{540}{4096}$$

$$p(4) = \frac{6!}{4!2!}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^2 = \frac{135}{4096}$$

$$p(5) = \frac{6!}{5!1!}\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^1 = \frac{18}{4096}$$

$$p(6) = \frac{6!}{6!0!}\left(\frac{1}{4}\right)^6\left(\frac{3}{4}\right)^0 = \frac{1}{4096}$$

La probabilidad de obtener los éxitos deseados, se consigue sumando las probabilidades individuales, lo que se representaría por $p(X) \geq 3$.

$$p(X \geq 3) = \frac{694}{4096} = 0.169$$

lo cual demuestra, como en el caso anterior, la baja probabilidad de que este tipo de familia tenga tantos hijos con ojos azules.

Ejemplo: La probabilidad de sacar exactamente dos caras en seis lanzamientos de una moneda:

$$6C2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{6-2} = \frac{6!}{2!4!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4 = \frac{15}{64}$$

Ejemplo: La probabilidad de que, al lanzar una moneda tres veces, aparezcan a) tres caras, b) dos caras y un sello, c) tres sellos.

$$a)\ p(\textit{tres caras}) = 3C3\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^0 = \frac{3!}{3!0!}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = 1\left(\frac{1}{8}\right)1 = \frac{1}{8}$$

$$b)\ p(\textit{dos caras}, 1\ \textit{sello}) = 3C2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = \frac{3!}{2!1!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1$$

$$= 3\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \frac{3}{8}$$

c) $p(1\,cara, 2\,sellos) = $ *por simetría con b)* $= \frac{3}{8}$

d) $p(tres\ sellos) = 3C0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^3 = $ *por simetría con a)* $= \frac{1}{8}$

Ejemplo: Una agencia de publicidad sostiene que 25% de los estudiantes que fuman, fuman Winston. Si obtenemos una muestra al azar de cuatro estudiantes que fuman y aquella afirmación es cierta, cuál es la probabilidad de que por lo menos uno de ellos fume tal marca? Suponiendo que el aserto es verdadero, la distribución del número de estudiantes que fuman Winston en una muestra de cuatro sería:

$$P(X) = \frac{4!}{X!(4-x)!}\left(\frac{1}{4}\right)^X\left(\frac{3}{4}\right)^{(4-X)}$$

donde X= 0, 1, 2, 3, 4

La probabilidad de encontrar por lo menos uno está dada por:

$$P(X \geq 1) = P(1) + P(2) + P(3) + P(4)$$

$$P(1) = \frac{4!}{1!(4-1)!}\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^3 = 4\left(\frac{1}{4}\right)\left(\frac{27}{64}\right) = \frac{108}{256}$$

$$P(2) = \frac{4!}{2!2!}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2 = 6\left(\frac{1}{16}\right)\left(\frac{9}{16}\right) = \frac{54}{256}$$

$$P(3) = \frac{4!}{3!1!}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right) = 4\left(\frac{1}{64}\right)\left(\frac{3}{4}\right) = \frac{12}{256}$$

$$P(4) = \frac{4!}{4!}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^0 = 1\left(\frac{1}{256}\right)1 = \frac{1}{256}$$

Entonces, $P(X \geq 1) = P(1) + P(2) + P(3) + P(4)$

$$= \frac{108}{256} + \frac{54}{256} + \frac{12}{256} + \frac{1}{256} = \frac{175}{256}$$

O, lo que es igual,

$$P(x \geq 1) = 1 - P(0) = 1 - \left[\left(\frac{4!}{0!4!}\right)\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^4\right]$$

$$= 1 - \frac{81}{256} = \frac{175}{256}$$

Ejemplo: La probabilidad de que un estudiante nuevo apruebe el curso es de 0,4; determinar la probabilidad de que, de cinco estudiantes nuevos, a) ninguno, b) uno y c) por lo menos uno apruebe el curso.

$$a)\ P(ninguno) = \frac{5!}{0!5!}(0.4)^0(0.6)^5 = 1 \times 1(0.07776) = 0.08$$

$$b)\ P(uno) = \frac{5!}{1!4!}(0.4)(0.6)^4 = (5)(0.4)(0.1296) = 0.26$$

$$a)\ P(al\ menos\ uno) = 1 - P(ninguno)$$
$$= 1 - 0.08 = 0.92$$

La distribución de frecuencias para un lanzamiento de dos dados, es como sigue:

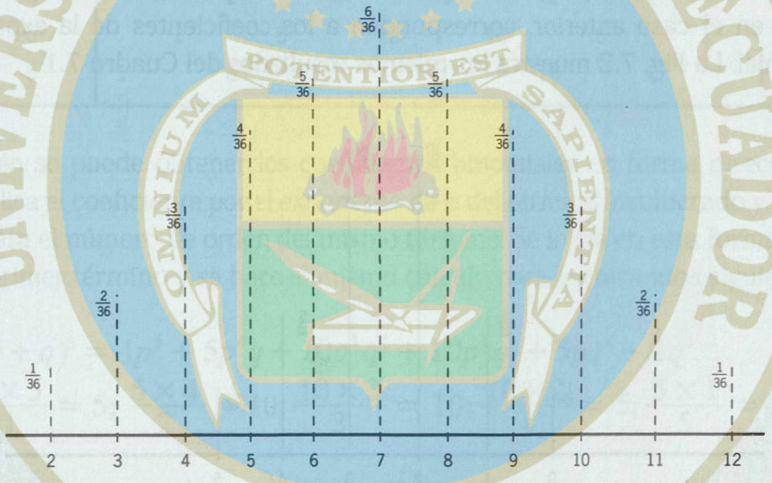


Fig. 7.1 Distribución de frecuencias para un lanzamiento de dos dados.

Para esta distribución binomial, vemos que la probabilidad de que los eventos mutuamente excluyentes 2, 3, 4, 5, 6 y 7 tengan lugar, va aumentando de 1/36 a 2/36, . . ., 6/36, respectivamente. En igual forma, disminuye para los eventos 8, 9, 10, 11 y 12. Los numeradores corresponden a los coeficientes de la expansión binomial, tal como se indica en la próxima Sección.

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO

La distribución de probabilidades para cinco lanzamientos de una moneda, está indicada en el Cuadro 7.1

CUADRO 7.1 Distribución de probabilidades para el binomial
$p = 1/2, N = 5$

Caras (X)	0	1	2	3	4	5
P(X)	1/32	5/32	10/32	10/32	5/32	1/32

Para esta distribución binomial vemos que la probabilidad de que tengan lugar los eventos mutuamente excluyentes (cuya suma es igual a 32/32 = 1), 0, 1 y 2 caras, va aumentando de 1/32 a 5/32 y 10/32; en igual forma, disminuye para los eventos 3, 4 y 5 caras, de 10/32 a 5/32 y 1/32. Los numeradores, como en el caso anterior, corresponden a los coeficientes de la expansión binomial. La Fig. 7.2 muestra los mismos resultados del Cuadro 7.1

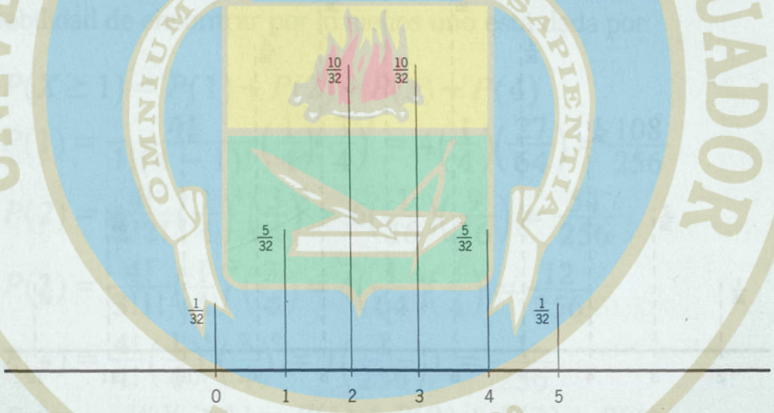


Fig. 7.2 Probabilidad de obtener caras para el binomial $p = 1/2, N = 5$

7.1.1 La expansión binomial

La distribución de probabilidades se llama distribución binomial, puesto que para $X = 0, 1, 2, 3, ..., N$, hay valores sucesivos correspondientes en la expansión binomial:

$$(p+q)^N = p^N + NC1p^{N-1}q + NC2p^{N-2}q^2 + \ldots + q^N$$

donde 1, *NC*1, *NC*2, se llaman coeficientes binomiales. Por ejemplo,

$$(p+q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$$

los coeficientes binomiales están dados por el **Triángulo de Pascal:**

Exponente	Coeficiente binomial
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1
6	1 6 15 20 15 6 1

También se puede obtener los coeficientes binomiales en forma directa: se multiplica el coeficiente por el exponente de p del término involucrado y se divide para el número de orden del mismo término. Se inicia en esta forma desde el primer término y se hace el mismo cálculo para los términos siguientes:

$$(p+q)^5 = 1p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + 1q^5$$

$$\frac{1\times 5}{1} = 5;\ \frac{5\times 4}{2} = 10;\ \frac{10\times 3}{3} = 10;\ \frac{10\times 2}{4} = 5;\ \frac{5\times 1}{5} = 1$$

7.1.2 Algunas propiedades de la distribución binomial

Mediante el uso de algunos procedimientos algebraicos, es posible hacer cálculos para la distribución binomial. Esos cálculos dan lugar a fórmulas para la media, la variancia y la desviación típica, que pueden emplearse en todos los problemas de tipo binomial. Por salir del alcance de este texto, no se describen los procedimientos algebraicos. Las tres propiedades más importantes de la distribución binomial son:

$$\mu = Np$$

$$\sigma^2 = Npq; \quad \sigma = \sqrt{Npq}$$

donde N es el número de observaciones, p la probabilidad de éxito y q la probabilidad de fracaso de un evento.

Ejemplo: Si la probabilidad de que las tuercas producidas por una máquina sean defectuosas es 0.1, encontrar a) la media, b) la desviación típica para la distribución de tuercas defectuosas, en un total de 400.

a) $\overline{x} = Np = 400(0.1) = 40$, es decir que podemos esperar que existan 40 tuercas defectuosas;

b) $\sigma^2 = Npq = 400(0.1)(0.9) = 36$; $\sigma = \sqrt{36} = 6$

7.2 La distribución normal

La distribución binomial de la Sec. 7.1 es la distribución teórica de frecuencias más usada para variables discretas.

En esta Sección se estudiará la distribución de frecuencias para casos de variables continuas, como el caso de altura, peso o ingresos de individuos. La variable pasa sin interrupción de un individuo al siguiente, una variable continua, sin límite con respecto al número de individuos con diferentes medidas.

Muchos fenómenos biológicos proporcionan datos que se distribuyen en forma suficientemente "normal", como para que esta distribución sea la base de la teoría estadística empleada por el investigador. El gráfico de esta distribución, la curva normal, llamada también curva de Laplace o de Gauss, tiene forma acampanada, siendo el tamaño de σ^2, el que determina qué tan aplanada o alargada sea la curva. La curva de la Fig. 7.3 es la gráfica de una distribución normal. Igual es el caso de la Fig. 2.1, sobre cociente intelectual y de la Fig. 2.2 sobre peso corporal.

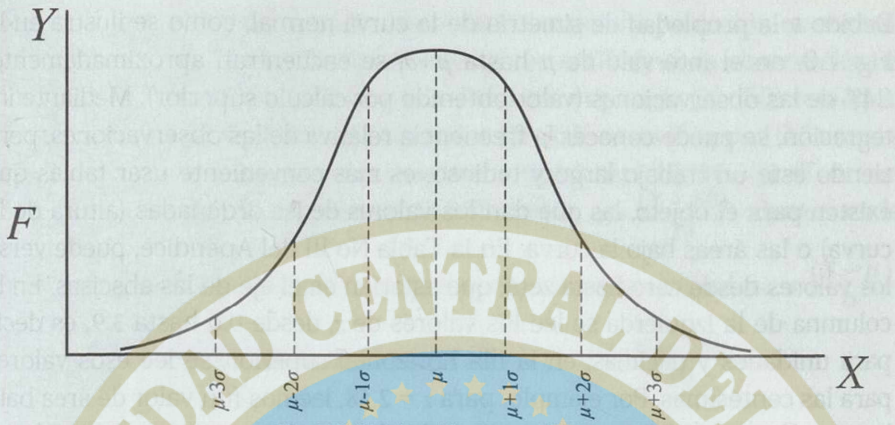


Fig. 7.3 Distribución normal típica.

La localización y forma de la curva normal están dados por los valores de μ y σ. El valor de μ establece el centro de la curva, en forma simétrica, es decir que 50% de las observaciones se hallan a la derecha de μ y 50% a su izquierda; en tanto que el valor de σ determina el grado de dispersión de los datos. En la Fig. 7.4 podemos apreciar tres curvas normales, todas con igual media, pero con diferente variancia (ver Sec. 4.6.3).

La curva **A** es la que menor grado de dispersión presenta; la curva **B** es intermedia y la curva **C** tiene la mayor variancia. En todas las curvas, el área total, o sea el total de las observaciones o individuos, es equivalente a 100% o uno.

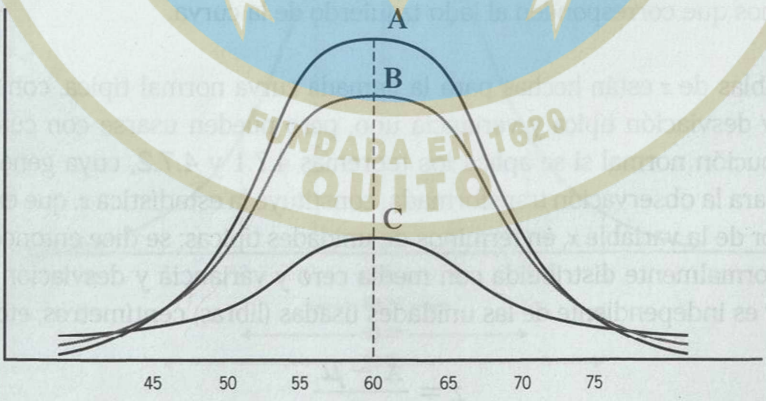


Fig. 7.4 Tres curvas normales con igual μ y diferente σ^2.

Debido a la propiedad de simetría de la curva normal, como se ilustra en la Fig. 7.3, en el intervalo de μ hasta $\mu+\sigma$, se encuentran aproximadamente, 34% de las observaciones (valor obtenido por cálculo superior). Mediante integración, se puede conocer la frecuencia relativa de las observaciones, pero siendo este un trabajo largo y tedioso, es más conveniente usar tablas que existen para el objeto, las que dan los valores de las ordenadas (altura de la curva) o las áreas bajo la curva. En la Tabla No III del Apéndice, puede verse los valores desde cero hasta zeta, que estarían en el eje de las abscisas. En la columna de la izquierda se lee los valores de z, desde 0.0 hasta 3.9, es decir para unidades y décimas; en la fila horizontal superior, se lee esos valores para las centésimas. Por ejemplo, para $z = 2.58$, leemos (un valor de área bajo la curva, que significa porcentaje de individuos u observaciones del total que es 100%), en la siguiente forma: columna izquierda, 2.5, seguimos en forma horizontal, hasta interceptar el valor correspondiente a la columna 8 = 0.4951. Si el valor del área bajo toda la curva equivale a 100% o uno, esto quiere decir que entre $z = 0$ y $z = 2.58$, se halla 49,51% del total de los individuos u observaciones.

Puesto que en el eje de las X se tiene los valores de z, desde μ (z=0), hasta el extremo de la curva $\mu + 3\sigma$ ($z = 3.99$), el valor derecho inferior de la Tabla No. III del Apéndice, muestra la cifra 0.5000, que corresponde justamente a la mitad del área total bajo la curva. Cuando existen valores de z negativos, por la propiedad de simetría de la curva, los valores tabulares son los mismos, pero sabemos que corresponden al lado izquierdo de la curva.

Las tablas de z están hechas para la llamada curva normal típica, con media cero y desviación típica y variancia uno, pero pueden usarse con cualquier distribución normal si se aplica los teoremas 4.7.1 y 4.7.2, cuya generalización para la observación transformada, constituye la estadística z, que expresa el valor de la variable x, en términos de unidades típicas; se dice entonces que z es normalmente distribuida con media cero y variancia y desviación típica uno, y es independiente de las unidades usadas (libras, centímetros, etc.).

$$z = \frac{X - \mu}{\sigma}$$

Ejemplo: Si la distribución de las observaciones de una población es normal, con media = 50 y desviación típica = 10, al restar 50 de cada observación, la media de $(X - 50)$ será igual a cero y la desviación típica de $(X - 50)$ será igual a 10 (Teorema 4.7.1). Ahora, si cada valor $(X - 50)$ se divide por 10, la nueva media será 0/10 = 0 y la nueva desviación típica será 10/10 = 1 (Teorema 4.7.2), de modo que la observación transformada $z = \frac{(X - 50)}{10}$, viene a ser la correspondiente a la fórmula de z, que, como sabemos es igual a $z = \frac{(X - \mu)}{\sigma}$, o sea que el valor de z de la observación 60 será: $z = \frac{60 - 50}{10} = 1$

El valor de z de la observación 45 será: $z = \frac{45 - 50}{10} = -0.5$, que es otra forma de decir que 60 es una desviación típica sobre la media 50 y que 45 es media desviación típica bajo la media 50. En la curva normal típica de la Fig. 7.4, se indica las áreas incluidas entre $z = -1$ y $+1$; $z = -2$ y $+2$; y $z = -3$ y $+3$, que corresponden al 68,27%, 95,45% y 99,73%, respectivamente, del área total que es igual a 100% o uno. De la Fig. 7.4 se desprende que el área bajo la curva termina prácticamente en -3σ y $+3\sigma$. Sin embargo, la ecuación de la curva va de $-\infty$ a $+\infty$.

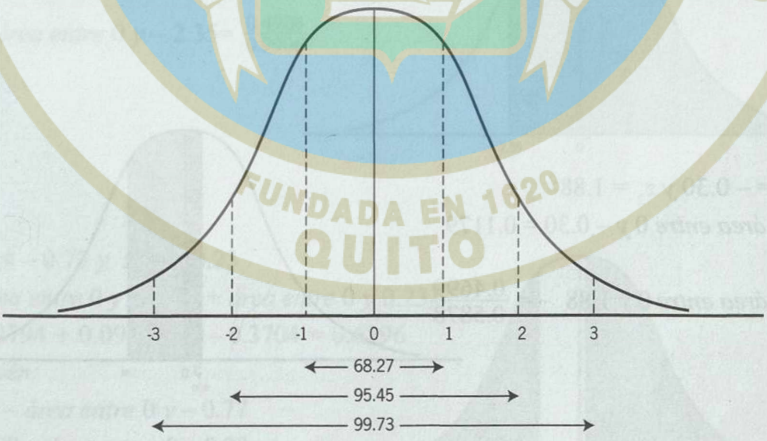
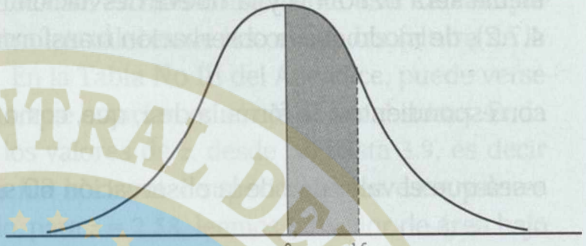


Fig. 7.5 Áreas bajo la curva normal.

Para encontrar el área o porcentaje de observaciones bajo la curva normal (zona sombreada), debe usarse la Tabla No III del Apéndice, en la siguiente forma:

a) $z = 1.6$

área entre 0 *y* 1.6 = 0.4452



b) $z = -0.97$

área entre 0 *y* 0.97 = 0.3340

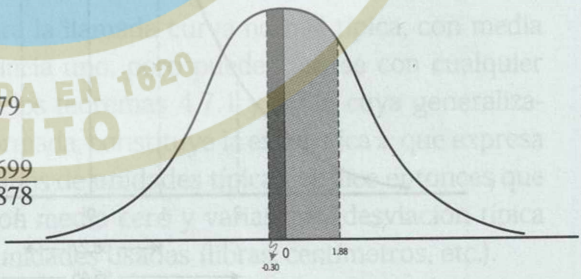


c) $z_1 = -0.30$ *y* $z_2 = 1.88$

área entre 0 *y* − 0.30 = 0.1179

+ *área entre* 0 *y* 1.88 = 0.4699

0.5878

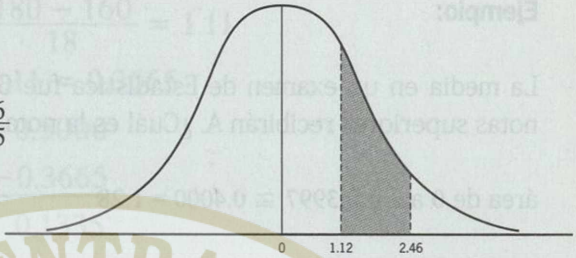


d) $z_1 = 1.12$ *y* $z_2 = 2.46$

área entre 0 *y* 2.46 = 0.4931

– *área entre* 0 *y* 1.12 = 0.3686

0.1245

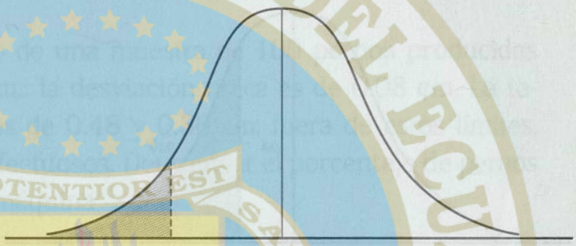


e) $z =< -2.08$

$\frac{1}{2}$ *curva* = 0.5000

– *área entre* 0 *y* – 2.08 = 0.4812

0.0188

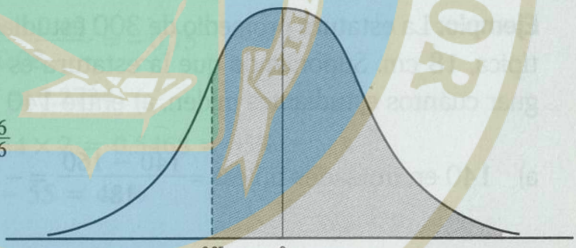


f) $z => -2.35$

$\frac{1}{2}$ *curva* = 0.5000

+ *área entre* 0 *y* – 2.35 = 0.4906

0.9906



g) $z_1 =< -0.77$ *y* $z_2 => 0.23$

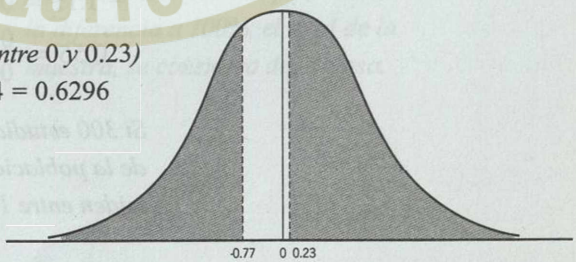
1 – (área entre 0 *y* – 0.77 + *área entre* 0 *y* 0.23*)*

*1 – (*0.2794 + 0.0910*)* = 1– 0.3704 = 0.6296

o también:

0.5000 – *área entre* 0 *y* – 0.77

+ 0.5000 – *área entre* 0 *y* 0.23



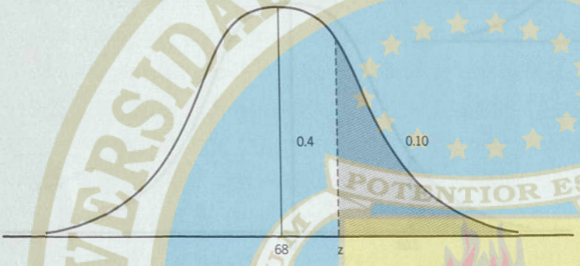
UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO

Ejemplo:

La media en un examen de Estadística fue 68/100; s = 12. Sólo el 10% de notas superiores recibirán A. ¿Cuál es la nota mínima para obtener A?

área de 0 a z = 0.3997 $\cong$ 0.4000 = 1.28

En el cuerpo de la Tabla de z podemos ver que la cifra más próxima a 0.4000, es 0.3997.



$$z = \frac{X - \overline{X}}{s};\ 1.28 = \frac{X - 68}{12}$$

$$(1.28)(12) = X - 68$$

$$X = (1.28)(12) + 68$$

$$X = 83.36$$

La nota mínima para obtener A es 83.4

Ejemplo: La estatura promedio de 300 estudiantes es de 160 cm; la desviación típica, 18 cm. Suponiendo que la estatura es normalmente distribuida, averiguar cuántos estudiantes miden, a) entre 140 y 170 cm. y b) más de 180 cm.

a) 140 en unidades típicas $= \frac{140 - 160}{18} = -1.11$

170 en unidades típicas $= \frac{170 - 160}{18} = 0.55$

$-1.11 = 0.3665$

$0.55 = \frac{0.2088}{0.5753}$

Si 300 estudiantes constituyen 100% de la población, 57%, es decir 171 individuos, miden entre 140 y 170 cm.

b) 180 en unidades típicas = $\frac{180 - 160}{18} = 1.11$

$$1.11 = 0.3665$$

$$0.5000$$

$$-0.3665$$

$$0.1335$$

De 300 estudiantes, 13.35%, es decir 40, miden más de 180 cm.

Ejemplo: El diámetro promedio de una muestra de 100 pernos producidos por una máquina es de 0.42 cm.; la desviación típica es de 0.08 cm. La tolerancia máxima en diámetro es de 0.48 y 0.36 cm; fuera de estos límites, los pernos son considerados defectuosos. Determinar el porcentaje de pernos defectuosos producidos por la máquina

0.48 en unidades típicas = $\frac{0.48 - 0.42}{0.08} = 0.75$

0.36 en unidades típicas = $\frac{0.36 - 0.42}{0.08} = -0.75$

$$0.75 = 0.2734$$

$$0.2734 \times 2 = 0.5468 = 55\%$$

$$100 - 55 = 45\%$$

Es decir que 55% de los pernos se ajustan a la tolerancia permitida; la diferencia a 100%, el total de la muestra, se considera defectuosa.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM
SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 8

PRUEBAS DE HIPÓTESIS

Se define a la hipótesis como una teoría tentativa o una suposición adoptada provisionalmente para explicar ciertos hechos y guiar la investigación de otros.

Toda comprobación de hipótesis, sea estadística o de otro tipo, incluye los siguientes aspectos:

a) Enunciación de la hipótesis. Podemos decir, por ejemplo, "los rendimientos de maíz en la zona de Puembo, son bajos".

b) Deducción de alguna o algunas teorías: "los rendimientos de maíz son bajos debido a la baja fertilidad de los suelos".

c) La verificación de estas teorías, en el campo de las ciencias experimentales, verificación experimental. En el ejemplo propuesto, se diseñaría un ensayo que incluya varios fertilizantes y, quizás, dosis de los mismos, frente a parcelas testigo.

d) Conclusiones en base a los resultados experimentales. Después de analizarlos, se podrá determinar qué elemento y en qué cantidad, debe aplicarse al suelo, para elevar los rendimientos de maíz.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST SCIENTIA QUAM TEMERITAS
FUNDADA EN 1620
QUITO

Cuando una persona se halla investigando sobre la naturaleza de un fenómeno cualquiera, consideraciones teóricas o el estudio de datos empíricos, le obligan a proponer una teoría, la que puede proporcionar un modelo con relación al fenómeno en estudio. Para que esa teoría tenga valor, debe estar de acuerdo con lo que en realidad sucede en la práctica; de otro modo, ella debe ser revisada o descartada.

Para decidir si es que la falta de acuerdo entre valores observados y teóricos es suficiente para rechazar la teoría propuesta, podemos usar una prueba estadística: la hipótesis. Podemos suponer que la teoría puede ser expresada como una hipótesis estadística, es decir aquella que especifica valores para uno o varios parámetros de una población teórica, de la que suponemos constituyen una muestra los datos tomados por nosotros. Normalmente, cuando se plantea una hipótesis, existe más de una alternativa por la que podemos optar, tal como se ilustra seguidamente.

Supongamos que la vida promedio de cierta marca de baterías X, es de 30 meses, con una desviación típica de cuatro meses. Si decidimos probar una nueva marca Y, cuyo precio es más bajo, vamos a esperar que la duración promedio de la segunda marca sea, por lo menos, igual a la que hemos venido usando; de otro modo no valdría la pena cambiar la marca conocida. Si se encuentra que la vida útil de Y es 24 meses, con una desviación típica de cinco meses, podemos plantear el problema de la siguiente forma:

$$H_0: \mu = 30$$
$$H_1: \mu < 30$$

Se plantea H_1 en la manera indicada, puesto que si la calidad de Y no es igual a la de X, entonces la marca Y sería inferior a X.

Supongamos ahora que el promedio de notas de un examen de ingreso del año anterior es de 8.5/10. El instructor quiere probar si la nueva clase tiene un promedio igual al del año pasado. Puesto que desconoce a los nuevos estudiantes, no tiene razón para creer que éstos serán mejores o peores que los anteriores. En consecuencia, el instructor plantea la hipótesis en la siguiente forma:

$$H_0: \mu = 8.5$$
$$H_1: \mu \neq 8.5$$

En ciertos casos, la hipótesis alternativa (H_1) podría ser también,

$$\mu_1 < \mu_2;\ \mu_1 > \mu_2 \text{ o } \mu_1 = \mu_2$$

8.1 La hipótesis nula, H_0

De ilustraciones anteriores se desprende que no existe diferencia entre rechazar la hipótesis de que las marcas de baterías, o la capacidad de dos promociones de estudiantes, sean igualmente buenas o que se acepte la hipótesis de que sus cualidades son diferentes. Teóricamente, sin embargo, es preferible plantear la hipótesis en el sentido de que las marcas son iguales en calidad, en lugar de la hipótesis de que son diferentes. Así se establece la hipótesis,

$$H_0: \mu_1 - \mu_2 = 0, \text{ o lo que es igual,}$$
$$H_0: \mu_1 = \mu_2$$

A este tipo de hipótesis se le conoce como **hipótesis nula**, debido a que supone que no existe diferencia entre tratamientos.

En la experimentación de campo, por ejemplo, el investigador plantea subconscientemente la hipótesis nula de que no existen diferencias entre las variedades o los tratamientos que va a aplicar en sus parcelas, con la esperanza de poder rechazarla y, después de tener la evidencia experimental, aceptar la hipótesis alternativa de que sí existen diferencias entre los tratamientos.

Es decir que se ha planteado una hipótesis, la hipótesis nula, con el objeto de utilizarla como un medio de trabajo. Sin embargo, la hipótesis nula no siempre plantea la no diferencia entre tratamientos. La hipótesis puede sostener que una variedad de trigo va a rendir una tonelada por hectárea, más que la variedad local. En este caso, se calcularía la estadística *t*, para lo que es necesario fijar el valor específico de μ, tal como se describe en el Cap. 10.

8.2 Prueba de hipótesis y nivel de significación

Al momento de plantear la hipótesis nula, se escoge el nivel de probabilidad o nivel de significación. Este nivel de significación teórico está dado por valores tabulares como los que constan en la Tabla No V del Apéndice, de acuerdo con el número de grados de libertad. Si el valor de t calculado es mayor que el valor tabular con el que se le compara, rechazamos la hipótesis nula y se acepta la alternativa; es decir se concluye que la hipótesis nula es falsa y la hipótesis alternativa verdadera. Si se escoge el nivel de significación de 0.05, se rechaza la hipótesis nula afirmando que existen diferencias significativas al 5%. Los procedimientos que nos permiten aceptar o rechazar hipótesis, se llaman pruebas de hipótesis o pruebas de significación.

Por ejemplo, antes de lanzar una moneda, suponemos que la moneda no está "cargada" y esperamos 50% de sellos y 50% de caras. Si al lanzar esa moneda 100 veces, registramos 75 caras y 25 sellos, rechazaríamos la hipótesis de que la moneda es correcta.

Generalmente, los investigadores usan los niveles de significación de 0.05 o 0.01, aun cuando puede usarse otros niveles, de acuerdo con la naturaleza del problema.

El nivel de **0.05, 5%** o **95%,** quiere decir que, de repetirse un evento 100 veces, existen 95 posibilidades de que obtengamos los mismos resultados iniciales; en otras palabras, existen 95 posibilidades entre 100, de que estemos en lo cierto y 5 de estar equivocados.

El nivel de **0.01, 1%** o **99%,** quiere decir que existe una posibilidad entre 100 de estar equivocados al tomar una decisión cualquiera. Es decir que, de acuerdo con la Fig. 8.1, si la hipótesis nula es correcta, el valor de z estará comprendido entre –2.58 y + 2.58, fuera de las zonas críticas o de rechazo de H_0.

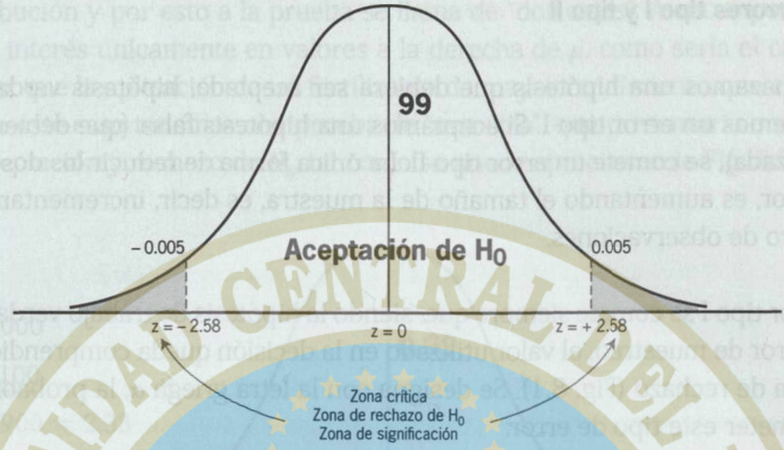


Fig. 8.1 Zonas de aceptación y rechazo de H_0, al 1% de significación. Prueba de dos "colas".

El valor de 2.58 viene dado por la resta de 0.5 (mitad de la curva) de 0.005 (mitad de 0.01, que es el nivel de probabilidad escogido):

$$\begin{array}{r} 0.5000 \\ \underline{-0.0050} \\ 0.4950 = 2.58 \end{array}$$

el valor de 0.4950 consta en el cuerpo de la Tabla de *z*; buscando hacia la columna de la izquierda, encontramos el valor 2.5 y hacia arriba, en la fila superior de la Tabla, vemos .08, con lo que el valor correspondiente de *z* es 2.58.

Si el nivel escogido hubiera sido el 5%, se reemplaza 2.58 por 1.96:

$$\begin{array}{r} 0.5000 \\ \underline{-0.0250} \\ 0.4750 = 1.96 \end{array}$$

La significación al 5% se indica poniendo un asterisco junto al valor calculado y dos asteriscos cuando la significación es al 1%; en este último caso se habla de un valor **altamente significativo**. La no significación se denota por **NS**.

8.3 Errores tipo I y tipo II

Si rechazamos una hipótesis que debiera ser aceptada, hipótesis verdadera, cometemos un error tipo I. Si aceptamos una hipótesis falsa (que debiera ser hechazada), se comete un error tipo II. La única forma de reducir los dos tipos de error, es aumentando el tamaño de la muestra, es decir, incrementando el número de observaciones.

El error tipo I se comete siempre que, siendo la hipótesis de trabajo verdadera, por error de muestreo, el valor utilizado en la decisión queda comprendido en la zona de rechazo (Fig. 8.1). Se designa con la letra griega α, la probabilidad de cometer este tipo de error.

Así mismo, se comete el error tipo II, siempre que, siendo falsa la hipótesis de trabajo, por error de muestreo, el valor utilizado en la decisión queda comprendido en la zona de aceptación (Fig. 8.1). La probabilidad de cometer este tipo de error, se designa con la letra griega β. Seguidamente, se presenta un resumen de las características de los dos tipos de error.

CUADRO No 8.1 Tipos de error I y II

Decisión	SITUACIÓN VERDADERA		Símbolo
	La H_0 es cierta	La H_0 es falsa	
Aceptación de la H_0	No hay error	Error tipo II	β
Rechazo de la H_0	Error tipo I	No hay error	α

8.4 Potencia de la prueba

Está estrechamente relacionada con la probabilidad de cometer un error tipo II. Es decir se refiere a la probabilidad de aceptar la hipótesis alternativa cuando la hipótesis nula es verdadera. La potencia de la prueba se denota por $1-\beta$.

8.5 Pruebas de una y de dos colas

En la Sec. 8.2 se ha presentado valores de z a la derecha e izquierda de μ. Es decir que, los valores pueden hallarse en los dos extremos o "colas" de la

distribución y por esto a la prueba se llama de "dos colas". A veces podemos tener interés únicamente en valores a la derecha de μ, como sería el caso de probar que la aplicación de un fertilizante da mayor rendimiento que el testigo. En este caso tenemos una prueba de "una cola", que presentaría una zona de aceptación y rechazo de H_0, tal como se puede apreciar en la Fig. 8.2.



Fig. 8.2 Zonas de aceptación y rechazo de H_0 al 1% de significación. Prueba de una "cola".

Al comparar la Fig. 8.1 con la 8.2, podemos concluir que las dos áreas sombreadas de la primera, son equivalentes al área sombreada de la Fig. 8.2.

Ejemplo: El fabricante de una medicina sostiene que el producto tiene 80% de efectividad para el control de una enfermedad. En una muestra de 300 individuos que tenían la enfermedad, la medicina alivió a 220 personas. Determinar si la afirmación del fabricante es cierta al 5% de probabilidades.

$$H_0: p = 0.8 \text{ y la afirmación es verdadera}$$
$$H_1: p < 0.8 \text{ y la afirmación es falsa}$$

Escogemos una prueba de una cola, puesto que nos interesa solo valores < 0.8, es decir, inferiores a 80% de efectividad del producto.

$$0.5000$$
$$\underline{-0.0500}$$
$$0.4500 = z = -1.65$$

De acuerdo con la Sec. 7.1.2:

$$\overline{X} = Np = 300(0.8) = 240$$
$$s = \sqrt{Npq} = \sqrt{300(0.8)(0.2)} = \sqrt{48} = 6.92$$
$$z = \frac{220 - 240}{6.92} = -2.89$$

El valor (absoluto) de z calculada es mayor que el tabular, es decir aquel cae en la zona de rechazo de H_0, por lo que aceptamos la hipótesis alternativa y concluimos que la afirmación del fabricante es falsa.

Ejemplo: En una muestra de 200 semillas, germinaron 172; al nivel del 1%, podemos rechazar la afirmación de que el poder germinativo de la semilla es, por lo menos, del 90%?

$$H_0 : p \geq 0.9$$
$$H_1 : p < 0.9$$

En este caso, escogemos una prueba de una cola –del lado izquierdo, negativo– puesto que nos interesa solo valores < 0.9

$$\overline{X} = 200(0.9) = 180$$
$$s = \sqrt{Npq} = \sqrt{(200)(0.9)(0.1)} = \sqrt{18} = 4.24$$
$$z = \frac{172 - 180}{4.24} = -1.88$$

$$0.5000$$
$$\underline{-0.0100}$$
$$0.4900 = z = -2.33$$

El valor absoluto de z calculada es menor que el tabular, por lo que aceptamos H_0: $p \geq 0.9$ y concluimos que el poder germinativo de la semilla llega, por lo menos al 90%.

En el caso de distribuciones binomiales, como hemos visto en los ejemplos anteriores, la estadística

$z = \dfrac{X - \mu}{\sigma} = \dfrac{X - Np}{\sqrt{Npq}}$; si dividimos numerador y denominador por N, tenemos que $z = \dfrac{X - p}{\sqrt{pq/N}}$

Hay situaciones en las que podemos trabajar con X o con valores fraccionales X/N, como en el siguiente ejemplo:

Nos dan un par de dados con el objeto de establecer si están o no cargados, al nivel del 5%. Si decidimos probar la H_0 de que la probabilidad de obtener 7 es igual a 1/6, H_0: $p = 1/6$; la H_1 sería $p \neq 1/6$; en este caso usaríamos una prueba de dos colas por cuanto la probabilidad puede ser mayor o menor que 1/6. Si tiramos los dados 300 veces y encontramos que el 7 aparece 38 veces:

$$z = \frac{X - p}{\sqrt{pq/N}} = \frac{\frac{X}{N} - p}{\sqrt{pq/N}}$$

$$= \frac{\frac{38}{300} - \frac{1}{6}}{\sqrt{\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{300}\right)}} = \frac{0.126 - 0.166}{0.02} = -2.0$$

Por tratarse de una prueba de dos colas restamos de 0.5 (la mitad de la curva), el valor de 0.025 (estamos trabajando al nivel del 0.05 de probabilidades); la diferencia es de 0.4750, que, en valores de $z = -1.96$. El valor calculado cae fuera de la zona de aceptación de H_0, en cuyo caso rechazamos H_0: $p = 1/6$, al nivel del 5%; sin embargo, no podríamos hacerlo al nivel del 1%. Por esta razón y por cuanto sabemos que la probabilidad de sacar 7 es 1/6, pensaríamos que 300 lanzamientos fue un número insuficiente de pruebas y repetiríamos el ensayo.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 9

LA DISTRIBUCIÓN DE X^2

En el Cap. 7 se trató de casos en los que la población consistía de dos clases de individuos y en los que el muestreo se refería a una o dos poblaciones. En el presente Capítulo vamos a considerar poblaciones clasificadas en más de dos categorías y a muestras extraídas de más de dos poblaciones.

Se define a X^2 (letra griega que se lee **Ji** cuadrado), como la suma de cuadrados de variables independientes, normalmente distribuidas, con media cero y variancia uno.

La distribución de X^2 depende del número de grados de libertad, existiendo una distribución (curva dentro del eje de coordenadas), para cada grado de libertad. En la Tabla de X^2 (Tabla No IV del Apéndice), se da el nivel de probabilidad a la cabeza de la tabla, los grados de libertad en la columna izquierda y los valores de X^2 en el cuerpo de la Tabla. El número de grados de libertad equivale al número de clases que se comparan menos uno.

La mayoría de los problemas en que se usa X^2, implican a la cola derecha de la distribución, al contrario de la distribución de z, en la que podemos estar interesados tanto en una, cualquiera, como en las dos colas de la distribución.

La distribución de X^2 se usa generalmente para la enumeración de datos. Sinembargo, X^2 es una distribución continua que se basa en la distribución normal.

9.1 Cálculo de X^2

Como se ha visto en el Cap. 6, los resultados provenientes de muestreo no siempre coinciden con resultados teóricos, de acuerdo a las reglas de probabilidad. Por ejemplo, aun cuando teóricamente esperaríamos obtener 250 caras y 250 sellos, en 500 lanzamientos de una moneda, llamaría la atención que obtengamos exactamente esa proporción en la práctica.

La prueba de X^2, o **prueba de bondad de ajuste**, nos indica cómo los valores observados se comparan con los valores esperados.

Supongamos que en una muestra dada, esperamos, de acuerdo a la teoría de probabilidades, que una serie de eventos $E_1, E_2, E_3, \ldots E_n$, se produzcan con frecuencias $e_1, e_2, e_3, \ldots, e_n$, llamadas frecuencias esperadas; después de realizado el experimento, observamos que dichos eventos se presentan con frecuencias $o_1, o_2, o_3, \ldots, o_n$, llamadas frecuencias observadas. La primera pregunta que se nos ocurre, en presencia de los resultados, es la siguiente: ¿se hallan éstos de acuerdo con los que esperábamos?; ¿son los resultados confiables?

La medida en que los valores observados y esperados discrepan, está dada por la estadística X^2, la que, en su forma más simple, se expresa así,

$$X^2 = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \ldots + \frac{(o_n - e_n)^2}{e_n}$$

$$= \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

Si el valor de $X^2 = 0$, las frecuencias teóricas y las observadas, se hallarían totalmente de acuerdo. Tal sería el caso al obtener exactamente 250 caras y 250 sellos, en 500 lanzamientos de una moneda. Mientras mayor es el valor de X^2, mayor sería la diferencia entre las frecuencias.

Ejemplo: En 500 lanzamientos de una moneda, se observaron 275 caras y 225 sellos. Probar la hipótesis de que la moneda no está "cargada", al nivel del 5 y del 1% de probabilidades.

$$X^2 = \frac{(275-250)^2}{250} + \frac{(225-250)^2}{250} = 5.0$$

Puesto que el número de categorías es $n = 2$ (caras y sellos), el número de grados de libertad sería 2–1 = 1. Los valores de $X^2_{0.05}$ y $X^2_{0.01}$ para un grado de libertad son 3.84 y 6.63, respectivamente.

El valor calculado de $X^2 = 5.0$ es mayor que 3.84; en consecuencia, podemos afirmar que las frecuencias esperadas difieren significativamente de las frecuencias observadas, y rechazamos la hipótesis de que la moneda no está cargada, al nivel de 5% de probabilidades. Sin embargo, puesto que 5.0 es menor que 6.63, no podemos rechazar la hipótesis al 1% de probabilidades. En este caso, lo correcto sería repetir la prueba.

Ejemplo: Alguien nos proporciona un par de dados, de los que sospechamos que están "cargados". Para probarlo, efectuamos 36 lanzamientos y obtenemos los resultados indicados abajo. ¿Podemos concluir que los dados están "cargados", al 1% de probabilidades?

Resultado	2	3	4	5	6	7	8	9	10	11	12
Frecuencia	1	0	5	0	4	5	6	5	0	10	0

H_0: los dados son correctos

H_1: los dados están "cargados"

La distribución de frecuencias para el lanzamiento de dos dados está indicada abajo; los numeradores de las fracciones que corresponden a cada resultado, serían las frecuencias esperadas, de un total de 36, como corresponde al total de las posibles combinaciones al lanzar dos dados.

Resultado	2	3	4	5	6	7	8	9	10	11	12
Frec. esperada	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
Frec. observada	1	0	5	0	4	5	6	5	0	10	0

$$X^2 = \frac{(1-1)^2}{1} + \frac{(0-2)^2}{2} + \frac{(5-3)^2}{3} + \frac{(0-4)^2}{4} + \frac{(4-5)^2}{5} + \frac{(6-5)^2}{6} + \ldots$$
$$+ \frac{(0-3)^2}{3} + \frac{(10-2)^2}{2} + \frac{(0-1)^2}{1} = 44.1$$

$X^2_{0.01}(10\ g.\ l.) = 23.21$; por cuanto el valor calculado de X^2 es mayor que el valor tabular, rechazamos H_0 y concluimos que los dados están "cargados", lo cual salta a la vista, puesto que sería raro que, de 36 lanzamientos, no salgan una sola vez el 3, 5, 10 o 12.

Ejemplo: En cruzamientos de dos tipos de fréjol, se encontraron, en la segunda generación, cuatro tipos de plantas. En una muestra de hojas de 1.301 plantas se encontró:

773	*verdes*
231	*cloradas*
238	*verdes - rayadas*
59	*doradas - verdes - rayadas*
1.301	

De acuerdo a la teoría mendeliana sobre la herencia, la proporción de tipos de hoja debe ser, 9: 3: 3: 1. La hipótesis nula sería de que los resultados están de acuerdo con la teoría mendeliana. El número de plantas dentro de cada clase, de acuerdo a las frecuencias esperadas, sería:

$$\frac{9}{16}(1.301) = 731.9$$
$$\frac{3}{16}(1.301) = 243.9$$
$$\frac{3}{16}(1.301) = 243.9$$
$$\frac{1}{16}(1.301) = \frac{81.3}{1301.0}$$

$$X^2 = \frac{(773.0 - 731.9)^2}{731.9} + \frac{(231.0 - 243.9)^2}{243.9} + \frac{(238.0 - 243.9)^2}{243.9} + \frac{(59.0 - 81.3)^2}{81.3}$$
$$= 9.25, \textit{ con } 4 - 1 = 3 \textit{ grados de libertad}$$

El valor X^2 para tres grados de libertad es 7,81. Por cuanto el valor calculado 9,25 es mayor que el tabular 7,81, estaríamos dispuestos a rechazar la hipótesis nula y aceptar la alternativa: los resultados obtenidos no concuerdan con la teoría sobre la herencia propuesta. Como en el ejemplo anterior, se justificaría repetir el ensayo.

Ejemplo: La relación hombres–mujeres en un pequeño pueblo de Escocia fue, en 1940, de 7 a 1, en favor de los hombres; es decir, había 7 mujeres por cada hombre. Supongamos que en una encuesta reciente se encontró que, en una muestra tomada al azar, formada por 300 personas, había 72 hombres. ¿Podríamos afirmar que la relación 7:1 ha cambiado?

Si la relación es de 7 a 1, una de cada ocho personas es hombre; en consecuencia, si la relación no ha cambiado, la proporción de hombres es $P = 1/8$.

Vamos a probar, entonces:

H_0: $p = 1/8$, *contra la hipótesis alternativa de dos colas*
H_1: $p \neq 1/8$, *al 5% de probabilidades*

Si X es el número de hombres:

$X = 72$, $N - X = 300 - 72 = 228$ *mujeres*

La frecuencia separada sería:

Hombres $= 1/8\ (300) = 37{,}5$
Mujeres $= 7/8\ (300) = 262{,}5$

$$X^2 = \frac{(72.0 - 37.5)^2}{37.5} + \frac{(228.0 - 262.5)^2}{262.5} = 31.74 + 4.53 = 36.27$$

En la Tabla No IV del Apéndice, vemos que el valor de $X^2_{0.05}(1\ g.\ l.) = 3.84$; por cuanto el valor calculado es mayor que el tabular, rechazaríamos H_0 y concluiríamos que, desafortunadamente, la relación ha cambiado.

9.2 Cuadros con dos criterios de clasificación

Llamados también cuadros de doble entrada, en los que las filas representan un criterio de clasificación y las columnas el otro. Para ilustrar el procedimiento, vamos a usar los siguientes ejemplos:

Para una prueba con pasta dentífrica, se dividió a cierto número de niños en tres grupos. Cada uno de ellos usó una marca distinta durante un año. Los resultados fueron:

Marca de pasta	Cantidad de caries nuevas			Total Filas
	0 – 1	2 – 3	4+	
A	40	10	10	60
B	50	20	20	90
C	60	10	20	90
Total columnas	150	40	50	240

¿Es cierta la afirmación que hacen los fabricantes de A, en el sentido de que hay "significativas diferencias" entre su producto y los otros, en cuanto a su capacidad para reducir el número de caries?

La hipótesis planteada sería:

H_0: *Las tres marcas de pasta son iguales.*
H_1: *La marca A disminuyó significativamente el número de caries, con relación a B y C.*

El cuadro presenta los valores observados. Para calcular los valores esperados, teóricos, procedemos de la siguiente forma, para llenar las celdas, dentro de cada columna.

$$e_{11} = \frac{(150)(160)}{240} = 37.5; \quad e_{21} = \frac{(150)(90)}{240} = 56.25; \quad e_{31} = \frac{(150)(160)}{240} = 56.25$$

$$e_{12} = \frac{(40)(60)}{240} = 10.0; \quad e_{22} = \frac{(40)(90)}{240} = 15.0; \quad e_{32} = \frac{(40)(90)}{240} = 15.0$$

$$e_{13} = \frac{(50)(60)}{240} = 12.5; \quad e_{23} = \frac{(50)(90)}{240} = 18.75; \quad e_{33} = \frac{(50)(90)}{240} = 18.75$$

El cuadro final de valores observados y esperados (entre paréntesis), sería como sigue:

Marca de pasta	Cantidad de caries nuevas		
	0 – 1	2 – 3	4+
A	40	10	10
	(37,5)	(10.0)	(12,5)
B	50	20	20
	(56,25)	(15,0)	(18.75)
C	60	10	20
	(56,25)	(15.0)	(18,75)

$$X^2 = \frac{(40.0-37.5)^2}{37.5} + \frac{(50.0-56.25)^2}{56.25} + \frac{(60.0-56.25)^2}{56.25} + \frac{(10.0-10.0)^2}{10.0} + \frac{(20.0-15.0)^2}{15.0} + \frac{(10.0-15.0)^2}{15.0} + \frac{(10.0-12.5)^2}{12.5} + \frac{(20.0-18.75)^2}{18.75} + \frac{(20.0-18.75)^2}{18.75} = 8.76$$

Los grados de libertad se calculan así:

$$g.l. = (filas - 1)(columnas - 1) = (3-1)(3-1) = 4$$

El valor tabular de $X^2_{0.05}\,(4\ g.l.) = 9{,}49$

$X^2_{0.01}\,(4\ g.l.) = 13{,}28$

Por cuanto el valor calculado de $X^2 = 8{,}76$ es menor que los valores tabulares al 5 al 1% de probabilidad, aceptamos la hipótesis nula y concluimos que las tres marcas de pasta dentífrica son iguales y, en consecuencia, es falsa la afirmación de los fabricantes de A, en el sentido de que su producto es superior a los otros. Los fabricantes de A pudieron sostener que, en total, su producto disminuyó 30 caries con relación a B y C o, en otras palabras, que fue 33% más efectivo; pero, como hemos visto, la diferencia no fue estadísticamente significativa.

Ejemplo: Para probar la efectividad de una vacuna, se administró el producto a 170 conejos; un grupo formado por 220, no recibió la vacuna, es decir, cons-

tituyó el lote testigo. Los 390 animales fueron infectados artificialmente con el agente causal de la enfermedad. Del grupo de animales vacunados, murieron ocho, en tanto que del testigo, murieron 19. ¿Podemos concluir que la vacuna fue efectiva? La hipótesis nula sería:

H_0: *La tasa de mortalidad entre los dos grupos es igual; o, lo que es lo mismo, la vacuna no fue efectiva.*

H_1: *La tasa de mortalidad para el grupo de animales vacunados es menor; es decir, la vacuna fue efectiva.*

El cuadro de doble entrada para los valores observados sería:

	Muertos	Vivos	Total
Vacunados	8	162	170
No vacunados	19	201	220
Total	27	363	390

Los valores esperados son:

$$e_{11} = \frac{(27)(170)}{390} = 11.76; \quad e_{21} = \frac{(27)(220)}{390} = 15.23$$

$$e_{12} = \frac{(363)(170)}{390} = 158.23; \quad e_{22} = \frac{(363)(220)}{390} = 204.76$$

El cuadro de valores observados y esperados (entre paréntesis) es:

	Muertos	Vivos
Vacunados	8	162
	(11,76)	(158,23)
No vacunados	19	201
	(15,23)	(204,76)

$$X^2 = \frac{(8.0 - 11.76)^2}{11.76} + \frac{(19.0 - 15.23)^2}{15.23} + \frac{(162.0 - 158.23)^2}{158.23} + \frac{(201.0 - 204.76)^2}{204.76} = 2.29$$

Los grados de libertad, como en el caso anterior:

$$g.l. = (filas - 1)(columnas - 1) = (2 - 1)(2 - 1) = 1$$

$X^2_{0.05}$ *(1 g.l.)* =3.84. Por cuanto el valor calculado es menor que el valor tabular, no podemos rechazar H_0 y concluimos que la vacuna no fue efectiva en cuanto a reducir la tasa de mortalidad.

Los ejemplos precedentes sirven para demostrar que, en estudios de este tipo, no se debe sacar conclusiones rápidas, previas al análisis estadístico. Es común que la prensa o revistas no científicas presenten resultados similares en forma de porcentajes y que podrían no estar basados en un análisis matemático. Sería precipitado concluir de los valores observados del ejemplo anterior, que el porcentaje de muertes dentro del lote de animales vacunados fue de 42%, con respecto al lote de animales no vacunados. No sólo es de interés conocer el número de muertes en cada grupo, sino que se debe tomar en cuenta el tamaño de la muestra, puesto que los números observados nos dicen algo cuando están relacionados con el numero total de individuos dentro del experimento. Supongamos que en un ensayo similar, el número de muertes permanecía igual, pero que el total de animales dentro de cada grupo era de 30. En este caso, X^2 sería igual a 8,14, mayor que el valor tabular de 3.84, en cuyo caso habríamos rechazado la hipótesis nula y habríamos concluido que la vacuna resultó efectiva. No es lo mismo comparar ocho muertes de entre 170 animales, que ocho muertes de entre 30.

9.3 Corrección de Yates

Esta corrección se usa generalmente, cuando el número de grados de libertad es uno y está dada por,

$$X^2 corregida = \frac{[(o_1 - e_1) - 0.5]}{e_1} + \frac{[(o_2 - e_2) - 0.5]}{e_2} + \ldots + \frac{[(o_i - e_i) - 0.51]}{e_i}$$

De la comparación de los valores de X^2, corregido y no corregido, puede sacarse la conclusión pertinente: en caso de que los dos valores sugieran conclusiones diferentes, la solución estaría en aumentar el número de observaciones, es decir, incrementar el tamaño de la muestra.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 10

LA DISTRIBUCIÓN DE *t*

Como en el caso de X^2, *t* presenta una distribución diferente, de acuerdo al número de grados de libertad. En la Tabla No V del Apéndice podemos ver que, para una muestra de 14 pares de observaciones, por ejemplo, ($n = 14 - 1 = 13$), en la columna de grados de libertad 13 y al nivel de probabilidades del 1% (0.01), en el cuerpo de la Tabla encontramos el valor 3.012, el mismo que debe ser sobrepasado por el valor de *t* calculado, en cuyo caso decimos que los resultados son significativos al 1%. Si el valor calculado fuera menor que 3.012, aceptaríamos la hipótesis nula de que los resultados no son significativos.

Conforme aumenta el número de grados de libertad, la distribución de *t* se aproxima a la normal. Se ha visto en capítulos anteriores que, para muestras mayores que 30 (30 observaciones), llamadas muestras "grandes", las distribuciones de muestreo eran aproximadamente normales. Para muestras menores que 30 observaciones, llamadas muestras "pequeñas", la aproximación de la distribución de muestreo a la normal, disminuye conforme se reduce el número de observaciones.

Así, William S. Gosset, industrial cervecero y estadístico, pudo apreciar que el uso de s para predecir σ, al calcular valores de *z*, no era confiable para

pequeñas muestras. Este autor ideó la estadística *t*, llamada "**t de Student**", seudónimo usado por Gosset.

Por definición, $t = \dfrac{\overline{X} - \mu}{s\overline{X}}$ En la práctica, para probar la hipótesis nula de que no hay diferencia entre dos juegos de muestras, se aplica,

$$t = \frac{\overline{X_1} - \overline{X_2}}{s(\overline{X_1} - \overline{X_2})} = \frac{\overline{d}}{s\overline{d}}$$

por la que *t* está dada por la diferencia entre medias de tratamientos, dividida por el error típico de la diferencia.

En definitiva, la prueba de *t* resulta un análisis de variancia para dos tratamientos.

10.1 Prueba de *t* para observaciones no pareadas

Tanto para observaciones pareadas como para las no pareadas, el valor $\overline{d}$, reemplaza el de $\overline{X_1} - \overline{X_2}$, proveniente de los dos juegos de muestras.

Para observaciones no pareadas,

$$s\overline{d} = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{s^2\left(\frac{n_1 + n_2}{n_1 n_2}\right)}$$

Ejemplo: El cuadro No 10.1 presenta la ganancia de peso en novillos con y sin suplemento proteínico.

$$H_0: \mu A = \mu B$$
$$H_1: \mu A \neq \mu B$$

CUADRO No 10.1 Ganancia de peso en novillos con y sin suplemento proteínico

A (sin suplemento)	B (con suplemento)
25	48
30	42
28	50
26	38
32	46
35	40
34	

$$\sum X_1 = 210 \qquad 264$$
$$\sum X_i^2 = 6.390 \qquad 11.728$$
$$\frac{(\sum X_i)^2}{n} = 6.300 \qquad 11.616$$
$$\overline{X} = 30 \qquad 44$$

Suma de Cuadrados

$$SC(X_1) = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1}$$
$$= 6.390 - \frac{44.100}{7} = 6.390 - 6.300 = 90$$
$$SC(X_2) = \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}$$
$$= 11.728 - \frac{69696}{6} = 11.728 - 11.616 = 112$$
$$s^2 = \frac{SC(X_1) - SC(X_2)}{(n_1 - 1) + (n_2 - 2)} = \frac{90 + 112}{6 + 5} = \frac{202}{11} = 18.36$$

18.36 es un estimado de la variancia común para los dos grupos; los grados de libertad = $(n_1 - 1) + (n_2 - 1) = 6 + 5 = 11$.

$$s\bar{d} = \sqrt{s^2\left(\frac{n_1 + n_2}{n_1 n_2}\right)} = \sqrt{18.36\left(\frac{42}{7+6}\right)} = \sqrt{59.32}$$

$$= 7.70 \textit{ libras.}$$

$$\textit{El valor de } t_{0.05} \textit{ (11 g. de l.)} = 2.201$$

$$\textit{El valor de } t_{0.01} \textit{ (11 g. de l.)} = 3.106$$

$$t = \frac{\bar{d}}{s\bar{d}} = \frac{30 - 44}{7.70} = \frac{-14}{7.70} = -1.82^{(N.S.)}$$

Es decir, la diferencia no es significativa, puesto que el valor calculado no sobrepasa el valor tabular. En consecuencia, aceptamos la H_0, de que no hay diferencia entre los tratamientos.

10.2 Prueba de *t* para observaciones pareadas

Frecuentemente, las observaciones dentro de un ensayo son pareadas: cuando se va a comparar dos raciones usando dos animales de cada una de 10 camadas de cerdos, asignando los animales de cada camada al azar, uno a cada ración; o en el caso de comparar el contenido de aceite de dos variedades de ajonjolí, sembradas en 14 localidades.

Ejemplo: Los rendimientos de dos variedades de avena, en 10 localidades fueron:

	Variedad A	Variedad B	$A - B = D$	$(A - B)^2 = D^2$
	65	53	12	144
	45	37	8	64
	55	46	9	81
	89	81	8	64
H_0: $\mu A = \mu B$	53	60	−7	49
H_1: $\mu A \neq \mu B$	64	54	10	100
	84	89	−5	25
	86	74	12	144
	52	59	−7	49
	32	29	3	9
	X_i 625	582	43	729

$\overline{X}$ 62.5 58.2 $\overline{d} = 4.3$

$$s^2\overline{d} = \frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)} = \frac{729 - 184.9}{10(9)} = \frac{544.1}{90} = 6.04$$

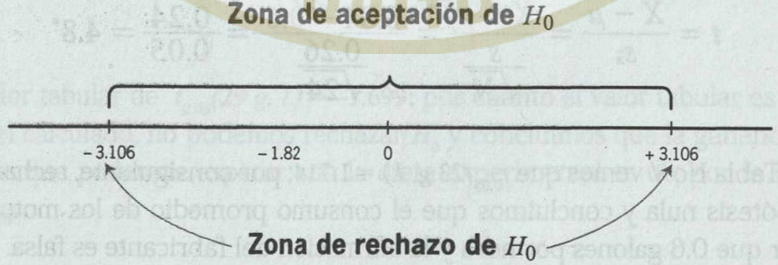
$s^2\overline{d} = 6.04$; $s\overline{d} = \sqrt{6.04} = 2.45$

$g.\,de\,l. = 10 - 1 = 9$; $t_{0.05}(9\ g.\,de\,1) = 2.262$

$$t = \frac{\overline{d}}{s\overline{d}} = \frac{4.30}{2.45} = 1.76^{(N.S.)}$$

Aceptamos H_0: no hay diferencia en rendimientos entre las variedades.

De los ejemplos propuestos anteriormente, se desprende que cuando el valor de *t* calculado es superior, en términos absolutos, al valor tabular, se rechaza la hipótesis nula de que la media del tratamiento A es igual a la de B y optamos por la hipótesis alternativa. Normalmente, no existe una sola hipótesis alternativa; pueden haber varios tipos de ellas, las que pueden ser designadas como alternativas de una o dos colas o lados. El uso de la prueba de una o dos colas está determinado por la forma de plantear el problema o la pregunta dentro del experimento. En los dos ejemplos anteriores, escogimos la prueba de dos colas por cuanto los tratamientos B (con suplemento proteínico y variedad B, respectivamente), pueden ser mayores o menores que los tratamientos A, en cuyo caso se justifica usar la hipótesis alternativa $\mu A \neq \mu B$. Rechazaremos entonces, la hipótesis nula cuando los valores absolutos de t calculados, sean mayores que los tabulares, en los dos lados de la distribución; es decir cuando aquellos se hallen en la región crítica o zona de rechazo, fuera de la zona de aceptación de la hipótesis nula.



En el primer ejemplo, el valor tabular de $t_{0.01}$ para 11 grados de libertad, es de 3.106; el valor calculado de t fue de –1.82, que cae en la zona de aceptación de la hipótesis nula. Por tanto, declaramos los resultados no significativos.

En resumen, para pruebas de dos colas o de dos alternativas, se lee el valor respectivo en la Tabla No V del Apéndice, de arriba hacia abajo, de acuerdo con los grados de libertad y el nivel de probabilidad escogido. Para pruebas de una cola, o de una sola alternativa, es decir cuando tenemos interés solamente en valores extremos a un solo lado de $\overline{X}$, se lee los valores de abajo hacia arriba, así mismo de acuerdo con los grados de libertad y el nivel escogido. En este último caso, si H_0: $\mu A = \mu B$, las alternativas pueden ser:

$$\mu A < \mu B \text{ o } \mu A > \mu B$$

Para ilustrar los casos en que se utilizan pruebas de una o de dos colas, presentamos seguidamente los siguientes ejemplos:

Ejemplo: Un fabricante de motores a gasolina sostiene que, en promedio, ellos gastan no más de 0.6 galones de gasolina por hora. Se probó una muestra de 24 motores y se encontró que el consumo promedio fue de 0.84 galones con $s = 0.26$. Con estos resultados, ¿podemos decir que la afirmación del fabricante es cierta?

$L\,H_0$ sería que la afirmación es correcta, en cuyo caso, H_0 : $\mu = 0.6$. Puesto.que nuestro interés sería rechazarla solamente en caso de que el consumo sea mayor que 0.6 galones, escogeríamos una prueba de una cola, con H_1: $\mu > 0.6$. Si el nivel de probabilidad es del 5%, vamos a comparar el valor calculado de t con el tabular para 23 grados de libertad.

$$t = \frac{\overline{X} - \mu}{s_{\overline{x}}} = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{N}}} = \frac{0.84 - 0.6}{\frac{0.26}{\sqrt{24}}} = \frac{0.24}{0.05} = 4.8^{*}$$

En la Tabla No V vemos que $t_{0.05}(23\ g.\ l.) = 1.714$; por consiguiente, rechazamos la hipótesis nula y concluimos que el consumo promedio de los motores es mayor que 0.6 galones por hora y la afirmación del fabricante es falsa.

Ejemplo: El pago promedio mensual por consumo de luz en la ciudad, el año pasado, fue de $ 320,00; quisiéramos saber si ha habido algún aumento este año. Para probarlo, tomamos una muestra al azar de varias familias y estimamos los pagos realizados por ellas. Vamos a utilizar la hipótesis de que no ha habido cambio, es decir, H_0: $\mu = 320$, contra la alternativa de dos colas, H_1: $\mu \neq 320$, al nivel de probabilidad del 1%.

Optamos por una prueba de dos colas porque el valor de las planillas puede haber aumentado o disminuido.

Tomamos una muestra de 50 familias y encontramos que el gasto promedio es de $ 280, con $s = \$ 70$.

$$t = \frac{\overline{X} - \mu}{s_{\overline{x}}} = \frac{280 - 320}{\frac{70}{\sqrt{50}}} = \frac{-40}{9.9} = -4.04$$

Valor que supera el de $t_{0.005}(49\ g.\ l.) = -2.684$. En consecuencia, rechazamos H_0 y concluimos que ha habido un cambio significativo en el valor de las planillas de luz.

Ejemplo: La ganancia promedio de 30 conejos, a los que se alimentó con una dieta experimental, fue de 275 gramos. $s_{\overline{x}} = 15.6g.$; al nivel del 5%, podemos concluir que la ganancia promedio con esa ración es no menor de 300 g.?

$$H_0 : \mu \geq 300$$
$$H_1 : \mu < 300$$
$$t = \frac{275 - 300}{15.6} = \frac{-25}{15.6} = -1.602$$

El valor tabular de $t_{0.05}(29\ g.\ l.) = -1.699$; por cuanto el valor tabular es mayor que el calculado, no podemos rechazar H_0 y concluimos que la ganancia promedio que podemos esperar con la dieta experimental es de, por lo menos. 300 g.

En los ejemplos propuestos podemos apreciar los siguientes aspectos:

1) La hipótesis nula se presenta siempre como una igualdad, es decir, H_0: $\mu A = \mu B$; la forma como se presenta la hipótesis alternativa depende del tipo de pregunta que plantee el problema:

$$H_1: \mu A \neq \mu B;\ \mu A > \mu B;\ \mu A < \mu B$$

2) Se usará una prueba de una o dos colas, dependiendo de nuestro interés y la forma cómo se plantea la pregunta dentro del problema. Si existe la posibilidad de que el tratamiento o factor B pueda ser tanto mayor como menor (cualquiera de los dos casos) que el tratamiento o factor A, usaremos una prueba de dos colas, planteando H_1: $\mu A \neq \mu B$; en caso de que tengamos interés en determinar si B tiene valores únicamente mayores o únicamente menores que A, es decir valores extremos solamente hacia el lado izquierdo o solamente al lado derecho de la distribución, usaremos una prueba de una cola, planteando H_1: $\mu A > \mu B$ o $\mu A < \mu B$.

3) En el caso de pruebas de una cola, leemos el valor tabular de t, para el nivel de probabilidad establecido, directamente en la Tabla No V, de abajo hacia arriba; para pruebas de dos colas, se lee el valor tabular, de acuerdo al nivel de probabilidad escogido, de arriba hacia abajo.

4) Cuando la diferencia entre dos tratamientos A y B es de tal magnitud que el valor calculado de t cae en una de las dos colas de la distribución (zonas críticas o de rechazo), rechazamos H_0 y optamos por la alternativa H_1.

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

Capítulo 11

CORRELACIÓN

En muchos problemas de investigación, tiene especial interés conocer el grado de relación que existe entre dos o más variables. Por ejemplo, ¿existe alguna relación entre dosis de fertilizante o fungicidas y el rendimiento de un cultivo?; ¿entre el grado de instrucción de una persona y su inteligencia?; entre la edad y la cantidad de cigarrillos que fuma un individuo?

La correlación mide el grado de la intensidad de asociación entre dos variables o el grado en que dos variables cambian una con respecto a la otra.

Los valores de correlación *(r)* varían de –1 a +1, cuando existe una relación perfecta negativa y positiva, respectivamente. Cuando no existe relación alguna entre las variables, $r = 0$. En la práctica, los valores se distribuyen entre los dos extremos y, más bien, debemos dudar de un valor de $r = 1.0$.

El diagrama de dispersión de los datos, o dispersograma, dentro de los ejes *X, Y*, da una idea del grado de asociación entre los variables. A cada valor de *X*, en el eje de las abscisas, corresponde uno de *Y*, en el eje de las ordenadas, ubicándose en esta forma el punto, dentro de los ejes de coordenadas.

En la Fig. 11.1 vemos que, en el primer caso, cuando la correlación es pequeña, r se aproxima a cero, es decir no existe relación entre las variables. Al centro se puede ver una relación positiva estrecha que, teóricamente, es igual a 1 y, en el tercer caso, una relación negativa, en la que $r = -1$.

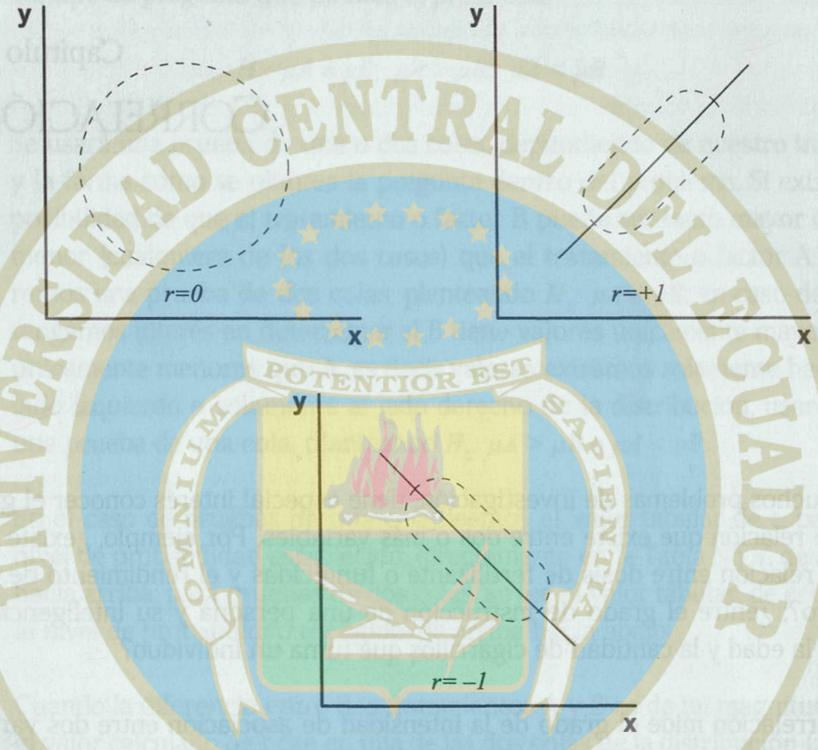


Fig. 11.1 Grado de relación entre dos variables.

A pesar de lo expuesto, puede ser difícil y riesgoso tratar de interpretar visualmente un valor de correlación, puesto que si no se escoge escalas adecuadas para los ejes de las X e Y, la simple observación puede sugerir una correlación falsa o esconder una verdadera. El primer caso sería el de un vendedor que, para impresionar al público, presenta una línea ascendente pronunciada, que representa volumen de ventas, la cual no guarda relación con la proporción real. Para evitar conclusiones falsas, el investigador debe indicar, junto con el diagrama de dispersión de los datos, el valor de la desviación típica y el valor de r.

11.1 Coeficiente de correlación

Cuando se trata de establecer el grado de relación entre dos variables, se calcula el coeficiente de determinación *(r^2 x 100)*, que determina el porcentaje de asociación entre las variables. Es decir que este valor denota el porcentaje de la variación de *Y* atribuible a *X*; el resto de esa variación no se puede explicar.

El coeficiente de indeterminación, está dado justamente por $1 - r^2 = I^2$, que constituye la proporción no explicable de una suma de cuadrados total.

$$r = \sqrt{\frac{\text{variación explicable}}{\text{variación total}}}$$

La raíz cuadrada del coeficiente de determinación es lo que se llama coeficiente de correlación *r*, el cual está dado por,

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{n}\right]}}$$

o sea que r es el cociente de dividir la suma de productos *XY[SP(XY)]*, por la raíz cuadrada de la suma de cuadrados de *X[SC(X)]* por la suma de cuadrados de *Y[(SC(Y)]*.

Supongamos que tenemos dos variables *X, Y,* con seis pares de observaciones correspondientes a dosis de fungicida *(X)* y rendimiento *(Y)*.

(X) Fungicida	(Y) Rendimiento
2	3
5	7
4	3
7	9
3	5
9	10
30	37

$$SC(X) = \sum_{i=1}^{6} X_i^2 - \frac{\left(\sum_{i=1}^{6} X_i\right)^2}{n}$$

$$= 2^2 + 5^2 + \ldots + 9^2 - \frac{(30)^2}{6}$$

$$= 184 - 150 = 34$$

$$SC(Y) = 3^2 + \ldots + 10^2 - \frac{(37)^2}{6}$$

$$= 273 - 228 = 45$$

$$SP(XY) = \sum_{i=1}^{6} XY - \frac{\left(\sum_{i=1}^{6} X\right)\left(\sum_{i=1}^{6} Y\right)}{6}$$

$$= (2\times 3) + (5\times 7) + \ldots + (9\times 10) - \frac{(30)(37)}{6}$$

$$= 221 - 185 = 36$$

$$r = \frac{36}{\sqrt{(34)(45)}} = +0.92^{++}$$

La significación del coeficiente de correlación *r* depende del número de pares de observaciones. Mientras mayor es su número, mayor es la probabilidad de obtener significación, puesto que el número de grados de libertad se basa en *(n – 2)*, en lugar de *(n – 1)*, como en casos y distribuciones anteriores.

El valor de *r* tabular debe ser sobrepasado por el valor calculado, para obtener significación. En nuestro ejemplo, los valores tabulares para 6 – 2 = 4 grados de libertad, a los niveles del 5 y del 1% son, 0.811 y 0.917, respectivamente; por cuanto el valor calculado en nuestro ejemplo, es de 0.92, podemos rechazar la hipótesis nula de que no existe relación entre las variables y aceptar la hipótesis alternativa: existe relación entre ellas, más de lo que pudiera atribuirse al azar.

El coeficiente de determinación viene dado por,

$$r^2(100) = (0.92)^2 \text{ X } 100 = 84.64\%$$

es decir que 85% del aumento en rendimiento, se puede atribuir a la aplicación del fungicida. El resto no se puede explicar.

11.2 Propiedades del coeficiente de correlación

a) *r* es independiente del origen de las variables; indica, más bien, la consistencia del cambio: no importa si comparamos lluvia con toneladas, libras con centímetros, etc.

b) Es independiente de la escala.

c) Es independiente del número de medidas, aun cuando generalmente se considera que son necesarios por lo menos 20 pares para tener confianza en un valor de *r* significativo.

d) Las variables deben ser pareadas por alguna razón lógica.

e) Las muestras deben ser tomadas al azar.

f) Correlación espuria es una condición en que las variables no están causalmente relacionadas entre sí sino que más bien, lo están a una tercera.

g) Se debe tener mucho cuidado al inferir causa y efecto de un coeficiente de correlación significativo.

Con relación al último literal, debe añadirse que, de todas las medidas estadísticas estudiadas, el coeficiente de correlación es la que está más sujeta a ser mal interpretada. Esto se debe a la suposición frecuentemente errónea de que por cuanto dos variables tienen relación, el cambio de una de ellas ocasiona un cambio en la otra. Así, muchas veces podemos usar un coeficiente de correlación significativo, para "probar" una relación de causa y efecto, que puede no existir.

Se ha demostrado que existe una correlación negativa entre el hábito de fumar y la obtención de buenas calificaciones. Si uno no fuma, podría usar este hecho como evidencia de que fumar es nocivo, de que fumar causa o se refleja en la obtención de notas bajas. Esto puede ser cierto, pero los fumadores podrían argüir que el obtener bajas calificaciones causa tensión nerviosa lo cual, a su vez, estimula al estudiante a fumar más. Podría suceder también que fumar y obtener malas notas no estén directamente relacionados entre sí, pero se hallan correlacionados porque los dos hechos tienen relación con un tercer factor, tal vez el grado en que el individuo se dedica a actividades sociales o que no tengan que vez con el estudio. Esto es justamente lo que se conoce como correlación espuria, tal como se anota en el literal f.

Finalmente, hay que anotar que, debido a la variación en las muestras, podemos obtener una correlación significativa aun cuando las variables no estén relacionadas; esto puede ser controlado seleccionando el nivel de significación.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 12

REGRESIÓN

Como se ha visto en el Capítulo precedente, en el campo de la investigación frecuentemente se encuentra dependencia entre dos variables. Así por ejemplo, decimos que existe relación entre la cantidad de fertilizante aplicado y el rendimiento de cultivos; o entre la longitud y diámetro de mazorca y el peso del grano; entre la altura de una persona y su peso, etc. etc.

El coeficiente de correlación determina la intensidad de asociación entre dos variables, pero nada nos dice sobre la gradiente de la línea de asociación, ni tiene capacidad de predicción.

Regresión tiene que ver con la dependencia de la variable *Y*, llamada variable dependiente, de otra variable *X*, la variable independiente. En matemáticas, se conoce a *Y* como función de *X*; en Estadística se describe a esta relación como regresión de *Y* en *X*. Es decir, en el caso de dosis de fertilizante *X*, y rendimiento de trigo, por ejemplo, a una serie de valores $X_1, X_2, X_3, \ldots, X_n$, corresponde otra de $Y_1, Y_2, Y_3, \ldots, Y_n$.

Como en el caso del diagrama de dispersión o dispersograma, en correlación, estos valores se disponen dentro de los dos ejes de coordenadas, y el dispersograma puede sugerir si existe una relación lineal como en la Fig. 12.1a, o si ella es curvilineal (Fig. 12.1b)

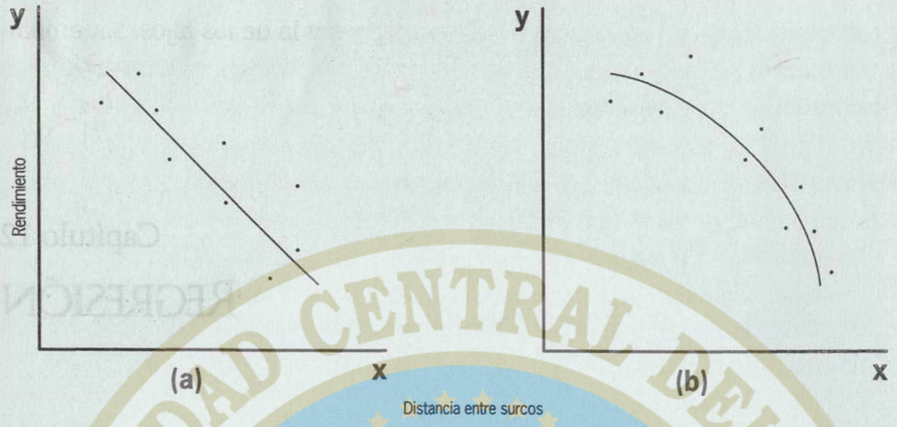


Fig. 12.1 Respuesta lineal y curvilineal.

En Estadística, no nos contentamos con la apreciación visual del dispersograma, sino que tratamos de encontrar ecuaciones que describan dichas curvas de respuesta, lo que se llama ajuste de curvas de respuesta.

Pero, además de determinar la función matemática, que nos diga de qué manera están relacionadas las variables, o sea que nos dé la "mejor" relación funcional entre ellas, es de interés saber con qué precisión se puede predecir el valor de una variable, si se conoce los valores de la otra. Puede tener interés tratar de predecir valores futuros como la producción de frutas, accidentes automovilísticos (compañías de seguro), posible ingreso de estudiantes a la Facultad, etc. O, puede ser posible observar un valor de X e imposible o poco práctico, observar el valor correspondiente de Y.

Sin embargo, es necesario advertir que, simplemente en base de nuestra suposición de que existe dependencia de Y en X, y por haber seguido el procedimiento mecánico de cálculo, no se debe concluir automáticamente que existe relación causal entre X e Y. Sólo el investigador, con conocimiento cabal del material y la forma como se llevó a cabo el experimento, está en capacidad de tomar esa decisión.

Si bien el término regresión no describe, dentro de la Estadística, los usos que tiene, su empleo es de origen histórico. Galtón, en 1886, postuló la teoría de

que existe relación entre la estatura de los padres y la de los hijos. Sin embargo, este autor encontró que la altura promedio de los hijos de matrimonios altos, no era tan grande como la de éstos, ni que los hijos de parejas de baja estatura, eran tan pequeños como sus padres. Galtón concluyó que parecía existir una tendencia del carácter altura, a acercarse a los valores promedios de la población; es decir una regresión hacia los valores promedios.

Resumiendo, podemos decir que los usos de regresión en Estadística, pueden ser:

a) Determinar si la variable dependiente Y, depende de X; es decir, nos da la medida de la proporción de cambio en Y, por una unidad de cambio en X.

b) Conocidos los valores de la variable independiente X, predecir valores de Y, desconocidos en ese momento.

c) Estudiar la forma de la curva de regresión: ¿es lineal?, ¿curvilineal?

d) Algunos investigadores pueden estar interesados en determinar el error en Y, dentro de un experimento, después de hacer los ajustes del efecto que puede tener X en Y.

e) Un investigador puede postular su teoría sobre causa y efecto: para probarla, puede usar varios métodos de regresión.

12.1 Ecuación de la línea recta

Si dentro del dispersograma, las observaciones se sitúan alrededor de la línea recta (respuesta lineal), podemos determinar, por geometría analítica, la ecuación de esa línea recta:

$$\widehat{Y} - \overline{Y} = b(X - \overline{X})$$
$$\widehat{Y} - \overline{Y} = bX - b\overline{X}$$
$$\widehat{Y} = \overline{Y} - b\overline{X} + bX; \quad \overline{Y} - b\overline{X} = a$$
$$\widehat{Y} = a + bX \quad (\textit{relación funcional})$$

donde:

$\widehat{Y}$ = *valor de la ordenada (valor predecido)*

X = *valor de la abscisa (valor conocido)*

$\overline{Y}$ = *media de Y*

$\overline{X}$ = *media de X*

b = *coeficiente de regresión*

La Fig. 12.2 ilustra la forma como se traza la línea recta, en base de su ecuación matemática,

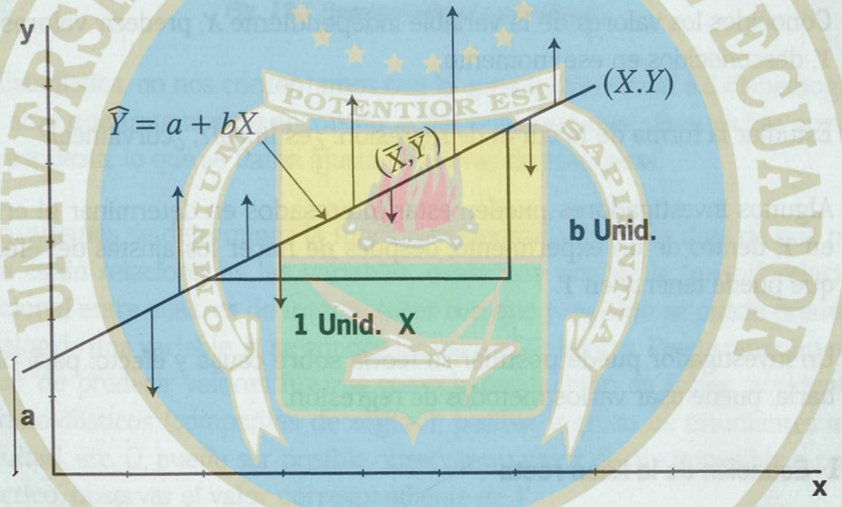


Fig. 12.2 Ecuación y trazado de la línea recta.

De la Fig. 12.2 se desprende que, cuando $X = 0$, $Y = a$, es decir que a es el punto donde la línea recta corta al eje Y, o sea el intercepto en Y. Cuando $a = 0$, la línea pasa por el punto de origen (0, 0).

El cambio de una unidad en el eje de las X, provoca un cambio b en el eje de las Y, de tal manera que b mide la gradiente de la línea de regresión. Cuando

b es positivo, las dos variables aumentan o disminuyen juntas. Tal sería el caso cuando X es kg. de fertilizante por hectárea, e Y, toneladas de grano por hectárea. Cuando b es negativo, mientras el valor de una variable aumenta, el de la otra disminuye: si X es distancia entre surcos, e Y rendimiento, a mayor distancia el rendimiento tiende a disminuir.

12.2 El método de los cuadrados mínimos

Se ha anotado antes, que la cualidad principal de establecer un valor de regresión, consiste en poder predecir resultados de Y, cuando sólo se conoce el de X. Sin embargo, estos parámetros no se pueden determinar sin error, debido a que los valores observados de la variable dependiente, rara vez coinciden con los valores esperados. Esto se demostró claramente en el Cap. 9, al tratar sobre X^2.

Cuando se traza líneas de regresión, es necesario ser objetivo al tratar de ajustar una serie de datos a esas líneas de regresión. Para ello, se debe escoger la línea de "mejor ajuste" de esos datos.

Como se ilustra en la Fig. 12.2, los datos observados se disponen alrededor del eje o línea de regresión. Se dice que la línea de regresión trazada es la que "mejor se ajusta" a los datos. Este criterio de "mejor", se refiere a los cuadrados mínimos, de acuerdo a los cuales la suma de cuadrados de las desviaciones de los puntos observados con respecto a la línea de regresión, debe ser un mínimo. En otras palabras, si se trazara otra línea de regresión, la suma de cuadrados de las desviaciones de los puntos observados, tendría un valor más grande.

12.3 Cálculo del coeficiente de regresión y de la ecuación de regresión

El Cuadro 12.1 presenta, a manera de ilustración, los pesos promedios de cinco camadas de cerdos y el consumo correspondiente de balanceado.

CUADRO 12.1 Pesos iniciales promedios de cinco camadas de cerdos y consumo de balanceado.

Peso corporal (X)	Consumo (Y)	Desviaciones		Cuadrados		Productos xy
		x	y	x^2	y^2	
85	720	+7	+31	49	961	217
75	680	−3	−9	9	81	27
80	700	+2	+11	4	121	22
68	630	−10	−59	100	3481	590
82	715	+4	+26	16	676	104
390	3445	0	0	178	5320	960
$\overline{X}$ 78	689					

Coeficiente de regresión $b = \frac{\sum XY}{\sum X^2} = \frac{960}{178} = 5.39\ lbs$ de balanceado por libra de peso corporal de los cerdos.

A este valor de b, el coeficiente de regresión, se ha llegado aplicando la fórmula,

$$b_{Y \cdot X} = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{(X - \overline{X})^2}, \text{ es decir}$$

$b_{Y \cdot X} = \frac{SP(XY)}{SC(X)}$, la suma de productos cruzados *XY*, dividida por la suma de cuadrados de *X*. La fórmula de trabajo usual es:

$$b_{Y \cdot X} = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}}$$

En el ejemplo propuesto,

$$SP(XY) = (85 \times 720) + (75 \times 680) + \ldots + (82 \times 715) - (390)(3445)/5$$
$$= 209.670 - 268.710 = 960$$
$$SC(X) = 85^2 + 75^2 + \ldots + 82^2 - (390)^2/5$$
$$= 30.598 - 30.420 = 178$$

$b_{Y \cdot X} = \frac{960}{178} = 5.39$, lo que se interpreta diciendo que, para conseguir el aumento de 1 *lb.* de peso de los cerdos, se necesita 5.39 *lbs.* de balanceado.

La ecuación de regresión se escribiría así:

$$\widehat{Y} - \overline{Y} = b(X - \overline{X})$$
$$\widehat{Y} - 689 = 5.39(X - 78)$$
$$\widehat{Y} = 689 + 5.39(X - 78)$$
$$\widehat{Y} = 689 - 420.42 + 5.39X$$
$$\widehat{Y} = 208.58 + 5.39X = a + bX$$

El intercepto en Y, $a = \overline{Y} - b\overline{X}$

$$a = 689 - 5.39(78)$$
$$= 89 - 420.42$$
$$= 268,58$$

La función de predicción puede apreciarse con el siguiente ejemplo: queremos determinar la cantidad de balanceado que consumiría un grupo de cerdos, cuyo peso promedio es de 90 *lbs.* Como puede verse en el Cuadro 12.2, no existe el valor de $X = 90$. Sin embargo, podemos estimar el valor correspondiente Y, aplicando la ecuación: $\widehat{Y} = a + bX$

$$\widehat{Y}(90) = 268.58 + 5.39(90)$$
$$= 268.58 + 485.10$$
$$= 753.68 \textit{ lbs. de balanceado.}$$

12.4 Valores ajustados

Son aquellos que se esperaría si todo los valores de Y tuvieran un valor promedio de X, o algún otro valor fijo. El cuadro 12.2 presenta los valores ajustados de consumo de balanceado, que serían aquellos que esperaríamos si todos los cerdos tuvieran un peso promedio de 78 libras.

CUADRO No 12.2 Valores de regresión y valores ajustados de consumo para los datos del Cuadro No 12.1.

Peso corporal *(X)*	Consumo *(Y)*	Consumo estimado $\widehat{Y}$	Desviación $d_{Y \cdot X} = Y - \widehat{Y}$	$d^2_{Y \cdot X}$	Consumo ajustado $689 + d_{Y \cdot X}$
85	720	726,73	−6,73	45.29	682,27
75	680	672,83	+7,17	51,41	696,17
80	700	699,78	+0,22	0.05	689,22
68	630	635,10	−5,10	26,01	683,90
82	715	710,56	+4.44	19,71	693,44
	3.445	3.445,00	0	142,47	3.445,00

$$\widehat{Y} = a + bX$$

$$= 268.58 + 5.39(85) = 726.73$$

12.5 Fuentes de variación en regresión

Tal como se indica ampliamente en el Capítulo 15, el análisis de variancia es un procedimiento aritmético que consiste en desdoblar la variación total en un juego de datos, de acuerdo con las diferentes fuentes de variación presentes. Para el caso de regresión lineal, la variación total puede desdoblarse en las siguientes fuentes: debido a regresión *(X)* y desviación de regresión (residual).

El análisis de variancia (ADEVA), para los datos del Cuadro 12.1 sería como sigue:

Fuente de Variación	*g. de l.*	*SC simból.*	*g. de l.*	*SC*	*CM*	*F*
X (debido a regres.)	1	$(\sum XY)^2 / \sum X^2$	1	5.177,53	5.177,53	109,02**
Residual (desv. regr.)	$n-2$	Por diferenc.	3	142,47	47,49	
Total	$n-1$	$\sum Y^2$	4	5.320,00		

$$SC(residual) = \sum Y^2 - \frac{(\sum XY)^2}{\sum X^2} = 5320.00 - (960)^2/178$$

$$5320.00 - 5177.53 = 142.47$$

$$SC(X), debido\ a\ regresión = \frac{(\sum XY)^2}{\sum X^2} = \frac{(960)^2}{178} = 5177.53$$

La prueba de F del ADEVA es una prueba de la hipótesis nula de que el coeficiente de regresión poblacional $\beta = 0$, o sea de que la variación en X no contribuye a la variación en Y.

El valor altamente significativo de F para X (debido a regresión), nos permite rechazar la H_0 y afirmar que el peso de los cerdos explica fácilmente la variabilidad en la cantidad de balanceado consumida.

12.6 Desviación y límites de confianza

Un estimado insesgado de la verdadera variancia en regresión, está dado por el cuadrado medio de la desviación de regresión, con n–2 grados de libertad,

$$s^2_{Y \cdot X} = \frac{\sum Y^2 - \dfrac{(\sum XY)^2}{\sum X^2}}{n-2}$$

en nuestro ejemplo, $= \dfrac{5.320.00 - (960)^2/178}{3} = 47.50$

La raíz cuadrada de este valor $s_{Y \cdot X} = 6.89$, se denomina la desviación típica de Y para X fija.

La variancia de un valor estimado $\widehat{Y} = \overline{Y} + bX$, de la media de una población, está dada por la suma de las variancias de $\overline{Y}$ y bX, así,

$$s^2_{Y \cdot X}\left[\frac{1}{n} + \frac{(X - \overline{X})^2}{\sum X^2}\right]$$

Esta variancia aumenta conforme X aumenta. Para establecer los límites de confianza para una media estimada, al nivel del 95%, tendríamos:

$$LC(\widehat{Y}) = \overline{Y} + bX \pm t_{0.05} s_{Y \cdot X} \sqrt{\frac{1}{n} + \frac{(X - \overline{X})^2}{\sum X^2}}$$

el valor para $t_{.05}$, con $n - 2$ *g. de l.*; en el ejemplo de los cerdos, los límites de confianza para $X = 90$, son,

$$LC = 689 + 5.39(90 - 78) \pm 3.182(6.89)\sqrt{\frac{1}{5} + \frac{12^2}{178}}$$

$$= 731.76 \text{ y } 775.60$$

A veces, es necesario determinar los límites de confianza para β, el parámetro estimado por b. Para ello usamos

$s^2_b = \frac{s^2_{Y \cdot X}}{\sum X^2}$, dentro de la ecuación $b \pm \frac{t_{0.05} s_{Y \cdot X}}{\sqrt{\sum X^2}}$

en el ejemplo, $5.93 \pm \frac{3.182(6.89)}{178} = 3.75 \text{ y } 7.03$

concluimos que β se encuentra entre 3.75 y 7.03 *lbs.* de balanceado por libra de peso corporal de los cerdos, con 95% de seguridad de estar en lo cierto.

12.7 Propiedades y suposiciones en regresión lineal

El problema de regresión lineal tiene las siguientes propiedades:

a) El punto $\overline{X}, \overline{Y}$, está situado sobre la línea de regresión.

b) La suma de las desviaciones de Y de la línea de regresión es igual a cero.

c) La suma de los cuadrados de las desviaciones es un mínimo. Es decir, si reemplazamos la línea de regresión calculada con cualquiera otra línea, la suma de cuadrados del nuevo grupo de desviaciones será un valor más grande.

En lo posible, las variables independientes *(X)*, deben ser igual o convenientemente espaciadas, a fin de tener mayor eficiencia en la conducción del experimento y en la interpretación de los datos: temperatura, distancias de siembra, dosis de herbicida, etc.

Cuando se supone que existe una variancia común, se puede usar el cuadrado medio del error, para obtener inferencias válidas sobre la media de una población, sin tomar en cuenta el valor de *X*. Si las variancias no son homogéneas, es necesario transformar los datos, a fin de homogeneizar las variancias. Tal sería el caso de contajes de insectos o daños causados por ellos, en que las variancias se ajustan a la distribución de Poisson y no a la normal.

Capítulo 13

COVARIANCIA

El análisis de covariancia **(ANACOVA),** es un procedimiento de gran importancia en experimentación y que, desafortunadamente, no es usado con la frecuencia que se debiera.

El ANACOVA utiliza el análisis de variancia (sobre el que se trata ampliamente en el Cap. 15) y regresión, para remover la variabilidad que generalmente existe en la variable independiente *X*, ajustar medias de tratamiento y, sobre esta base, estimar con mayor fidelidad el efecto de *X* sobre la variable dependiente *Y*.

Para los datos del Cuadro 12.1, en el que se presenta pesos corporales de cerdos *(X)* y el consumo de alimento balanceado *(Y)*, si se hubiera usado varias raciones o dietas, las diferencias entre los efectos de raciones podían o no podían ser significativas; es decir puede o no puede haber diferencias en cuanto a la calidad de las raciones. Sin embargo, antes de sacar conclusiones al respecto, debemos preguntarnos: al existir variación en los pesos promedios iniciales de las camadas *(68 y 85 lbs.)*, podemos atribuir la diferencia en los pesos finales de los cerdos (a la terminación del ensayo), a la bondad de algunas de las raciones? O, quizás, buena parte de las diferencias en los pesos finales se debe a la diferencia en los pesos iniciales de las camadas? Con respecto

a esta última pregunta, es sabido que lechones grandes y de mayor peso al nacimiento, son más agresivos y, en consecuencia, consumen mayor cantidad de alimento y ganan peso más fácilmente que lechones pequeños y livianos.

Al ajustar los valores de *Y*, de acuerdo con los valores correspondientes de *X*, es posible analizar e interpretar con precisión los resultados del ensayo. Por medio del ANACOVA, la contribución al error experimental para ganancia de peso –contribución que, como se ha anotado, puede deberse a diferencias de peso inicial– puede ser calculada y eliminada.

En general, en ensayos de alimentación de animales, las diferencias entre medias de tratamientos no ajustadas, puede deberse al valor nutritivo de las raciones, a diferencias en la cantidad consumida por los animales o a las dos causas. Tal sería el caso de una dieta de alto valor nutritivo en la que, un exceso de harina de pescado, puede volver impalatable al alimento; o cuando se adiciona niveles subidos de melaza o torta de algodón a la dieta: en el primer caso pueden producirse diarreas y en el segundo intoxicaciones por el contenido de gosipol de la torta.

Si, en los casos indicados, se ajusta las diferencias entre ganancia de peso, a un consumo igual de alimento para todas las raciones, esas medias ajustadas indicarán si, en realidad, las dietas tienen valor nutritivo diferente.

Para ensayos de tipo agronómico, el uso del ANACOVA es también de gran utilidad. Supongamos el caso de parcelas experimentales en las que se prueba el rendimiento de algunas variedades de maíz o niveles de fertilizante o fungicida; si, por una granizada o ataque de plagas o pájaros antes de la cosecha, se presentan daños en las parcelas, el análisis de variancia de datos no ajustados puede no reflejar el potencial de rendimiento de acuerdo con los tratamientos establecidos. Al remover, por medio del ANACOVA, la variación de rendimiento por parcela, podremos evaluar fielmente los resultados; es decir, se supone que la población de plantas es igual en todas las parcelas y, después de este ajuste, será posible medir los rendimientos potenciales de cada tratamiento.

13.1 Usos del ANACOVA

a) Aumentar la precisión de experimentos. En estos casos, el covariado *X* (ejemplo de las raciones consumidas por cerdos), es la medida que se toma dentro de cada unidad experimental (animales), antes de aplicar los tratamientos (dietas), medida que tiende a predecir la respuesta de la unidad experimental *(X)*, al tratamiento *(Y)*.

Como se ha anotado antes, una fuente importante de error proviene del azar, del hecho de que, por randomización, cerdos más grandes sean asignados a ciertas raciones. Al remover las diferencias de peso inicial, se obtiene un error experimental de menor tamaño y es posible comparar medias de tratamientos con mayor confianza. Este es posiblemente el uso más común del ANACOVA.

b) Interpretación de datos. Los procedimientos aritméticos (ADEVA) y las técnicas estadísticas (regresión), permiten al investigador comprender los principios que determinan ciertos resultados en un ensayo experimental. Por ejemplo, podemos suponer que ciertos tratamientos influyen no sólo en la variable dependiente *(Y)*, sino también en la variable independiente *(X)*. La utilización del ANACOVA nos permitirá confirmar o negar esta suposición.

c) Ajustar medias de tratamiento frente a posibles fuentes de error. Los ejemplos de consumo de alimentos por animales y de parcelas de maíz afectadas por causas ambientales, tipifican este uso del ANACOVA. Después de ajustar medias de tratamientos, es importante determinar si los tratamientos están influyendo sobre la variable independiente *(X)*, en cuyo caso la interpretación de los datos variaría. Esto se debe a que las medias de tratamientos ajustadas, estiman o predicen valores, suponiendo que las medias de tratamientos de la variable independiente *X* son iguales.

El ajuste remueve parte de los efectos de tratamientos cuando las medias de la variable independiente son afectadas por los tratamientos; en estos casos, se debe interpretar los resultados del ensayo con cuidado.

d) Control del error. El objetivo del ANACOVA es incrementar la precisión al medir efectos de tratamientos removiendo, por regresión, influencias ambientales que no han podido ser controladas en forma eficiente por el diseño experimental.

Básicamente, se usa el ANACOVA para controlar el error y ajustar medias de tratamientos, cuando la variable independiente mide efectos ambientales y no está influenciada por los tratamientos. En un ensayo de aplicación de úrea a parcelas de raigrás, la cantidad de nitrógeno aplicada, puede determinar diferencias en la población de plantas por parcela. Cuando la variable independiente (población), está afectada por los tratamientos, el análisis de rendimiento de forraje ajustado para las diferencias en población, remueve parte del efecto de la aplicación de nitrógeno y, en esta forma, es posible equivocarse en la interpretación de los datos.

e) Estudiar el tipo de respuesta. Cuando se trata de relacionar aplicación de úrea y rendimiento de forraje, es interesante estudiar si el tipo de respuesta es lineal o curvilineal. Este sería un caso típico de regresión, cuando se establece un ensayo de fertilización como el anotado, en varias localidades y a lo largo de varios años.

f) Desdoblamiento de la covariancia en sus partes constitutivas; y,

g) Estimar datos perdidos.

13.2 Suposiciones en el ANACOVA y el modelo lineal aditivo

Las suposiciones en covariancia deben ser una combinación de aquellas propuestas en el ADEVA y en regresión:

1. Las medias deben medirse sin error y tener valores fijos.

2. Debe haber homogeneidad en la regresión, luego de haber removido los efectos de bloques y tratamientos; además, se supone que la regresión de Y en X, es lineal e independiente de bloques y tratamientos.

3. Los residuos (ε_{ij}), son normal e independientemente distribuidos, con media cero y variancia común.

Para un diseño de bloques completos al azar, la descripción del modelo lineal aditivo, esta dado por,

$$Y_{ij} = \mu + \tau_i + \rho_j + \beta(X_{ij} - \overline{X}..) + \varepsilon_{ij}$$

donde: μ= *media*
X_{ij} = *variable independiente o concomitante*
Y_{ij} = *variable dependiente*
τ_i = *efecto de tratamientos*
ρ_j = *efecto de bloques*
β = *coeficiente de regresión*
$\overline{X}..$ = *media de la variable independiente*
ε_{ij} = *error*

13.3 Modelo matemático

El Cuadro No 13.1 presenta el modelo del ANACOVA y el procedimiento para obtener los valores de acuerdo a las diferentes fuentes de variación.

El análisis de los valores ajustados representa los puntos alrededor de la línea de regresión. La suma de cuadrados debida a regresión, significa la dependencia de *Y* en *X*. Si el valor de *F* (cociente entre el cuadrado medio de tratamientos, sobre el cuadrado medio del error), resulta significativo, existen diferencias entre tratamientos. La Tabla No VI del Apéndice, da los valores de *F*.

CUADRO No 13.1 ANACOVA para un diseño en bloques completos al azar.

Fuente de Variación	g. de l.	SC(X) ($\sum X^2$)	SP(XY) ($\sum XY$)	SC(Y) ($\sum Y^2$)	Valores ajustados		
					g. de l.	S.C	C. M.
Total	$n-1$	$\sum_i^n X^2 - \frac{\left(\sum_i^n X\right)^2}{n}$	$\sum_i^n XY - \frac{\left(\sum_i^n X\right)\left(\sum_i^n Y\right)}{n}$	$\sum_i^n Y^2 - \frac{\left(\sum_i^n Y\right)^2}{n}$	$n-2$	$\sum Y^2 - \frac{(\sum XY)^2}{\sum X^2}$	
Rep.	$r-1$	$\frac{\sum_i^r \left(\sum_i^t X\right)^2}{t} - FC^{[1]}$	$\sum_i^r \left(\sum_i^t X\right)\left(\sum_i^t Y\right) - FC^{[2]}$	$\frac{\sum_i^r \left(\sum_i^t Y\right)^2}{t} - FC^{[3]}$			
Trat.	$t-1$	$\frac{\sum_i^t \left(\sum_i^r X\right)^2}{r} - FC$	$\sum_i^t \left(\sum_i^r X\right)\left(\sum_i^r Y\right) - FC$	$\frac{\sum_i^t \left(\sum_i^r Y\right)^2}{r} - FC$			
Error	$(t-1)(r-1)$	Dif.	Dif.	Dif.			
Tratamientos + Error					$(n-1)(t-1)-1$	$\sum Y^2 - (\sum XY)^2/\sum X^2$	SC/g. de l.
Error					$(t-1)(r-1)-1$	$\sum Y^2 - (\sum XY)^2/\sum X^2$	SC/g. de l.
Trat. Ajustados					$t-1$	SC(T+E) - SC(E)	SC/g. de l.

$$[1]\ FC = \frac{\left(\sum_i^n X\right)^2}{n}$$

$$[2]\ FC = \frac{\left(\sum_i^n X\right)(\sum Y)}{n}$$

$$[3]\ FC = \frac{\left(\sum_i^n Y\right)^2}{n}$$

FC = Factor de corrección
g. de l. = Grados de libertad
SC: Suma de cuadrados
CM: Cuadrado medio

13.4 Ilustración

A manera de ejemplo, se presenta seguidamente los resultados de un ensayo con cinco camadas de cerdos (repeticiones) y cuatro dietas (tratamientos), donde,

X: peso inicial de los cerdos (var. independ.)
Y: peso final de los cerdos (var. dependiente)

TRATAMIENTOS

Camada	1		2		3		4		Total	
	X	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*
I	4	10	1	5	2	5	3	5	10	25
II	2	5	3	9	6	12	1	7	12	33
III	4	8	2	5	4	10	4	6	14	29
IV	6	15	1	1	6	10	2	9	15	35
V	4	12	3	5	2	8	1	3	10	28
	20	50	10	25	20	45	11	30	61	150

	ADEVA Y			ADEVA X		
Fuente	*g. l.*	*SC*	*CM*	*g. l.*	*SC*	*CM*
Rep.	4	16,0	4,0	4	5,20	1,3
Tratam.	3	85,0	28,33(NS)	3	18,15	6,05(NS)
Error	12	122,0	10,16	12	29,60	2,47
Total	19	223,0		19	52,95	

Para los ADEVAS de X e Y, se procede exactamente como en la Sec. 16.1.

Como vemos no hay diferencia en peso inicial (variabilidad neta) ni en peso final debido a tratamientos.

Cálculo ANACOVA:

SP*XY*:

$$Total: (4 \times 10) + (2 \times 5) + \ldots + (1 \times 3) - FC = 85.50$$

$$FC = \frac{(61)(150)}{4 \times 5} = 457.50$$

Bloques: $(10 \times 25) + (12 \times 33) + \ldots + (10 \times 28)/4 - FC = 6.75$

Tratamientos: $(20 \times 50) + (10 \times 25) + \ldots + (11 \times 30)/5 FC = 38.50$

Error = por diferencia = 40,25

ANACOVA

Fuente	*g. l.*	SC(*X*)	SP*XY*	SC(*Y*)	VALORES AJUSTADOS			
					g. l.	SC	CM	F
Total	19	52,95	85,50	223,0				
Bloques	4	5,20	6,75	16,0				
Tratam.	3	18,15	38,50	85,0	3	9,85	3,28	0,54(NS)
Error	12	29,60	40,25	122,0	11	67,27	6.12	
Trat. + E.	15	47,75	78,75	207,0	14	77,12		

$$SC\ Reg.(Error): \frac{SC(Y) - (SPXY)^2}{SC(X)} = 122.0 - \frac{(40.27)^2}{29.60} = 67.27$$

$$SC\ Trat. + Error = \frac{SC(Y) - (SPXY)^2}{CS(X)} = 207.0 - \frac{(78.75)^2}{47.75} = 77.12$$

Tratamientos ajustados: $77.12 - 67.27 = 9.85$

$$F = \frac{3.28}{6.12} = 0.54\ NS$$

Para ensayos de ganadería muchas veces nos interesa trabajar con ganancias de peso diario, en lugar de pesos iniciales y finales. Restamos $Y - X$; en el ejemplo anterior tenemos:

TRATAMIENTOS

Camada	1	2	3	4	
I	6	4	3	2	Los valores X quedan iguales.
II	3	6	6	6	
III	4	3	6	2	
IV	9	0	4	7	
V	8	2	6	2	

ADEVA Y

Fuente	*g.l.*	*SC*	*CM*
Bloques	4	7.70	1.92
Tratam.	3	26.15	8.71 NS
Error	12	71.10	5.92
Total	19	104.95	

ANACOVA

Fuente	*g. l.*	SC(*X*)	SP*XY*	SC(*Y*)	VALORES AJUSTADOS			
					g.l.	*SC*	*CM*	*F*
Bloques	4	5,20	1.55	7.70				
Tratam.	3	18.15	20.35	26.15	3	9.85	3.28	0,54(NS)
Error	12	29.60	10.60	71.10	11	67.27	6.12	
Trat. + E	15	47.75	30.95	97.25	14	77.12		

Como vemos, no han variado los resultados anteriores. Da igual trabajar con pesos iniciales y finales o con las ganancias diarias (diferencia).

El ANACOVA confirma al ADEVA: no hay diferencias significativas, pero no nos dice si debemos ajustar las medias de tratamientos.

Antes del ajuste, determinamos el coeficiente de regresión *b*:

$$b = \frac{SPXY}{CS(X)} = \frac{40.25}{29.60} = 1.36$$

Es decir que por un aumento de 1 en el peso inicial *X*, hay un aumento de 1.36 en el peso final *Y*.

Para determinar si debemos ajustar medias de tratamientos, vemos la homogenidad de regresiones, $H_0 = \beta = 0$; si el valor de F es significativo, $H_0 \neq 0$, ajustamos las medias.

$$SC \text{ debido a Reg.} : \frac{SPXY^2}{SC(X)} = \frac{40.25}{29.60} = 54.73 \qquad CM = SC/1 = 54.73$$

$$F = \frac{CM \text{ regresión}}{CM \text{ residual}} = \frac{54.73}{6.12} = 8.94 ** \ (1.11 \ g \text{ de } l)$$

Aceptamos H_i $\beta \neq 0$. Las líneas no son paralelas, no son homogéneas y debemos hacer ajuste de medias.

13.5 Ajuste de Medias: $\widehat{Y}_i = \overline{Y}_i - b_{YX}(\overline{X}_i - \overline{X}..)$

$\widehat{Y}_i$: *media de cada grupo de tratamientos (Y)*
b_{YX}: *coeficiente de regresión*
$\overline{X}_i$: *media de cada grupo (X)*
$\overline{X}..$: *media general (X)* $61/20 = 3.05$

Tratamiento	$(\widehat{Y}_i) = \overline{Y}_i - b_{YX}(\overline{X}_i - \overline{X}..) = \widehat{Y}$
1	10 – 1.36(4.0 – 3.05) = 8.708
2	5 – 1,36(2,0 – 3,05) = 6.428
3	9 – 1,36(4,0 – 3,05) = 7.708
4	6 – 1,36 (2,2 – 3.05) = 7.156

Si se encuentra que las medias de tratamientos no ajustadas resultan significativas, pero no las ajustadas, esto querría decir que las diferencias encontradas entre las no ajustadas sólo reflejarían diferencias entre el covariado, más no entre los tratamientos.

La variancia de la diferencia entre dos medias ajustadas es:

$$s^2\overline{d} = s^2_{y \cdot x}\left[\frac{2}{r} + \frac{(\overline{X}_i - \overline{X}_j)^2}{SC(X)}\right]^{con} (r-1) - 1(t-1) g.l.$$

para las medias 1 y 2 :

$$s^2\overline{d} = 6.12\left[\frac{2}{5} + \frac{(4-2)^2}{29.60}\right] = 3.275; \ s\overline{d} = 1.81 \ lbs.$$

Se usa esta fórmula cuando la investigación está avanzada. Sin embargo, cuando se comienza una investigación, se usa un solo valor promedio, dado por Finney (1946), para comparar todas las medias. Se sugiere esta ecuación cuando los grados de libertad del error son 20 o más.

$$s\bar{d} = \sqrt{\frac{2s^2_{Y\cdot X}}{r}\left[1 + \frac{SC(trat)}{(t-1)SC(E)}\right]}$$

$$= \sqrt{\frac{2(6.12)}{5}\left[1 + \frac{18.15}{3 \times 29.60}\right]} = 1.717 lbs.$$

13.6 Ganancia en Precisión por Covariancia

CM Error (Y) antes del ajuste $= \frac{122.0}{12} = 10.16$ *con* 12 *g. l.*

CM Error después del ajuste = 6.12 con 11 *g. l.*

Ajustamos este valor con la fórmula,

$$s^2_{Y\cdot X}\left[1 + \frac{SC(trat)}{(t-1)SC(E)}\right] = 6.12\left[1 + \frac{18.15}{3 \times 29.60}\right] = 7.37$$

$$\text{Eficiencia} = \frac{s^2_Y}{s^2_{Y\cdot X}} \times 100$$

$$= \frac{10.16}{7.37} \times 100 = 138\%$$

100 repeticiones con ANACOVA, son tan eficientes como 138 sin covariancia.

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1826
QUITO


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO



Segunda Parte

PRINCIPIOS DE DISEÑO EXPERIMENTAL


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 14

EXPERIMENTACIÓN

14.1 Introducción

La experimentación tiene por objeto comprobar en la práctica, una hipótesis formulada sobre el valor de un procedimiento, factor o elemento; generalmente se usa un "testigo" o "control", que constituye el individuo o parcela experimental, que no recibe el tratamiento, cuyo efecto se desea medir.

Por ejemplo, si se va a probar en una región el grado de adaptación y rendimiento de una nueva variedad de cebada, se planeará el ensayo de tal manera de incluir variedades locales como testigos. Si la nueva variedad presenta mayor resistencia a enfermedades, insectos, al encamado y, en consecuencia, da mayores rendimientos por hectárea, esa variedad será recomendada para la zona.

Para poder expresar su potencialidad genética, la variedad mejorada debe disponer de adecuadas cantidades de nutrimentos en el suelo. En el ejemplo anterior, interesa al investigador diseñar un ensayo de fertilización, en el que se prueben varias dosis de nitrógeno, fósforo y potasio. El nitrógeno puede ser aplicado en su totalidad a la siembra, o puede ser usado en aplicación fraccionada. Esta sería una variable adicional: épocas de aplicación. En igual

forma, se puede probar sistemas y densidades de siembra, épocas de cosecha, labores culturales, uso de insecticidas y fungicidas, etc., etc.

En investigación con animales, se puede probar balanceados, sistemas de pastoreo, carga animal, tipos de vacuna, uso de vitaminas, minerales, etc. En tales casos, se forma varios grupos de animales, tan homogéneos como sea posible, para la investigación. Para la formación de los grupos, se debe tener en cuenta aspectos como edad, peso, sexo, tratamiento anterior, etc. Mientras a varios grupos se administra diferentes dosis de vitaminas, por ejemplo, un grupo testigo no las recibe; el análisis estadístico dará información sobre el incremento de peso, altura, precocidad de los animales que recibieron la vitamina, frente a los que no la recibieron.

Hay varias definiciones sobre lo que es un experimento. Steel y Tome (1960), lo consideran como una pregunta planeada, que va a detectar nuevos hechos o confirmar o denegar los resultados de ensayos anteriores. Esos hechos pueden referirse a recomendaciones de nuevas variedades, procedimientos, pesticidas, etc. Así se plantean los factores en estudio que forman el plan o diseño del experimento.

Se ha definido también al experimento como el conjunto de reglas usadas para sacar la muestra de la población y, una vez concluido el ensayo, obtener información sobre la población.

Se considera a la investigación como un arte y como una ciencia. El investigador se apoya en ésta, para la aplicación y utilización de todos los procedimientos aritméticos disponibles y, en este sentido la investigación es una ciencia exacta.

La capacidad del profesional para determinar los objetivos del ensayo en forma clara y precisa, así como para interpretar los resultados y obtener inferencias válidas, convierte a la investigación en un arte.

14.2 Objetivos

De lo expuesto se deduce que diseñar un experimento, significa planear un trabajo de modo que reúna información que sea pertinente al problema bajo investigación.

El diseño de un experimento es, entonces, la secuencia completa de pasos previstos de antemano, para asegurar la obtención de datos apropiados, los que permitirán un análisis objetivo, que conduzca a deducciones válidas con respecto al problema en estudio. Esta definición sugiere que el investigador debe conocer el problema, así como los objetivos específicos que persigue con su trabajo. Concretamente, el profesional debe preguntarse: 1) ¿Cuáles son los caracteres o variables que van a ser analizados? ¿Cómo se va a medir su efecto?; 2) Una vez que se ha establecido la o las variables, que van a ser estudiadas, ¿existe uniformidad general con respecto a otros factores?; 3) Cuál va a ser el procedimiento a seguirse?; y 4) ¿En qué forma van a ser analizados los datos?

Resumiendo, los objetivos de la experimentación deben plantearse como preguntas que pueden ser contestadas, hipótesis sujetas a prueba o efectos a ser estimados. Los objetivos deben plantearse en forma clara y precisa; a veces, por un exceso de ambición, se incluyen más variables de las que se puede manejar y, posteriormente, interpretar; o, en la toma de datos, se incluyen más datos de lo que sería práctico anotar. Por esta razón, es importante clasificar los objetivos como principales y secundarios, puesto que algunos tipos de diseño dan mayor precisión para ciertas comparaciones que para otras.

Si bien los ensayos factoriales permiten analizar varios factores a la vez, ahorrando tiempo y espacio, en un ensayo muy complicado, puede ser difícil interpretar ciertas interacciones de alto grado.

Los recursos de tiempo, espacio, dinero y grado de precisión requerida, darán una pauta para el tipo de diseño a usarse.

Finalmente, debe anotarse que el objetivo final de la experimentación es incrementar rendimientos por unidad de superficie, así como la calidad del pro-

ducto, para beneficio del consumidor y del país. La experimentación pura, académica o especulativa debiera ser llevada a cabo en las universidades, en las cuales varios proyectos de ese tipo podrían constituir tesis de grado para estudiantes. Aquellos estudios promisorios de laboratorio, invernadero y cámaras de crecimiento, podrían pasar al campo de la investigación aplicada, en las estaciones experimentales.

14.3 Fases de la experimentación

Todo investigador tiene una lista tácita de pasos a seguir cuando diseña un ensayo. En general, esos pasos coinciden con las diferentes etapas mencionadas al describir el Método Científico (Sec. 1.5), y son:

a) Enunciado del problema.

b) Formulación de la o las hipótesis.

c) Selección del procedimiento experimental y del diseño.

d) Ejecución del experimento.

e) Aplicación de los procedimientos estadísticos a los resultados.

f) Interpretación de resultados.

g) Análisis económico de los mismos, a fin de valorar la conveniencia práctica de adoptar el tratamiento cuya superioridad ha sido demostrada por el experimento.

h) Obtención de conclusiones y recomendaciones.

i) Publicación de los resultados.

La selección del procedimiento experimental y del diseño comprende, en el campo agrícola, la elección del terreno donde va a establecerse el lote experi-

mental; la determinación del número y superficie de la parcela experimental; la forma y dimensiones de la misma; su distribución en el campo y la asignación de los diferentes tratamientos a las parcelas.

La determinación de rendimiento por parcela y por hectárea, se reduce a pesar y registrar el producto de la parcela experimental. Más difícil resulta la determinación de calidad o de características organolépticas del cultivo. En muchos ensayos de intervalos de pastoreo, por ejemplo, interesa determinar la diferencia en contenido de proteína, fibra, etc., de las muestras de cada tratamiento.

14.4 Clases de experimentos

Básicamente existen cuatro clases de experimentos: **preliminares, críticos, demostrativos y ensayos regionales**.

En un experimento **preliminar**, el investigador comienza con un gran número de tratamientos, los que le darán una pauta para el trabajo futuro. Así, el fitomejorador inicia su trabajo con miles de líneas de trigo, por ejemplo, que, en su fase final, se reducirán a dos o tres, como futuras variedades. En el caso de fertilización, se combina en un experimento factorial múltiples tratamientos, para luego reducirlos, según se presente la curva de respuesta. Este último caso se refiere al **experimento crítico**, en el que es posible detectar diferencias más pequeñas entre tratamientos.

Los experimentos **demostrativos** están en el campo de la Extensión Agrícola y tratan de demostrar a los agricultores las ventajas de una nueva variedad o recomendar el uso de nuevas prácticas culturales, como uso de fertilizantes, pesticidas, etc., obtenidas en la Estación Experimental.

Los **ensayos regionales** son aquellos que usa el investigador para confirmar los resultados obtenidos en la Estación Experimental. Para condiciones como las del Ecuador, los factores ambientales y microambientales presentan variaciones tan drásticas en cortas distancias, que a veces, los resultados obtenidos en la Estación Experimental no se cumplen en localidades situadas a cortas distancias. Esas variaciones se refieren a suelo, clima, topografía, etc.

En el campo agropecuario, existen algunos factores –algunos de ellos fuera de nuestro control– que influyen en el resultado final que se llama producción por hectárea. Los aspectos de clima y suelo son, probablemente, los más importantes.

Como premisa, debemos aceptar que el ensayo experimental, llevado a nivel de Estación Experimental, tiene aplicación limitada: las conclusiones pueden ser aplicadas a condiciones similares a aquellas en que se realizó el ensayo. Para poder tener confianza en los resultados y ampliar el alcance de las inferencias, es necesario repetir el ensayo en diferentes lugares, sobre distintos suelos, en condiciones climáticas variadas.

Con este criterio, se establecen los ensayos regionales, para los que se seleccionan localidades representativas de grandes zonas, a las que se espera beneficiar con la introducción de una nueva variedad o práctica agropecuaria. Es decir que, en esta forma, es posible generalizar la hipótesis formulada para el ensayo original.

Los adelantos conseguidos por las técnicas de laboratorio y la experimentación en macetas, no ha eliminado la necesidad de "pasar" este tipo de investigación al campo, que, en definitiva, es a donde se trata de llevar la aplicación de los resultados obtenidos. Sólo una pequeña parte de los varios factores de suelo y clima que afectan al desarrollo de las plantas, puede ser trasladada al invernadero, a cámaras de crecimiento, fitotrones, etc. Sin embargo, esta fase de la investigación es de suma importancia puesto que proporciona pautas y modelos de respuesta. Así, en el invernadero, se usan soluciones nutritivas y elementos radiactivos en frascos de cristal, para medir la absorción y translocación de elementos en la planta. Se emplea la técnica del "tratamiento completo menos uno", con el objeto de establecer la curva de respuesta en presencia y ausencia de ciertos elementos. Para trabajos de fitomejoramiento, es posible producir dos o tres generaciones en un año, mientras que en el campo se tendría solo una. En este último caso, existe el riesgo de perder material valioso por factores climáticos como heladas, granizadas, etc.; si se está trabajando en mejoramiento de alfalfa, es factible inducir floración mediante el uso de luz artificial en el invernadero; resulta más cómodo realizar cruzamientos manuales; no existe el peligro de aborto de flores por llovizna, etc.

Sin embargo, sería peligroso entregar al agricultor una recomendación en base a resultados obtenidos en macetas, o una variedad de una especie cuya semilla ha sido cosechada en el invernadero.

De lo expuesto, se puede concluir que el experimento de campo es insustituible ya que, en esta forma, sometemos a la planta a las diversas condiciones que interaccionan para producir un resultado, en condiciones que las tendría el agricultor en la hacienda.

14.5 Tratamientos

Se ha usado repetidamente la palabra "tratamiento", en secciones precedentes. Como algunos otros términos usados en Estadística, la palabra tratamiento fue aceptada en la literatura técnica, debido a su empleo en la experimentación agrícola. Igual es el caso de expresiones como **bloques, parcela, testigo,** entre otras, que son de uso común en muchos campos de experimentación no agrícola.

Para el investigador, la palabra **tratamiento**, denota diferentes procedimientos o variables, cuyos efectos van a ser medidos y comparados. En la selección de tratamientos, es importante definir claramente cada tratamiento y la forma cómo el conjunto de ellos van a permitir alcanzar los objetivos del experimento. El efecto del tratamiento puede ser medido en la unidad experimental o en una parte de ella, lo que constituye la unidad de muestreo, existiendo mayor variabilidad en la primera.

En la Sec. 14.1 se anota algunos ejemplos de tratamientos en experimentación agrícola y animal.

En estudios sociológicos y psicológicos, los tratamientos se pueden referir a edad, sexo, grado de educación, religión, etc.

En la industria, el productor de detergentes puede establecer como tratamiento el tipo de agua (dura o suave), temperatura del agua, duración del lavado, marca y tipo de lavadora.

Se ha mencionado también la palabra testigo o control, como un tratamiento que generalmente se incluye en el diseño del experimento. Muchas veces se usa estas palabras para referirse a un tratamiento sobre el que no tenemos interés especial pero que puede ser necesario para demostrar, por comparación, que los otros tratamientos son eficaces.

En una prueba de fertilización, se puede aplicar cuatro dosis de P^2O^5: 0, 50, 100 y 150 *kg./Hectárea;* la "dosis" o "nivel" cero, constituye el testigo o control.

Para ilustrar la necesidad de usar o no un testigo, supongamos el caso de la aplicación de tres fertilizantes potásicos (fuentes), en suelos de origen volcánico de la Sierra; las tres fuentes potásicas contiene igual cantidad de potasio. El testigo sería la dosis "cero potasio". Para esta situación, pueden darse tres casos: a) Si la efectividad de la aplicación de potasio ha sido suficientemente demostrada por ensayos anteriores, queda por determinar cual de las tres fuentes es superior: en este caso, no hay necesidad de usar un testigo; b) La aplicación de potasio es generalmente efectiva, pero ocasionalmente no se detecta diferencias con el testigo. Esta situación puede presentarse cuando el suelo tiene un alto contenido de cenizas volcánicas; en este caso debe incluirse un testigo; c) Igual sería el caso cuando se desconoce el efecto de la aplicación de potasio en una zona nueva.

14.6 Selección de tratamientos

En muchos casos, la selección de tratamientos tiene una importancia decisiva en la precisión de un experimento. Se puede incrementar el grado de precisión, al probar más de una variable, en lugar de diseñar un ensayo para cada variable. Por ejemplo, si tratamos de introducir cinco nuevas variedades de alfalfa a una localidad (prueba de adaptación y rendimiento), se puede añadir la variable "épocas de corte", con lo que no solamente se obtiene doble información dentro del mismo experimento, sino que se aumenta el grado de precisión, al disponer de mayor número de grados de libertad para el error.

Se presenta otra situación, en la que la selección de las dosis adecuadas de un elemento, es de importancia. Si planteamos la hipótesis de que la respuesta

a un fertilizante o fungicida es lineal, nos interesa conocer la gradiente de la línea de respuesta. Es decir, tiene importancia saber hasta qué cantidad de elemento aplicado, el rendimiento asciende; bien podría suceder que, pasado cierto nivel "óptimo", la línea de respuesta se estabiliza o comienza a descender.

El criterio general es que, por lo menos, se debe tener tres dosis que darán una idea sobre la curva de rendimiento, dentro del eje de coordenadas. Por ejemplo, si la recomendación de una casa comercial, es la de usar 4 kg. de producto activo por hectárea, haríamos bien en probar, además, 2 y 6 kg. por hectárea. Como se verá en el capítulo sobre factoriales, el número de niveles de un factor y su espaciamiento, tienen mucho que ver con la precisión del experimento.

14.7 Unidad experimental

Se llama **unidad experimental** a la unidad material a la cual se aplica un solo tratamiento (que puede ser una combinación de varios factores), dentro de una repetición o bloque del ensayo.

La unidad puede ser una parcela de terreno, un animal, un paciente de hospital, una hoja, un árbol, una maceta, un pedazo de alambre, un lote de semilla, etc., etc. Como en otros casos mencionados antes, el término usado en experimentación agrícola "parcela experimental", ha sido extendida a otros campos, para denotar aquellas unidades experimentales.

La característica general de tales unidades, es la de que presentan resultados diferentes, aun cuando se las someta al mismo tratamiento: estas diferencias, grandes o pequeñas, constituyen, en parte, lo que se llama **error experimental**, como se detalla en la Sec. 14.10.

Es conveniente tener bloques compactos y parcelas largas y angostas. En cuanto al tamaño de la parcela, como se anota en la Sec. 14.8, es preferible aumentar el número de repeticiones y disminuir el tamaño de la parcela. Sin embargo, el tamaño de la parcela debe ser tal, que permita operaciones de labranza, por ejemplo, tal como las podría hacer el agricultor, que es a quien llegarán las recomendaciones.

Otro aspecto que limita el tamaño de la parcela, se refiere a los "bordes" u "orillas". Hay tratamientos (uso de herbicidas, insecticidas y fertilizantes), para los que se debe dejar surcos bordes a fin de evitar que, por efecto del viento, o por lixiviación, los tratamientos aplicados a una parcela afecten a las parcelas contiguas. A este respecto, existe discusión sobre el hecho de si se debe dejar "calles" entre parcelas. No hay duda de que mientras más grande es el bloque, mayor será el tamaño del error experimental, por el aumento en la heterogeneidad del suelo. En general, es aconsejable disponer las parcelas unas junto a otras.. sin calles, descartando, para la toma de datos, uno o más surcos bordes. Por ejemplo, si la parcela tiene siete surcos, se puede eliminar uno a cada lado y cosechar los cinco centrales. También se debe eliminar de 50 a 100 cm. hacia los extremos de los surcos, según la longitud de los mismos. En caso de siembras al voleo, se eliminará una faja alrededor de la parcela. Otra ventaja de no dejar calles entre parcelas se refiere al hecho de que, en esa forma, disminuye el área de limpieza a mano que, generalmente, hay que hacer en todo ensayo. Se debe, sin embargo, dejar calles entre repeticiones.

Existen excepciones en lo que se relaciona con dejar calles entre parcelas: para ciertos ensayos de Entomología, por ejemplo, es necesario circular entre las parcelas para efectuar la aplicación de insecticidas con rociadoras de cierto tamaño.

De todos modos, el buen criterio del investigador y el conocimiento sobre la variabilidad del suelo, indicarán si se debe dejar calles y cual debe ser su ancho.

Los efectos de bordes son conocidos por todo investigador: en parcelas de cereales, pastos, papa, es fácil ver que los surcos externos de la parcela, aquellos que no tiene competencia por el lado exterior, desarrollan mejor que los surcos interiores. La cosecha de toda la parcela sesgaría la toma de datos, aumentando la variancia del error, por aumento de la heterogeneidad entre las parcelas. Es importante que los surcos o franjas bordes reciban el mismo tratamiento que la parcela neta.

14.8 Repeticiones

Se llama **repetición, replicación, bloque o reproducción,** al conjunto básico de tratamientos dentro del ensayo, los que, en experimentación agrícola, van uno al lado del otro en el terreno, formando un "bloque" más o menos compacto.

Cualquiera que sea la fuente del error experimental, la repetición del experimento básico disminuye el error asociado con la diferencia entre los resultados promedios para dos tratamientos, con la condición de que se haya tomado ciertas precauciones (como la **randomización**), a fin de evitar que ciertos tratamientos resulten favorecidos en una repetición más que en otra. En esta forma, los errores que afectan, positiva o negativamente, a cualquier tratamiento, se equilibran con el aumento en el número de repeticiones.

La heterogeneidad del suelo es universal. En consecuencia, debe evitarse, para fines de experimentación, terrenos de fuerte pendiente, que presenten erosión, acumulación de materia orgánica, etc.; en lo posible, el terreno debe ser plano.

En el pasado, se usaba parcelas experimentales grandes, sin repetición. Su tamaño variaba de 100 a 8.000 metros cuadrados, con el criterio de que representaban mejor las condiciones normales de una hacienda. Este tamaño de parcela está bien para ensayos demostrativos, cuando se desea presentar hechos conocidos y cuando las diferencias entre tratamientos son grandes.

La investigación moderna ha demostrado que es posible controlar la variación del suelo, disminuyendo el tamaño de la parcela y aumentando el número de repeticiones. Se ha probado, así mismo, que el incremento del número de repeticiones es dos veces más efectiva que el aumento del largo del surco.

Debe aclararse que múltiples lecturas o mediciones de un mismo lote (de terreno, de semillas, de algo producido en serie), no representan necesariamente repeticiones verdaderas. Por ejemplo, supongamos que se desea probar el poder germinativo de dos variedades de raigrás, A y B, y se toma varias muestras de semilla de cada una. Si se intentara un análisis estadístico a partir

de las muestras, veremos que no se dispone de una estimación válida del error para probar la diferencia entre variedades. La variación entre semillas dentro de variedades, proporciona una estimación válida del error para indicar sólo la variabilidad de variedad a variedad. Se podría usar como repeticiones muestras de varias cosechas de cada variedad. En el caso citado, los datos de germinación no constituyen repeticiones verdaderas sino solamente medidas repetidas tomadas de la misma unidad experimental.

Sin embargo, algunos investigadores en Ciencia Animal, aceptan los varios pastoreos o cortes producidos por el mismo potrero, como una forma de repetición.

Resumiendo, podemos afirmar que, cuando un tratamiento aparece más de una vez en el experimento, se dice que existe repetición. Las funciones que cumplen las repeticiones son:

a) Estimar el error experimental.

b) Aumentar la precisión del experimento, reduciendo la desviación típica de medias de tratamientos.

c) Incrementar el alcance las inferencias.

d) Ayudar a controlar el error experimental.

e) Al aumentar el número de repeticiones, la desviación típica de la media y la desviación típica de la diferencia entre tratamientos disminuye:

N° Repeticiones	$s\overline{X}$
1	.420
2	.298
3	.243
4	.210
5	.188
6	.172
•	•
•	•
12	.121

De la comparación de valores de $s\overline{X}$, se desprende que el valor de $s\overline{X}$ para 1 y 4 repeticiones, baja exactamente a la mitad para 4 repeticiones. Si bien el valor de $s\overline{X}$ para 12 repeticiones es de .121, razones de orden práctico y económico, como se anota en la Sec. 14.9, hacen que, para experimentos de campo convencionales, la mayoría de investigadores usen de tres a seis repeticiones.

En definitiva, el número de repeticiones tiene que ver con el número de grados de libertad *(g. de l.)*, lo cual, a su vez, influye en el grado de precisión con que va a compararse medias de tratamientos. Supongamos que se van a probar cinco variedades (tratamientos), en un diseño de bloques completos al azar, con dos y seis repeticiones. El esquema del análisis de variancia (ADEVA) sería, en cada caso:

5 TRATAMIENTOS, 2 REPET.		5 TRATAMIENTOS, 6 REPET.	
F. de V.	*g. de l.*	F. de V.	*g. de l*
Total	9		29
Reps.	1		5
Tratam.	4		4
Error	4		20

Como se podrá apreciar en el Cap. 16, cuando se presente el diseño de bloques completos al azar, los 20 *g. de l.* del error, en el segundo caso, permitirán comparaciones mucho más precisas entre medias de tratamientos, que los 4 *g. de l.* en el primer caso.

Es importante que el área incluida dentro de cada repetición y que contiene a todos los tratamientos, sea lo más uniforme posible. Sinembargo, no es necesario, y podría ser no deseable, tener condiciones uniformes entre repeticiones: así se limitaría el alcance de las inferencias tratándose de poblaciones grandes. Si el lote de terreno escogido para el ensayo está ubicado en una pendiente, se puede ubicar una o dos repeticiones juntas en condiciones similares y las otras dos repeticiones más abajo, en igual forma.

En muchos ensayos de campo –el caso de pastos o árboles frutales– el ensayo se lo mantiene por varios años. Por otro lado, las condiciones climáticas

varían de año a año y es importante conocer su efecto sobre los tratamientos. Así mismo se usan varias localidades para probar variedades o tratamientos. Las repeticiones en el tiempo (años) y en el espacio (localidades), pueden ser consideradas como formas amplias de repetición. El objeto es ampliar el alcance de las inferencias.

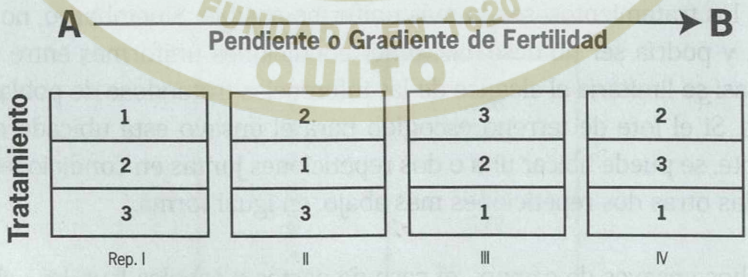
14.9 Disposición de los bloques

El arreglo de los bloques en el campo tiene importancia cuando los tratamientos se refieren a prácticas culturales que requieren del uso de equipos de labranza, sistemas de riego, equipos de aspersión como bombas de carretilla para aplicación de insecticidas, herbicidas, etc.

Por otro lado, como se indica en la Sec. 16.4, se debe dejar la máxima variación entre bloques, a fin de que las unidades experimentales, dentro del bloque, compartan condiciones uniformes, es decir, presenten variaciones mínimas.

En la Sierra ecuatoriana, frecuentemente se encuentra localidades que reúnen condiciones ambientales deseables para llevar a cabo un ensayo pero, en las que la gradiente del terreno obliga al investigador a disponer los bloques y parcelas de modo diferente al que adoptaría en terreno plano.

Cuando la pendiente del terreno, o la gradiente de fertilidad, van de A a B, como se indica en la Fig. 14.1, resulta ventajoso colocar los bloques uno a continuación de otro, a lo largo de la pendiente AB. Los tratamientos se sortean dentro de cada bloque, quedando, por así decirlo, bajo la misma cota.



14.1 Disposición de los bloques a lo largo de la pendiente AB.

14.10 Factores que determinan el número de repeticiones

a) **Precisión requerida:** mientras más pequeña sea la desviación de la hipótesis nula que tratamos de detectar (es decir, mientras más pequeña sea la diferencia entre medias de tratamientos), mayor debe ser el número de repeticiones.

b) **Naturaleza del material experimental.** Algunos suelos, así como cierto tipo de material experimental, son más heterogéneos que otros, en cuyo caso, se debe aumentar el número de repeticiones.

c) **Diseño usado.** Cuando el número de tratamientos es crecido y exige el uso de unidades experimentales heterogéneas, el error experimental por unidad aumenta.

d) **Presupuesto.** Se refiere a la mano de obra para mantenimiento del ensayo y la toma de datos.

e) **Cantidad de semilla disponible.** Generalmente, el fitomejorador obtiene algunos gramos de semilla en la primera fase del mejoramiento. En este caso, no es posible tener el número de repeticiones requerido.

f) **Tierra disponible.**

14.11 Error experimental

Los resultados experimentales varían no sólo por la acción de los tratamientos, sino también debido a variaciones ambientales, las que tienden a enmascarar el efecto de los tratamientos. Para expresar estas variaciones, generalmente se usa el término error experimental, en el que error no quiere decir "equivocación", sino que incluye todo tipo de variación externa, ajena al material experimental.

Error experimental sería, entonces, el fracaso de llegar a resultados idénticos en dos unidades experimentales tratados en la misma forma. También se define al error experimental como la medida de la variación que existe entre

las observaciones de unidades experimentales, bajo el mismo tratamiento; es decir, la variación no proveniente de los tratamientos.

Existen dos clases de variación: la variación inherente al material experimental al cual se aplica los tratamientos, y aquella que proviene de la falta de uniformidad en la conducción física del experimento.

Si los resultados obtenidos tienen la precisión requerida, como para llegar a conclusiones válidas, ninguna de las dos fuentes de variación deben preocupar al investigador.

Resumiendo, la variación de los datos refleja: 1) errores de experimentación, 2) errores de observación, 3) errores de medición, 4) la variación del material experimental (entre unidades experimentales) y 5) efectos combinados de todos los factores extraños a los tratamientos en estudio.

14.12 Control del error experimental

En muchos campos de investigación, de acuerdo a la disponibilidad de tiempo que se puede asignar a un experimento, el error experimental puede influir de tal manera en los resultados, que solamente es posible detectar las diferencias entre tratamientos que son de cierta magnitud. En consecuencia, se ha dedicado bastante atención a los métodos de investigación que permiten aumentar la precisión en los experimentos.

Esos métodos se refieren a 1) Aumentar el tamaño del experimento (no precisamente en cuanto a superficie), usando más repeticiones o incluyendo tratamientos adicionales; 2) Refinar la técnica experimental, usando un diseño eficiente y tomando los datos en forma cuidadosa; 3) Manejando el material experimental de tal modo que permita reducir la variabilidad dentro del mismo; es decir, usando material homogéneo y 4) Usando unidades experimentales del tamaño y forma adecuados.

Con relación a la técnica usada, por ejemplo, en la aplicación de fungicidas, la misma persona debe hacer todas las aplicaciones en un mismo bloque. Se

debe cosechar toda una repetición en el mismo día. Se usará la misma balanza para pesar el rendimiento de toda una repetición, etc.

14.13 Selección del sitio y del material experimental

Antes de diseñar el experimento, es importante seleccionar el sitio (campo) o el área (población) en que se va a llevar a cabo el experimento. En experimentación agrícola, debe escogerse una localidad que sea representativa de una zona grande, a la que se va a beneficiar con las conclusiones obtenidas. De otro modo, las recomendaciones llegarán tan sólo al propietario de la hacienda en la que se hizo el ensayo. Una vez escogida la localidad (hacienda), se debe seleccionar un lote de terreno que sea tan homogéneo como sea posible, evitando pendientes y sitios irregulares. Como se ha indicado antes, la heterogeneidad del suelo es universal e, inclusive en terrenos planos, la fertilidad varía de un metro cuadrado a otro. Mientras mayor sea la uniformidad general del lote, mayores son las probabilidades de reducir el error experimental y de obtener conclusiones válidas y confiables del ensayo.

En segundo lugar viene la selección de la unidad experimental. En experimentación agrícola se han hecho numerosos estudios sobre la variabilidad de varios cultivos en parcelas de diferente forma y tamaño. La idea es obtener la máxima precisión de acuerdo al tiempo, dinero y mano de obra disponibles.

Así mismo, con respecto al material experimental, se deberá usar aquel que es representativo, o sea el más usado en la zona. Si se trata de iniciar un ensayo de fertilización en maíz, en una región alejada del Oriente ecuatoriano, por ejemplo, se deberá usar una muestra del maíz que usan los agricultores de la zona. Un ensayo de fertilización usando semilla seleccionada o certificada, de otra zona, no beneficiaría a los agricultores actuales, que se encuentran alejados de la ciudad, quienes, por el momento, no tiene acceso o interés en adquirir semillas mejoradas. Si, en forma simultánea, se puede introducir y probar la semilla mejorada, el beneficio para la región sería mayor.

La adecuada selección es más difícil de llevar a cabo en investigación económica o sociológica. Al probar algún método de manejo de haciendas, el

experimento puede basarse en entrevistas a los agricultores y obtener su participación. El éxito de este trabajo dependerá de la habilidad del investigador para obtener una muestra representativa de los agricultores de la zona, a los que beneficiarán las recomendaciones que se desprendan del estudio.

14.14 Precisión relativa del experimento

Se ha mencionado algunos aspectos relacionados con la forma de reducir el tamaño del error experimental y aumentar la precisión del experimento. En cuanto al refinamiento de la técnica y al procedimiento experimental, se puede anotar lo siguiente:

a) Se debe tratar de aplicar los tratamientos en forma uniforme; esto incluye también la toma de datos: si, en un ensayo de fertilización de alfalfa con 4 repeticiones, el número de parcelas es más o menos crecido, y no es posible cortar y pesar todo el ensayo en el mismo día, se debe completar este trabajo en dos o tres repeticiones, dejando la cuarta para el día siguiente. En esta forma, no habrá diferencias entre tratamientos dentro de cada repetición.

b) Se debe tratar de controlar las variaciones externas de tal modo que todos los tratamientos actúen bajo idénticas condiciones. Si la variable en estudio es fungicidas, el único elemento que debe variar de una parcela experimental a otra es tipos y/o dosis de fungicida. Todas las parcelas deben recibir idéntica preparación del terreno, iguales sistemas, densidad y época de siembra y cosecha, labores culturales, riegos, fertilización, la misma variedad, etc., etc.

c) El investigador usa en forma rutinaria una serie de aparatos para la toma de datos; es importante calibrarlos periódicamente a fin de controlar su precisión. Tal es el caso de balanzas de reloj que se usan para efectuar pesajes en el campo: frecuentemente los resortes se estiran o, cuando se guarda las balanzas en lugares húmedos, se oxidan. De todos modos, una vez que se ha comenzado a tomar los datos, se debe usar el mismo equipo o instrumental, por lo menos hasta terminar una repetición completa.

d) Frecuentemente, el investigador deja el trabajo rutinario en manos de asistentes, los que, por más entrenamiento que tengan, carecen de la experiencia del profesional. Es conveniente, entonces, vigilar el trabajo de campo o laboratorio a ciertos intervalos. Esto es especialmente cierto en lo que se refiere a calificaciones en las que se usan escalas relativas de medida: porcentaje de enfermedad, vigor de plántulas, etc.

e) Se ha anotado antes, que la precisión de un experimento está afectada por el número de grados de libertad del error. Los grados de libertad dependen del número de tratamientos, repeticiones y del diseño. El aumento en precisión es importante cuando existen menos de 20 grados de libertad para el error. Como se puede apreciar de los siguientes valores, pasados los 20 grados de libertad, el valor de t, disminuye muy poco.

g. de l.	$t_{.05}$
5	2.57
10	2.23
20	2.09
60	2.00

El procedimiento de Fisher, citado por Cochran y Cox (1957), al comparar la precisión de dos diseños, indica la siguiente ecuación para determinar la Eficacia Relativa de dos diseños:

$$E.R. = \frac{\frac{(n_1+1)}{(n_1+3)}s^2{}_1}{\frac{(n_2+1)}{(n_2+3)}s^2{}_2} = \frac{(n_1+1)(n_2+3)s^2{}_2}{(n_2+1)(n_1+3)s^2{}_1}$$

donde $s^2{}_1$ y $s^2{}_2$ son las variancias del primero y segundo diseños y n_1 y n_2, sus grados de libertad.

f) Finalmente, la precisión de un experimento puede aumentarse, mediante el uso de la técnica llamada covariancia, tal como se vio en el Cap. 13.

14.15 Randomizadón

Debido a la variabilidad del suelo y para evitar que un tratamiento sea favorecido o resulte en desventaja, en forma sistemática, en las varias repeticiones, Fisher ideó la técnica llamada **randomización** o **azarización**. Su función es darnos una estimación válida, insesgada del error experimental, de medias de tratamientos y de sus diferencias, e incluye el uso aleatorio de sistemas, como lanzar una moneda o el uso de tablas de números al azar, similares a los que se emplea para el juego de la lotería.

Las parcelas contiguas tienden a ser correlacionadas, por lo que cualquier parcela debe recibir al azar a cualquier tratamiento.

Los diseños sistemáticos, en los que los tratamientos son aplicados en forma sistemática, generalmente sobre o subestiman el error experimental. La randomización tiende a destruir la correlación entre errores y hacer válidas las pruebas de significación.

Cochran y Cox (1957), anotan que la randomización es similar a un seguro, en que es una precaución que se toma contra problemas que pueden o no pueden ocurrir y, si ocurren, pueden o no pueden ser serios.

El ejemplo más común de randomización estaría dado por el sorteo o rifa de algún objeto, para lo que se distribuyen boletos entre un grupo de personas. Si se coloca papeles o fichas numeradas en un ánfora y se supone que se hallan completamente mezcladas, sin ningún orden ni agrupamiento particular, cualquier secuencia en que salgan, habrá de considerarse al azar. A veces, cuando el investigador tiene pocos tratamientos, recurre a este método. Sin embargo, siendo un sistema exactamente igual, es preferible recurrir a una tabla de números al azar, en las que el proceso de sorteo de cifras ha sido realizado previamente (Tabla N° I del Apéndice).

14.16 Selección del diseño experimental

El número y la naturaleza de los tratamientos a estudiarse determinan el tipo de diseño que debe usarse. Antes de decidir sobre este aspecto, se deberá

considerar aspectos como grado de precisión requerida, disponibilidad de semilla, tierra, equipos, mano de obra y similitud de los tratamientos con las prácticas de uso común por los agricultores. Si, en este último caso, aplicamos tratamientos ideados por nosotros, que incluyen prácticas no usadas o de difícil aplicación por los agricultores, los resultados tendrán utilidad muy limitada.

El diseño de bloques completos al azar es, quizás, el más usado por su flexibilidad, sencillez, tanto en el diseño como en el análisis estadístico. Este diseño funciona bien hasta con 20 tratamientos. Si el número de tratamientos es mayor, deberá usarse otros diseños, incluyendo, posiblemente, experimentos en confundido.

Si las variables dentro del experimento son riego y fertilización, es más fácil regar una parcela grande que una subparcela, en cuyo caso se usaría el diseño de parcelas divididas (Cap. 18).

Muchas veces, en especial en ensayos que se usarán para tesis de grado, se realiza el trabajo experimental sin planificarlo en forma completa y se trata de adaptar los datos obtenidos a un diseño dado. Es obvio suponer los errores que van a producirse en esa forma. El estudiante debe, antes de iniciar el ensayo, poner por escrito los objetivos principales y secundarios del trabajo, el procedimiento y, por lo menos, el esquema del ADEVA, de acuerdo al cual se analizarán los datos.

Esto es también importante en la Estación Experimental, donde, en caso de ausencia del autor del ensayo, otra persona puede continuar el trabajo experimental y analizar los datos sin problemas.

14.17 Inferencia estadística

Como se vio en la Sec. 4.11, la solución estadística a un problema de estimación, consiste en afirmar que la verdadera diferencia entre tratamientos, está comprendida dentro de ciertos límites, indicando la probabilidad de que esa afirmación sea correcta. Un ejemplo ilustrará esta situación:

Queremos determinar si la utilización de un herbicida es ventajosa. El costo del producto es tal, que su uso será económico sólo en caso de que permita obtener rendimientos, en cultivos de alfalfa, de 16 toneladas de forraje fresco por corte y por hectárea. Si, después de llevar a cabo el experimento, se encuentra que los límites de confianza al 95% de probabilidades son 17 y 19 toneladas, estaríamos seguros al usar el herbicida. Pero, si aquellos son 12 y 14 toneladas, descartaríamos el uso del producto. Habría duda en caso de que los límites fueran 14 y 18 toneladas, porque podría haber una pequeña ganancia o una pequeña pérdida. En todo caso, habría riesgo al hacer la recomendación sobre el uso del herbicida. Lo conveniente sería repetir el ensayo.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 15

ANÁLISIS DE VARIANCIA: ORDENACIÓN UNILATERAL

15.1 El análisis de variancia

En el Cap. 10 se presentó los métodos estadísticos que permiten el análisis de datos provenientes de dos muestras o dos tratamientos. En ese capítulo se vio que, por medio de la prueba de t de Student, es posible determinar si la variedad A rinde estadísticamente más que la variedad B; o si alambre eléctrico con aislamiento plástico dura más que alambre sin aislamiento, etc.

Sin embargo, raras veces el trabajo del investigador se limita a estudiar dos tratamientos a la vez. Con el objeto de ganar tiempo, esfuerzo y para aumentar la precisión del experimento, generalmente se estudian varios factores a la vez, incluyendo algunos niveles dentro de cada factor.

Una vez completada la toma de datos dentro de un experimento, el investigador puede observar las medias de tratamientos y pensar que el tratamiento A es superior al tratamiento B. Supongamos, por ejemplo, que de un ensayo de aplicación de cal a parcelas de alfalfa, se ha observado que el testigo rindió ocho toneladas de forraje; con dos toneladas de cal, se cosechó 12 toneladas de forraje y con cuatro toneladas de cal, el rendimiento fue de 14 toneladas de forraje. ¿Sería posible determinar, de la simple observación de estos valo-

res, que existen diferencias significativas entre tratamientos?; ¿diferencias en rendimiento que justifiquen económicamente aplicar dos o cuatro toneladas de cal por hectárea? Obviamente, la respuesta es negativa. Debemos recurrir al análisis estadístico de los datos, para, de acuerdo con sus resultados, establecer si debemos aceptar o rechazar la hipótesis nula de que no existen diferencias entre tratamientos.

Este análisis estadístico es lo que se llama análisis de variancia **(ADEVA)**, que fue introducido por Sir Ronald Fisher y es en esencia, un procedimiento aritmético que consiste en desdoblar una suma de cuadrados total (variación total), en fuentes de variación reconocidas, incluyendo la variación que no se ha podido medir (proveniente de la variabilidad inherente al material experimental o de la falta de homogeneidad del ambiente en el que se realizó el experimento), fuente de variación a la que se conoce como **residuo, resto o error experimental.**

En un diseño de bloques completos al azar (Cap. 16), la fuente de variación total se desdobla en aquella proveniente de la agrupación de los tratamientos en bloques o repeticiones; la variación causada por los tratamientos; y la variación inherente al material experimental, así como la originada por el ambiente heterogéneo en que se llevó a cabo el ensayo. Como se anotó en el párrafo anterior, a esta fuente no controlable, se le conoce como error experimental.

El **ADEVA** es utilizado en todos los campos de investigación, cuando los datos son medidos cuantitativamente, es decir, cuando las observaciones se hallan en forma de números. Su uso ha significado una gran expansión para el diseño experimental y, como veremos adelante, cada diseño tiene su propio ADEVA.

Las suposiciones básicas del **ADEVA**, a partir del cual se efectuarán las pruebas de significación son:

a) Los efectos de tratamientos y los efectos ambientales, son aditivos; y,

b) El error experimental constituye un elemento al azar, normal e independiente distribuido, con media cero y variancia común.

15.2 Grados de libertad

Antes de presentar el primer tipo de diseño, conviene aclarar lo que significa esta expresión. Se llama grados de libertad *(g. de l.)*, al número de comparaciones independientes menos uno, que puede hacerse en un juego de datos. En general, se dice que es el número de comparaciones independientes, menos el número de restricciones impuestas, que puede hacerse en un grupo de datos.

Si suponemos que el rendimiento de cinco variedades de un cultivo (muestra), es de 15, 16, 17, 18 y 19; la media es 17; las desviaciones de los números con respecto a su media (que deben sumar cero), serían,

15,	16,	17,	18,	19
−2	−1	0	+1	+2

Es decir que en esta muestra formada por cinco datos, uno queda fijo (la restricción viene dada por $\overline{X} = 17$), porque $\overline{X}$ ha sido usada como el origen para las desviaciones. Entonces quedan cuatro valores (cuatro grados de libertad), que pueden ser comparados en forma independiente con la media.

O sea que, en general, el número de grados de libertad de una muestra de datos, está dado por $n - 1$, el total de observaciones menos uno. La variancia de la muestra está dada, entonces, por

$$s^2 = \frac{\sum(X - \overline{X})^2}{n - 1} = \frac{\textit{suma de cuadrados corregidos}}{n - 1}$$

La variancia, o cuadrado medio, es el cociente que resulta de dividir la suma de cuadrados (S.C.) para el número de grados de libertad. Dentro del **ADEVA** y para obtener el valor de **F** (que denota la significación entre tratamientos), se divide el cuadrado medio (C.M.) de tratamientos para el cuadrado medio del error. Al comparar el valor de F calculado con el valor de F tabular (Tabla No VI del Apéndice), para el número de grados de libertad de tratamientos y del error, podemos establecer si existe significación o no; habrá significación en caso de que el valor de F calculado, sea mayor que el tabular, al nivel de significación escogido (generalmente del 5 y del 1%). Cuando el valor de F

calculado es mayor que el de F tabular, podemos rechazar la hipótesis nula de que no existen diferencias entre tratamientos y aceptamos la hipótesis alternativa. En este último caso, el siguiente paso sería efectuar las comparaciones entre medias de tratamientos, a fin de conocer cuáles de ellos son superiores.

15.3 Diseño completamente al azar (DCA)

Es el diseño más simple y se le usa cuando las unidades experimentales son homogéneas, cuando la variación entre ellas es muy pequeña. Tal es el caso de experimentos de laboratorio, cámaras de crecimiento, invernadero, en los que las condiciones ambientales son controladas. Esta es una prueba con un solo criterio de clasificación equivalente a la prueba no pareada de *t*.

Entre las ventajas se puede anotar:

a) Es fácil de planear.

b) Es flexible en cuanto al número de tratamientos y repeticiones: el límite está dado por el número de macetas o unidades experimentales, en general.

c) No es necesario que el número de tratamientos sea igual al número de repeticiones.

d) No se calcula parcelas perdidas.

e) Por no tener muchas restricciones, aumenta el número de grados de libertad para el error.

Entre las desventajas:

a) No es eficiente con material heterogéneo.

b) Puesto que no existen restricciones en cuanto a la randomización, el error experimental incluye la variación total entre las unidades experimentales, excepto aquella debida a tratamientos. Es deseable agrupar las unidades

experimentales, de tal modo que la variación entre éstas dentro de cada grupo, sea menor que la variación entre unidades experimentales en diferentes grupos. Esta condición se cumple en el diseño de bloques completos al azar y cuadrado latino, como se verá en el Cap. 16.

Seguidamente se presenta el modelo matemático para el diseño completamente al azar (DCA),

F. de V.	g de l.	S C	C M	F
Total	$rt-1$	$\sum_1^{rt} X^2 - \left(\sum_1^{rt} X\right)^2 \Big/ rt$		
Tratam.	$t-1$	$\sum_1^{t}\left(\sum_1^{r} X\right)^2 \Big/ r - FC^{[1]}$	*SC/g. de l.*	C. M. Trat./ C. M. Error
Error	$t(r-1)$	*Total – tratam.*	*SC/g. de l.*	

$$[1]\ F.C.\,(\textit{Factor de Corrección}) = \left(\sum_1^{rt} X\right)^2 \Big/ rt$$

Donde t = número de tratamientos.
r = número de repeticiones.
SC = suma de cuadrados.
CM = cuadrado medio.

15.3.1 DCA con repetición desigual

A veces se da el caso de que, por insuficiencia de material para todos los tratamientos o, porque se han perdido unidades experimentales para ciertos tratamientos, no se dispone de igual número de observaciones por tratamiento. Esta es una de las ventajas del DCA, para el que se puede analizar los datos directamente, sin tener que calcular parcelas perdidas. El siguiente ejemplo ilustrará el método de cálculo.

CUADRO No 15.1 Rendimiento en g./maceta

REPETICIONES	TRATAMIENTOS				
	1	2	3	4	
I	45	35	34	41	
II	46	33	34	41	
III	49	..	35	44	
IV	44	..	34	43	
V	..	..	33	41	
VI	..	..	..	42	
VII	..	..	..	44	
VIII	..	..	..	41	
IX	..	..	..	41	
	184	68	170	378	$X.. = 800$
$\overline{X}$	46	34	34	42	$\overline{X}.. = 800/20 = 40$

Factor de Corrección $(F.C.) = \left(\sum_{1}^{20} X..\right)^2 / 20$

$= (800)^2/20 = 32.000$

$S.C.(Total): \sum_{1}^{20} X^2 - F.C. = 45^2 + 46^2 + 49^2 + ... + 41^2 + 41^2 - F.C. = 464$

$S.C.(Tratamiento) = \frac{(184)^2}{4} + \frac{(68)^2}{2} + \frac{(170)^2}{5} + \frac{(378)^2}{9} - F.C. = 432$

$S.C.(Error) = Por\ diferencia, 464 - 432 = 32$

ADEVA

F. de V.	g. de l.	S.C.	CM.	F_c	F_t	
					.05	.01
Total	19	464				
Tratam.	3	432	144	72**	3,24	5,29
Residuo (Error)	16	32	2			

El valor altamente significativo de F, nos permite rechazar la hipótesis nula, de que no hay diferencias entre tratamientos; existen diferencias al nivel del 1%. Con este resultado, procederíamos a determinar cuáles tratamientos son

los mejores, usando las comparaciones entre tratamientos, tal como se verá en la Sec. 15.4.

$$\textit{El coeficiente de variación}\ (C.V.) = \frac{\sqrt{C.M.\ Error}}{\overline{X}}(100) = \frac{1.41}{40}(100) = 3.52\%$$

15.3.2 DCA con igual repetición

En un ensayo con macetas, se aplicaron cinco tratamientos, a clones de raigrás, cuatro macetas por tratamiento; los rendimientos fueron los que se anotan en el Cuadro No 15.2. La hipótesis nula propuesta sería:

H_0: *No existe diferencia entre tratamientos.*
H_1: *Existen diferencias entre tratamientos, más allá de lo que puede atribuirse al azar.*

CUADRO No 15.2 Respuesta de clones de raigrás a la aplicación de cinco tratamientos (4 macetas por trat.), en g./maceta

	TRATAMIENTOS					
Maceta	**1**	**2**	**3**	**4**	**5**	
1	101	51	83	67	29	
2	93	61	68	40	45	
3	93	59	72	46	51	
4	96	58	75	52	42	
	$\sum$ 383	229	298	205	167	= 1.282
	$\bar{x}$ 95.75	57.25	74.5	51.25	41.75	

$$FC = \sum (X_i)^2 / rt = (1.282)^2 / 4 \times 5 = 82.176$$

$$SC(tot.) = \sum (X_i)^2 - FC = 101^2 + 93^2 + \ldots + 51^2 + 42^2 - FC$$
$$= 90.334 - 82.176 = 8.168$$

$$SC(trat.) = \sum (X_i)^2 / r - FC = \frac{383^2 + 229^2 + \ldots + 167^2}{4} - FC$$
$$= 89.462 - 82.176 = 7.286$$

$$SC\ (Error) = Por\ diferencia = 8.168 - 7.286 = 882$$

ADEVA

F. de V.	*g. l.*	S C	C M	F
Total	19	8.168		
Trat.	4	7.286	1.821,5	30.98**
Error	15	882	58,8	

La prueba de F, resultante de dividir,

$$F = \frac{C.M.(Tratam.)}{C.M.(Error)} = \frac{1.821,50}{58,8} = 30.98\ **, con\ 4\ y\ 15\ g.\ de\ l.$$

da un valor altamente significativo, que nos permite rechazar H_0 y aceptar H_1: existen diferencias entre tratamientos. En efecto, si comparamos el valor de F calculado con el de F tabular, vemos que, para los niveles de 5 y 1%, los valores tabulares son, 3,06 y 4,98, respectivamente, los que son ampliamente superados por el valor calculado.

Si el cuadrado medio de tratamientos es menor que el cuadrado medio del error, se declara los resultados no significativos. El valor significativo de F implica que los tratamientos no pertenecen a poblaciones con μ común. Nos dice que los tratamientos difieren en forma significativa, pero no indica cuál de ellos es el mejor y cómo difieren unos de otros. Para poder saberlo, apelamos a las comparaciones entre tratamientos, base de algunas de las cuales son valores como el error típico de la media ($s\overline{X}$) y el error típico de la diferencia entre medias de tratamientos ($s\overline{d}$).

En nuestro ejemplo, esos valores son:

$$s\overline{X} = \sqrt{\frac{s^2}{r}} = \sqrt{\frac{58.8}{4}} = \sqrt{14.70} = 3.83$$

$$s\overline{d} = \sqrt{2\frac{s^2}{r}} = \sqrt{2\frac{58.8}{4}} = \sqrt{29.40} = 5.42$$

El coeficiente de variación,

$$C.V. = \frac{\sqrt{s^2\,(Error)}}{\overline{X}}(100) = \frac{\sqrt{58.8}}{61.4}(100) = 11.96\%$$

15.4 Comparaciones entre tratamientos

En el Cap. 10 se vio que la prueba de t de Student, permite la comparación entre dos medias de tratamientos. Esto se desprende claramente de la relación,

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s\overline{d}}$$

Cuando existen varios tratamientos en un experimento, la comparación de pares de tratamientos, en forma separada, puede ser un trabajo tedioso y largo. Seguidamente se presenta algunas comparaciones entre tratamientos, las que tienen como base un principio similar y difieren en el grado de estrictez, en cuanto a declarar más o menos diferencias significativas entre medias de tratamientos. De acuerdo con esto, podemos decir que en orden de mayor a menor liberalidad, es decir de menor a mayor estrictez, podemos ordenar a las pruebas más comúnmente usadas, en la siguiente forma:

Diferencia mínima significativa (DMS), Duncan, Tukey y Scheffé.

15.4.1 Diferencia mínima significativa

La **DMS** es la prueba más fácil de calcular y, probablemente, la más comúnmente usada para comparar pares de medias de tratamientos,

$$DMS = t_{(g.\,de\,l.\,Error)} X s_{\overline{d}}$$
$$= t_{(g.\,de\,l.\,Error)} X \sqrt{2s^2(Error)/r}$$

Para los datos del Cuadro No 15.2, el valor de $t_{.05}$ para 15 *g. de l.* del error, es de 2,131; $t_{.01}$ = 2,947; luego,

$$DMS(_{.05}) = 2.131\sqrt{2(58.8)/4}$$
$$= 2.131(5.42)$$
$$= 11.55$$
$$DMS(_{.01}) = 2.947(5.42)$$
$$= 15.97$$

Las medias entre tratamientos son:

$$\overline{X}_1 - \overline{X}_2 = 95.75 - 57.25 = 38.50;$$
$$\overline{X}_3 - \overline{X}_4 = 74.50 - 51.25 = 23.25$$

Hasta aquí en las dos comparaciones, cada media aparece una sola vez, por lo que se dice que tales comparaciones son independientes u ortogonales; si se quisiera comparar $\overline{X}_5$ con cualquiera otra media, se rompería la ortogonalidad (independencia), que es una suposición básica para el uso de la **DMS**. Sin embargo, esto no quiere decir que no se debe hacer tal comparación: como se verá más adelante, si la comparación ha sido planeada antes de iniciar el experimento y el investigador cree que existe una buena razón para hacerla, se puede usar la DMS en forma válida.

Por cuanto las diferencias encontradas, $\overline{X}_1 - \overline{X}_2 = 38.50$ y $\overline{X}_3 - \overline{X}_4 = 23.25$, sobrepasan el valor de **DMS**, al nivel del 5 y del 1%, (11,55 y 15,97, respectivamente), declaramos que las diferencias son altamente significativas.

Desventajas del uso de la DMS

a) La **DMS** no toma en consideración el número de tratamientos dentro de un experimento; es decir, no ofrece protección cuando el número de medias aumenta, a partir de dos. En este caso, puede darse la situación de que demasiadas medias aparecen como significativas, razón por la que se considera a la **DMS** como una prueba demasiado liberal; la **DMS** da pruebas correctas cuando se compara dos medias de tratamientos.

b) No se debe usar la **DMS**, a menos que el valor de F del ADEVA, dé un valor significativo.

c) Al usar esta prueba, la probabilidad de cometer un error tipo I en comparaciones individuales aumenta.

d) Cuando se hacen comparaciones no ortogonales (una media está involucrada en más de una comparación), puede concluirse erróneamente, que hay mayor cantidad de información de la que realmente existe.

Ventajas del uso de la DMS

a) Es un valor fácil de calcular y de simple utilización.

b) La prueba es válida cuando se hacen comparaciones planeadas de antemano, comparaciones que se identifican con los objetivos del experimento.

c) La DMS da resultados satisfactorios cuando se compara cada una de las medias con el testigo: $\overline{X}_2$ *vs.* $\overline{X}_1$; $\overline{X}_3$ *vs.* $\overline{X}_1$, $\overline{X}_4$ *vs.* $\overline{X}_1$.

15.4.2 Prueba de rango múltiple de Duncan

La ventaja de la prueba consiste en el hecho de que no se necesita que el valor de F sea significativo para poder usarla. Es una prueba más estricta que la DMS y permite comparar todas las medias entre sí, sin restricciones. En el caso del Cuadro No 15.2, al haber cinco medias, es posible efectuar 5(5 – 1)/2 comparaciones. Para aplicar la prueba, se procede en la siguiente forma:

a) Encontrar el valor de $s_{\overline{X}}$,

$$s_{\overline{X}} = \sqrt{(\textit{cuadrado medio del error})/r}, \textit{ en el ejemplo,}$$

$$= \sqrt{\frac{58.8}{4}} = \sqrt{14.70} = 3.83$$

En la Tabla de rangos mínimos de Duncan (RMD), se da los valores respectivos, de acuerdo al número de *g. de l.* del error y a los niveles del 5 y del 1%. En la parte superior de la Tabla, consta el número de medias que se comparan. Para nuestro ejemplo, los valores respectivos para 15 *g. de l.*, al 5% de probabilidad, desde dos hasta cinco medias, constan más abajo. Luego se multiplican estos valores por el de $s_{\overline{X}}$ y se obtiene los rangos mínimos significativos (**RMS**).

Valores para medias	2	3	4	5
RMD	3,01	3,16	3,25	3,31
RMS (RMD $X s_{\overline{X}}$)	11.53	12,10	12,45	12,68

Las diferencias entre medias tienen que sobrepasar el valor de **RMS**, para ser significativas. Cuando se redondea las medias, los valores de **RMS**, deben ser redondeados en igual forma.

b) Se coloca las medias en orden, de menor a mayor.

Tratamiento No:	(5)	(4)	(2)	(3)	(1)
$\overline{x}$:	41,75	51,25	57,25	74,50	95,75

c) Comparaciones entre medias: Teóricamente, se resta el valor más pequeño del más grande y, si su diferencia es mayor que el valor de **RMS** para cinco medias, se le declara significativo; luego, se resta el segundo valor más pequeño del más grande y, si su valor es mayor que el de **RMS** para cuatro medias, se le declara significativo; en esta forma se sigue hasta restar el segundo valor más grande del más grande, es decir,

95,75 – 41,75 = 54,0 > 12,68, *significativo*
95,75 – 51,25 = 44,5 > 12,45 "
.....
51,25 – 41,75 = 9,5 < 11,53, *no significativo*

En la práctica, para ahorrar tiempo, cuando el número de medias es crecido, se comienza restando de la media mayor, el valor de **RMS** para cinco medias: 95,75 – 12,68 = 83,07. Todas las medias menores que 83,07 se declaran significativamente diferentes de (1); es decir, que con esta operación, ya sabemos que el tratamiento (1) es significativamente superior a todos los otros; luego, 74,50 – 12,45 = 62,05; también (3) es significativamente superior a los tratamientos (5), (4) y (2); 57,25 – 12,10 = 45,15; aquí vemos que (2) no es significativamente diferente de (4), pero lo es de (5). Esto lo podemos comprobar restando 57,25 – 51,25 = 6,0, que resulta menor que 11,53, el valor de **RMS** para dos medias; en cambio, 57,25 – 41,75 = 9,5, que es menor que **RMS** para dos medias.

Estos resultados se resumen en la siguiente forma:

Tratamiento No:	(5)	(4)	(2)	(3)	(1)
$\overline{X}$:	41,75	51,25	57,25	74,50	95,75
	d	cd	c	b	a

lo que se interpreta así: dos medias que comparten la misma línea o letra, no difieren estadísticamente, son estadísticamente iguales o no tienen diferencias significativas entre ellas. Cuando las medias van colocadas en columna, la presentación no varía:

Tratamiento No	$\overline{X}$	Diferencia
(5)	41,75	d
(4)	51,25	cd
(2)	57,25	c
(3)	74,50	b
(1)	95,75	a

15.4.3 Prueba de Tukey

Es un procedimiento similar al de la **DMS**, en cuanto se refiere a que es necesario un solo valor para determinar la significación de las diferencias. Es una prueba de gran adaptabilidad y superior a la **DMS**, porque la unidad considerada es el experimento mismo.

La prueba consiste en calcular un valor **D**, que es el producto de $s_{\overline{X}}$ y un factor **Q**, tomado de la Tabla de puntos de porcentaje de rangos mínimos, para el 5 y 1% de probabilidades; y de acuerdo a los *g. de l.* del error y del número de tratamientos involucrados,

$$\mathbf{D} = \mathbf{Q}(g.\ de\ l.\ Error)s_{\overline{X}}$$

En el ejemplo que estamos siguiendo, el valor de Q para 15 *g. de l.* cinco tratamientos y al 5% es de 4,37

$$\mathbf{D} = (4{,}37)(3.83) = 16{,}74$$

Tal como en el caso de la prueba de Duncan, los datos se resumen así,

Tratamiento No:	(5)	(4)	(2)	(3)	(1)
$\overline{X}$:	41,75	51,25	57,25	74,50	95,75
	c	c	c	b	a

Para que dos medias sean declaradas significativas, la diferencia entre ellas debe sobrepasar el valor de **D**.

Como vemos, esta prueba toma en cuenta el número de tratamientos, a diferencia de la DMS: mientras el número de tratamientos aumenta, el valor de **Q** también aumenta. Por esta razón, se considera a la prueba de Tukey más estricta que a la DMS y a la de Duncan. En efecto, mientras que por la prueba de Duncan encontramos cuatro diferencias significativas, por la de Tukey sólo tres lo fueron.

15.4.4 Prueba de Scheffé

Se le considera una prueba más estricta que las anteriores y es de fácil aplicación, puesto que se utiliza los valores de la Tabla de F (la misma que se usa para determinar si existe o no significación entre tratamientos, después del **ADEVA**). El valor de S calculado para comparar medias de tratamientos está dado por:

$$S = (F_0)(s_{\bar{d}}), \textit{en donde}$$
$$F_0 = \sqrt{F(p-1)}$$

donde **F** es el valor tabular de acuerdo al número de *g. de l.* del error y al nivel escogido; y *(p – 1)* el número de tratamientos menos uno. En nuestro ejemplo,

$$F_0 = \sqrt{(3.06)(4)} = \sqrt{12.24} = 3.50$$
$$s_{\bar{d}} = \sqrt{\frac{2s^2}{r}} = \sqrt{\frac{2(58.8)}{4}} = \sqrt{29.40} = 5.42$$
$$s = (3.50)(5.42) = 18.97$$

Como en las pruebas de Duncan y de Tukey, arreglamos el orden de las medias y tenemos:

Tratamiento No:	(5)	(4)	(2)	(3)	(1)
$\overline{X}$:	41,75	51,25	57,25	74,50	95,75
	c	c	bc	b	a

Como vemos, mientras que con la prueba de Tukey los tratamientos (2) y (3) fueron diferentes, de acuerdo a la de Scheffé, no lo son.

Otras pruebas similares, de uso menos común que las anotadas, son la de Student-Newman-Kewl y la de Dunnett. Esta última se usa cuando se quiere comparar todas las medidas de tratamientos contra el testigo o control.

15.5 Modelo Lineal

Una observación X_{ij}, para DCA, está formada por varios componentes,

$$X_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

la media general, más el efecto proveniente de tratamientos, más un elemento al azar, que es el error experimental para cada observación.

Una suposición básica dentro del **ADEVA**, es que este elemento al azar (e) es normal e independientemente distribuido, con $\mu = 0$ y σ^2 común.

La independencia es necesaria para poder estimar μ y σ^2 y la normalidad para poder realizar pruebas de significación. Dependiendo de los efectos de tratamientos τ_i, podemos tener los siguientes modelos:

Modelo I o Fijo. De acuerdo al cual, el conjunto de tratamientos estudiados, constituye toda la población y las inferencias obtenidas de los resultados, son válidas sólo para esa población o experimento. Es decir, los efectos de tratamientos son fijos y $\Sigma\tau_i = 0$. La repetición del experimento incluiría el mismo juego de tratamientos.

Modelo II o al azar. La repetición del experimento incluiría un nuevo juego de τ_i, pero de la misma población de τ_i. Es decir que los τ_i estudiados, constituyen una muestra sacada de una gran población de τ_i y las inferencias obtenidas se aplican a esta última. En cada experimento tendríamos una muestra al azar de la población total.

El Cuadro No 15.3 presenta los cuadrados medios esperados, de acuerdo a los dos modelos.

CUADRO N° 15.3 Cuadrados medios esperados para los Modelos I y II.

F. de V.	g. de l.	Modelo I	Modelo II
Total	$(t \times r) - 1$		
Tratam.	$t-1$	$\sigma^2 + \frac{r\sum \tau_i^2}{t-1}$	$\sigma^2 + r\sigma^2\tau$
Error	$t(r-1)$	σ^2	σ^2

15.6 Análisis de variancia con sub-muestras

En algunos casos es necesario hacer algunas observaciones dentro de la unidad experimental, la unidad a la que se ha aplicado un tratamiento. Estas observaciones se hacen en sub-muestras o unidades de muestreo.

Por ejemplo, en el caso de aplicación de fungicidas a plantas en macetas, podemos estudiar algunas o todas las plantas dentro de una maceta; esta última constituye la unidad experimental. Dos fuentes de variación que contribuyen a la variancia y que se aplica a comparaciones entre medias de tratamientos son:

1.– La variancia entre plantas bajo el mismo tratamiento, es decir entre plantas dentro de macetas.

2.– La variación entre plantas en diferentes macetas bajo el mismo tratamiento, es decir la variación entre macetas dentro de tratamientos.

Los cuadrados medios para los dos tipos de variación se conocen como error de muestreo y error experimental, respectivamente.

En el ejemplo mencionado, una observación X_{ijk} tiene 3 sufijos:

$i = 1,2,3, \ldots .n$ *tratamientos* *(t)*
$j = 1,2,3, \ldots .n$ *repeticiones* *(r) macetas*
$k = 1,2,3, \ldots .n$ *muestras* *(m) submuestras o plantas*

Seguidamente, se presenta el Cuadro de **ADEVA**, el mismo que incluye las fuentes de variación, el cálculo de los grados de libertad y la suma de cuadrados para cada una de ellas. No consta en el Cuadro del cálculo de "macetas" o unidad experimental (que equivale a "huertos" en el ejemplo numérico que sigue), aún cuando es indispensable conocer la suma de cuadrados respectiva, para poder calcular, por diferencia, la suma de cuadrados de la fuente "macetas dentro de tratamientos" y que constituye el error experimental; así como para obtener la suma de cuadrados del error de muestreo. El cuadrado medio de tratamientos se divide para el cuadrado medio del error experimental, a fin de obtener el valor de **F** y establecer si hay diferencias significativas entre tratamientos.

ADEVA

Fuente de variación	g. de l.	Suma de cuadrados
Total	$trm - 1$	$\sum X^2{}_{ij}k - F.C.$ $\quad F.C. = X^2 \ldots / {}_{ij}k$
Tratamientos	$t - 1$	$\sum X^2{}_{i..} / {}_{j}k - FC$
Macetas/Tratamientos (Error Experimental)	$t(r-1)$	*SC macetas – SC tratamientos*
Plantas/Macetas (Error Muestreo)	$tr(m-1)$	*SC Total – SC macetas*

Modelo lineal: $X_{ij}k = \mu + \tau_i + \sum_{ij} + \delta_{ij}k$

donde: μ = *media general*
τ_i = *efecto de tratamientos*
$\sum_{ij}$ = *error experimental*
$\delta_{ij}k$ = *error de muestreo (contribución de plantas dentro de macetas).*

Supongamos que deseamos determinar el contenido de ácido ascórbico de tres variedades de pera, en seis huertos diferentes, usando cuatro frutos por huerto. El Cuadro No 15.4 presenta los valores obtenidos.

CUADRO No 15.4 Contenido de ácido ascórbico en tres variedades de pera, en mg./100 g.

Variedad	Huerto	Fruto				Total Huerto	Total Variedad
		1	2	3	4		
A	1	2.0	2.1	1.9	2.0	8.0	
	2	2.2	2.3	2.2	2.1	8.8	
	3	2.4	2.3	2.2	2.3	9.2	
	4	1.9	1.8	2.2	2.0	7.9	
	5	2.4	2.5	2.3	2.4	9.6	
	6	2.5	2.5	2.4	2.5	10.0	53.5
B	1	2.4	2.3	2.4	2.1	9.2	
	2	2.0	2.1	1.9	1.9	7.9	
	3	2.2	2.2	2.4	2.3	9.1	
	4	2.6	2.5	2.5	2.7	10.3	
	5	1.8	2.0	2.1	1.9	7.8	
	6	2.5	2.6	2.6	2.7	10.4	54.7
C	1	1.8	1.7	1.7	1.9	7.1	
	2	2.0	2.1	1.9	2.1	8.1	
	3	1.6	1.7	1.7	1.7	6.7	
	4	1.9	1.8	1.9	2.0	7.6	
	5	2.0	2.2	2.0	2.2	8.4	
	6	1.6	1.8	1.7	1.6	6.7	44.6
							152.8

Por ser poco práctico tomar todos los frutos de cada huerto, dentro de la zona, hemos seleccionado seis huertos y cuatro frutos, escogidos al azar. Estos frutos constituyen las submuestras.

El cálculo de las sumas de cuadrados y grados de libertad, es como sigue:

$$Factor\ de\ Corrección (F.C.) = (152.8)^2/72 = 324.28$$

$$SC(Total) = (2.0)^2 + (2.1)^2 + (1.9)^2 + \ldots + (1.6)^2 - FC = 6.20,$$

$$con(vhf) - 1\ g.\ de\ l. = (3 \times 6 \times 4) = 72 - 1 = 71\ g.\ de\ l.$$

$$SC(Variedades) = (53.5)^2 + (54.7)^3 + (44.6)^2/24 - FC = 2.53, con$$

$$v - 1\ g.\ de\ l. = 3 - 1 = 2\ g.\ de\ l.$$

$$SC(Huertos) = (8.0)^2 + (8.8)^2 + (9.2)^2 + \ldots + (6.7)/4 - FC = 5.65$$

$$SC(Huertos/Var.), Error\ Experimental = SC(Huertos) - SC(Variedades)$$

$$= 5.65 - 2.53 = 3.12\ con\ v(h-1) = 3(6-1) = 15 g.\ de\ l.$$

$$SC(Frutos/Huertos), Error\ de\ muestreo = SC(Total) - SC(Huertos)$$

$$= 6.20 - 5.65 = 0.55, con\ vh(f-1) = 3 \times 6(4-1) = 54\ g.\ de\ l.$$

El ADEVA sería como sigue:

ADEVA

Fuente de variación	g. de l.	SC	CM	Fc	Ft	
					5%	1%
Total	71	6.20				
Variedades	2	2.53	1.26	6.0+	3.68	6.36
Huertos/Var. (Error experim.)	15	3.12	0.21			
Frutos/huertos (Error muestreo)	54	0.55	0.01			

El valor de F calculado $(F_c) = 1.26/0.21 = 6.0$, es mayor que el valor tabular (F_t) al 5%. En consecuencia, rechazamos la hipótesis nula de que las tres variedades son iguales en su contenido de ácido ascórbico y procedemos a establecer la superioridad de alguna de las variedades, usando la prueba de Rango Múltiple de Duncan.

$$s^2_{\bar{x}} = \frac{0.21}{24};\ s_{\bar{x}} = \sqrt{\frac{0.21}{24}} = 0.094$$

Valores para medias	2	3
RMD	3.01	3.16
RMS = RMD ($s_{\bar{x}}$)	0.282	0.297

Variedad	B	A	C
$\overline{X}$:	2.28	2.22	1.86
	a	a	b

Es decir, no existe diferencia estadística entre las variedades A y B, pero las dos dieron mayor contenido de ácido ascórbico que C, al nivel del 5% de significación.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 16

ANÁLISIS DE VARIANCIA: ORDENACIÓN MÚLTIPLE

Como se ha indicado en el Cap. 15, se usa el diseño completamente al azar (**DCA**), solamente cuando la única fuente de variación que se anticipa, es aquella proveniente del efecto de tratamientos, es decir cuando se da por descontado que las unidades experimentales, cuando están sujetas al mismo tratamiento, responden en igual forma. Tal sería el caso de experimentos llevados a cabo en invernaderos, cámaras de crecimiento, fitotrones, etc., lugares donde existe control sobre las condiciones ambientales.

En numerosos experimentos de campo y de otro tipo, resulta muy difícil obtener uniformidad dentro de las unidades experimentales. En el Cap. 14 se vio que, por ser la heterogeneidad del suelo universal, es necesario restringir la variabilidad del mismo, agrupando los tratamientos en bloques o repeticiones; esto quiere decir que parcelas adyacentes van a responder en forma similar, en tanto que aquellas distantes unas de otras, darán una respuesta diferente.

En el caso de ensayos con animales, aquellos de una misma edad, sexo y pesos similares, darán respuestas uniformes, en contraste con animales de características diversas.

Seguidamente se presenta el diseño de bloques completos al azar (BCA), que es, probablemente, el diseño más común en experimentación, por las ventajas que ofrece.

16.1 Diseño de bloques completos al azar (BCA)

El objetivo del diseño es reunir las unidades experimentales –a las que se aplicará los tratamientos– en bloques de cierto tamaño, de tal modo de que los tratamientos participen por igual de condiciones tan uniformes como sea posible, dentro de cada bloque. La variabilidad entre unidades experimentales de diferentes bloques, será mayor que entre unidades dentro del mismo bloque. En consecuencia, las diferencias que se encuentren entre unidades, se deberá, en su mayor parte, a diferencias entre tratamientos; la diferencia que no se debe a tratamientos, es removida por el diseño y forma parte del error experimental.

De acuerdo con esto, es fácil ver que la variabilidad entre bloques, no afecta a las diferencias entre medias de tratamientos, puesto que en cada bloque aparece una vez cada tratamiento. Por esto se afirma que bloques y tratamientos son ortogonales.

Durante el curso del experimento, todas las parcelas dentro del bloque. deben ser tratadas en igual forma, exceptuando la aplicación de los tratamientos cuyo efecto se trata de medir. Así, si los tratamientos son fertilizantes, todos los otros factores como preparación del suelo, sistema, época y densidad de siembra, labores culturales, uso de pesticidas, riegos, variedad, etc., deben ser exactamente iguales para todas las parcelas. Si no fuera así, estaríamos introduciendo fuentes adicionales de variación, cuyo efecto sería imposible de medir, las que enmascararían el efecto de los tratamientos en estudio. Los tratamientos deben ser randomizados dentro de cada repetición.

Este diseño se caracteriza por el equilibrio que presenta, por ser fácil de planear y cuyo procedimiento de cálculo es simple.

La única desventaja radica en el hecho de que, cuando el número de tratamientos es alto, aumenta la superficie del terreno dentro de cada bloque y, en consecuencia, el tamaño del error experimental.

El Cuadro No 16.1 presenta la nomenclatura comúnmente usada en el caso de ordenación doble, es decir cuando se tiene dos variables como en la situación de **BCA**, con tratamientos y repeticiones.

CUADRO N° 16.1 Ordenación de doble entrada para dos variables.

	TRATAMIENTOS				
Repeticiones	**1**	**2**	**3**	$X_{(i)}$	$\overline{X}_{(i)}$
I	X_{11}	X_{21}	X_{31}	$X.1$	$\overline{X}_{\cdot 1}$
II	X_{12}	X_{22}	X_{32}	$X.2$	$\overline{X}_{\cdot 2}$
III	X_{13}	X_{23}	X_{33}	$X.3$	$\overline{X}_{\cdot 3}$
IV	X_{14}	X_{24}	X_{34}	$X.4$	$\overline{X}_{\cdot 4}$
$X_{(j)}$	$X_{1\cdot}$	$X_{2\cdot}$	$X_{3\cdot}$		$X..$
$\overline{X}_{(j)}$	$\overline{X}_{1\cdot}$	$\overline{X}_{2\cdot}$	$\overline{X}_{3\cdot}$		$\overline{X}..$

En donde el primer subíndice se refiere a la columna y el segundo a la fila; por ejemplo, X_{32} es la observación que corresponde al tercer tratamiento en la segunda repetición. A la derecha del cuadro se escribe $X_{\cdot 1}$, que se define como el total de las X_{ij} para esa fila;

$$\overline{X}_{1\cdot} = \frac{(X_{11} + X_{12} + X_{13} + X_{14})}{4}; \quad \overline{X}_{\cdot 1} = \frac{(X_{11} + X_{21} + X_{31})}{3}$$

La media de los doce valores X_{ij} $(X_{11}, X_{12}, X_{13}, \ldots, X_{34})$ se escribe como $\overline{X}..$

En el Cuadro No 16.2 se presenta los resultados de un ensayo de adaptación de seis variedades de fréjol, en el que se usaron cuatro repeticiones por tratamiento.

CUADRO No 16.2 Prueba de adaptación de seis variedades de fréjol (rendim. expresados en g./parcela).

REPETICIONES

Variedades		I	II	III	IV	Σ	$\overline{X}$
Carate	(6)	42	46	38	41	167	41,75
P.R.	(T)	32	38	31	30	131	32,75
Cal.	(7)	25	32	28	26	111	27,75
R.K.	(9)	18	20	26	24	88	22,00
Max.	(8)	35	42	46	40	163	40,75
C.R.	(10)	36	25	22	26	109	27,25
	Σ	188	203	191	187	769	

$$\text{Factor de Corrección } (F.C.) = \frac{\sum (X_{ij})^2}{rt} = \frac{X^2..}{rt} = \frac{769^2}{6 \times 4} = 24.640$$

$$\text{SC (Total)} = \sum X_{ij}^2 - FC$$
$$= 42^2 + 32^2 + 25^2 + \ldots + 26^2 - FC$$
$$= 26.205 - 24.640$$
$$= 1.565$$

$$\text{SC (Rep)} = \frac{\sum X^2_{\cdot j}}{t} - FC = \frac{188^2 + 203^2 + 191^2 + 187^2}{6} - FC$$
$$= \frac{148.003}{6} - FC$$
$$= 24.667 - FC = 27$$

$$\text{SC (Varied.)} = \frac{\sum X^2_{i\cdot}}{r} - FC = \frac{167^2 + 131^2 + \ldots + 109^2}{4} - FC$$
$$= \frac{103.565}{4} - FC$$
$$= 25.891 - FC = 1.251$$

$$\text{SC (Error)} = SC(Total) - SC(Rep.) - SC(Varied.)$$
$$= 1.565 - 27 - 1.251 = 287$$

ADEVA

F. de V.	*g. de l.*	*S.C.*	*C.M.*	Fc	Ft	
					5%	1%
Total	23	1.565				
Rep.	3	27	9	<1		
Varied.	5	1.251	250.20	13.08**	2.90	4.56
Error	15	287	19.13			

El valor calculado de F (F_c), para probar la hipótesis nula de que no existen diferencias entre variedades, está dado por 250.20/19.13 = 13.08**, con 5 y 15 *g. de l.* Al contrastar este cociente con los valores tabulares (F_t), encontramos que es altamente significativo, puesto que supera no sólo al valor de F al 5% (2,90), sino al valor de F al 1% (4,56). De acuerdo con estos resultados vemos que existe evidencia para rechazar la H_0 y aceptar H_1: existen diferencias reales entre variedades. Las comparaciones entre medias de tratamientos nos permitirán establecer la superioridad estadística de unas variedades sobre otras. Para aplicar la DMS, calculamos el valor de $s_{\bar{d}}$ y encontramos el valor de t, para 15 *g. de l.*

$$DMS_{.05} = (s_{\bar{d}})(t_{.05})$$

$$s_{\bar{d}} = \sqrt{2\frac{(19.13)}{4}} = \sqrt{9.56} = 3.09$$

$$t_{.05} = 2.131 (para\ 15\ g.\ de\ l.)$$

$$DMS = (3.09)(2.131) = 6.58$$

Las variedades entre las que exista una diferencia mayor que 6,58 son estadísticamente diferentes.

Para la prueba de Duncan, vemos los valores de **RMD** para 15 *g. de l.*, al 5%, y los multiplicamos por el de $s_{\bar{X}}$ = 2,18.

Valores para medias	2	3	4	5	6
RMD	3.01	3.16	3.25	3.31	3.36
RMS	6.56	6.88	7.08	7.21	7.32
Tratamiento No: R.K(9)	C.R.(IO)	Cal(7)	P.R.(T)	Max(8)	C(6)
$\overline{x}$: 22,0	27.25	27.75	32.75	40.75	41.75
				a	a
		bc	b		
	c	bc			

Como podemos ver, la variedad testigo (P.R.), ha dado igual rendimiento que la Cal. (7) y C.R. (10), ubicándose aquella en segundo lugar, en cuanto a rendimiento.

El coeficiente de variación, $C.V. = \sqrt{s^2}/\overline{X} = \sqrt{19.13}/32.04 = 14\%$

Con relación a un valor significativo de F para repeticiones, existe el criterio erróneo de que, en ese caso, el ensayo no tiene valor, o de que los resultados son poco confiables. En realidad, un valor de F significativo para bloques, nos está indicando que al usar **BCA** se ha incrementado el grado de precisión en comparación con **DCA**, puesto que ha existido heterogeneidad en el suelo y hemos tenido éxito al estratificar o agrupar los tratamientos en bloques, reconociéndoles como una fuente adicional de variación, la misma que ha sido eliminada por medio del error experimental.

Por otro lado, al existir heterogeneidad en el suelo, ha sido posible probar las variedades, bajo condiciones diferentes y así se puede ampliar el alcance de las inferencias; es decir, las conclusiones servirían para áreas más amplias, que en el caso de que el terreno hubiera sido uniforme.

16.1.1 Modelo lineal aditivo

En el diseño de BCA, una observación X_{ij}, está formada por,

$$X_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

la media general por observación, más el efecto de tratamientos (τ_i), más el efecto promedio de bloques (β_j), más un elemento al azar, que es el error experimental para cada observación. Dependiendo de los efectos de tratamientos y bloques, podemos tener tres modelos:

Modelo I o Fijo. Efecto de tratamientos y bloques fijos, cuando los tratamientos estudiados constituyen toda la población, en cuyo caso, los resultados se aplican al lugar y año en que se llevó a cabo el ensayo.

Modelo II o al azar. Los tratamientos y bloques del ensayo, han sido escogidos al azar de una población grande de tratamientos y bloques. Las conclusiones se aplican a toda esta población de la que fueron seleccionados.

Modelo Combinado. Cuando se supone que los efectos de bloques son al azar y los tratamientos fijos. En este caso, se tendrá el cuidado de ver si los bloques son realmente representativos de una población, sobre la que se obtendrá las inferencias.

El Cuadro No 16.3 presenta los cuadrados medios esperados, de acuerdo a los tres modelos.

CUADRO No 16.3 Cuadrados medios esperados para los Modelos I, II y Combinado.

F. de V.	*g. de l.*	**Modelo I**	**Modelo II**	**M. Combinado**
Total				
Rep.	$r-1$	$\sigma^2 + t\sum\beta^2/r-1$	$\sigma^2 + t\sigma^2\beta$	$\sigma^2 + \sigma^2\beta$
Trat.	$t-1$	$\sigma^2 + r\sum\tau_i^2/t-1$	$\sigma^2 + r\sigma^2\tau_i$	$\sigma^2 + r\sum\tau^2/t-1$
Error	$(r-1)(t-1)$	σ^2	σ^2	σ^2

16.1.2 Cálculo de parcelas perdidas

Cuando se usa el **DCA**, la pérdida de una parcela no afecta al procedimiento de cálculo, el mismo que puede continuarse sin problemas. Con BCA, puede darse el caso de que un ataque de pájaros o insectos, haya destruido una par-

cela completa, o que, en un ensayo con animales, uno de ellos haya muerto. Para tales casos, es necesario estimar el valor faltante, antes de proceder con el **ADEVA**. Para ello, se utiliza la fórmula siguiente, para calcular el valor de una parcela perdida:

$$X = \frac{rB + tT - G}{(r-1)(t-1)}, \text{ donde,}$$

X = *Parcela perdida*
r = *Número de repeticiones*
B = *Total del bloque donde se perdió la parcela*
t = *Número de tratamientos*
T = *Total de la hilera en la que se perdió la parcela*
G = *Gran total general*

Una vez estimado el valor perdido, se llena la celda respectiva con ese valor y se procede con el **ADEVA**, en forma normal, restando un *g. de l.* de los *g. de l.* para total y para el error.

Cuando existen varias parcelas perdidas, se coloca en el cuadro de datos, valores aproximados, menos uno, el cual es calculado usando la fórmula indicada. Luego se procede a calcular el siguiente, hasta completar las celdas faltantes. Una vez completado el cálculo para todas las parcelas perdidas, se resta un *g. de l.* de los *g. de l.* de total y del error, por cada parcela perdida.

Con el conocimiento de la variabilidad del suelo, se podría eliminar las parcelas perdidas y analizar los datos como si se tratara de **DCA**.

En caso de existir parcelas perdidas con BCA, el valor de

$$s_{\bar{d}} = \sqrt{s^2\left[\frac{2}{r} + \frac{t}{r(r-1)(t-1)}\right]}$$

16.1.3 Grado de eficiencia de BCA vs. DCA

Cuando se quiere comparar el grado de eficiencia conseguido con **BCA**, en comparación con **DCA**, se usa la siguiente ecuación, para calcular el error de **DCA**.

$$E_e(DCA) = \frac{n_b E_b + (n_t + n_e)E_e}{n_b + n_t + n_e}$$

donde E_b y E_e son los cuadrados medios para bloques y para el error, y n_b, n_t y n_e, son los *g. de l.* para bloques, tratamientos y para el error. La eficacia relativa (E.R.), de **BCA** en relación a **DCA**, está dada por,

$$E.R. = \frac{E_e(DCA)}{E_e(BCA)} \frac{(n_1 + 1)(n_2 + 3)100}{(n_1 + 3)(n_2 + 1)}$$

donde n_1 son los grados de libertad para BCA y n_2 los grados de libertad para **DCA**.

Si el resultado fuera 80%, esto significaría que 80 repeticiones con **DCA**, son tan eficientes, o dan tanta información, como 100 repeticiones con **BCA**. Si el valor fuera 200%, diríamos que **BCA** es dos veces más eficiente que **DCA**.

En definitiva, para usar **BCA** debemos identificar las fuentes de variación que resultan más notorias con bloques. Es decir, dejaríamos la máxima variación entre bloques y la mínima variación entre unidades experimentales dentro del bloque.

La forma y orientación de los bloques y de las parcelas, tiene importancia. Si se nota heterogeneidad en el suelo, la forma rectangular es preferible para las parcelas, tratando de que el conjunto de ellas formen un cuadrado (el bloque). En esa forma, las parcelas pueden compartir las mismas condiciones de suelo. En el caso extremo de un terreno pendiente, las unidades experimentales irían unas junto a otras, sobre la misma cota, por así decirlo, tal como se vio en la Sec. 14.9.

16.2 Diseño de bloques completos al azar con sub-muestras

En algunos casos, al investigador no le interesa tomar datos de toda la unidad experimental; para ahorrar tiempo y trabajo, o por conveniencia, él puede tomar datos en partes o porciones de la unidad experimental, las que constituirían sub-muestras. En este caso, para el ADEVA, es necesario distinguir entre error experimental, tal como se vio en el ADEVA para el diseño de bloques completos al azar, y el error de muestreo, al que dan lugar las sub-muestras tomadas de la unidad o parcela experimental. Se espera que el error experimental sea de mayor magnitud que el error de muestreo; es decir, que la variación entre unidades experimentales va a ser mayor que la variación entre sub-muestras dentro de la unidad experimental. Para determinar si hay significación entre bloques o tratamientos, se divide el cuadrado medio de estas fuentes de variación para el cuadrado medio del error experimental. El ADEVA pertinente sería:

Fuente de variación	g. de l.	SC	
Total	$trs - 1$	$\sum X^2_{ij}\, k - FC$	$FC = X^2 \ldots /_{ij}\, k$
Rep.	$r - 1$	$\sum X^2_{.j.}/ts - FC$	

Fuente de variación	g. de l.	SC	
Tratam.	$t - 1$	$\sum X^2_{i..}/rs - FC$	
Error exp.	$(r - 1)(t - 1)$	$\sum X^2_{ij.}/s - FC - SC$	$Rep. - SC\ Tratam.$
Muest. / unid. exp. (Error de muestreo)	$rt(s - 1)$	$\sum X^2_{ij}\, k - X^2_{ij.}/s$	

Es decir, que para obtener la suma de cuadrados del error experimental (E_e), se calcula la suma de cuadrados de las unidades experimentales (*UE*) y se resta la suma de cuadrados de repeticiones y de tratamientos; para el cálculo de la suma de cuadrados de la unidad experimental se aplica la fórmula:

$$SC\ (UE) = \sum X^2_{ij.}/s - FC$$

Una vez calculado el error experimental se puede restar de la suma de cuadrados total, la suma de cuadrados de repeticiones, de tratamientos y del error

experimental, para obtener la suma de cuadrados del error de muestreo (*Em*), o aplicar la fórmula directamente:

$$SC\ (Em) = \sum X^2_{ij}\, k - X^2_{ij.}\, /s$$

Para ilustrar el procedimiento de cálculo, supongamos que se ha aplicado dos insecticidas en parcelas de trigo, a fin de controlar el ataque de pulgón. Si A fue el testigo y B y C los dos productos comerciales, hemos cosechado tres sub-parcelas (sub-muestras), dentro de cada parcela o unidad experimental; el diseño usado fue de bloques completos al azar, con seis repeticiones; los datos, registrados como gramos de trigo por sub-muestra, constan en el Cuadro No 16. 4.

CUADRO No 16.4 Respuesta del trigo a la aplicación de dos insecticidas (g./sub-muestra).

Tratamientos	Repeticiones						
	1	2	3	4	5	6	Total Tratam.
A	45	30	17	43	32	53	
	50	40	57	23	39	46	
	39	35	42	11	47	43	
Total UE	134	105	116	77	118	142	692
B	68	50	51	60	85	83	
	40	84	83	64	74	74	
	49	74	66	65	87	64	
	157	208	200	189	246	221	1.221
C	53	54	24	57	76	62	
	46	63	85	33	48	57	
	43	112	88	28	42	59	
	142	229	197	118	166	178	1.030
Total repet.	433	542	513	384	530	541	2.943

$FC = (2.943)^2/54 = 160.393{,}50$

$SC\ (Total) = (45)^2+(50)^2+(39)^2+...+(57)^2+(59)^2–FC = 22.149{,}50$

$SC(Rep.) = (433)^2 + (542)^2 + ... + (541)^2/3–FC = 2.435{,}28$

$SC\ (Tratam.) = (692)^2 + (1.221)^2 + (1.030)^2/18 - FC = 7.973,44$
$SC(UE) = (134)^2 + (157)^2 + \ldots + (178)^2/3 - FC = 12.947,50$

$SC\ (Error\ experimental) = SC\ (UE) - SC\ (Rep.) - SC\ (Tratam.)$
$= 12.947,50 - 2.435,28 - 7.973,44 = 2.538,78$

$SC\ (Error\ de\ muestreo) = \sum X^2_{ij}k - X^2_{ij\bullet}/s$
$= 182.543 - 173.341 = 9.202,00$, *o también por diferencia:*
$SC\ (Total) - SC\ (Rep.) - SC\ (Tratam.) - SC\ (Ee)$
$= 22.149,50 - 2.435,28 - 7.973,44 - 2.538,8$
$= 9.202,00.$

El ADEVA respectivo sería:

Fuente de variación	*g. de l.*	*SC*	*CM*	*F*
Total	53	22.149,50		
Rep.	5	2.435,28	487,06	
Tratam.	2	7.973,44	3.986,73	15,70**
Error experimental	10	2.538,78	253,88	
Error de muestreo	36	9.202,00	255,61	

El valor tabular de $F_{.01}$ *(2,10 g. de l.)* = 7,56; en consecuencia, declaramos los resultados altamente significativos. Para establecer la superioridad de los tratamientos, aplicamos la Prueba de Rango Múltiple de Duncan.

$$S^2_{\bar{X}} = \frac{253.88}{18};\ S_{\bar{X}} = \sqrt{\frac{253.18}{18}} = 3.76$$

Valores para medias	2	3
RMD	3,15	3,30
RMS= RMD ($s_{\bar{X}}$)	11,84	12,40

Tratamiento	B	C	A
$\bar{X}$:	67,83	57,22	38,44
	a	a	b

Es decir que no hubo diferencia estadística entre los insecticidas B y C, pero los dos permitieron obtener mayor rendimiento de trigo, que el que se obtuvo de las parcelas testigo (tratamiento A). La selección de B o C estaría dada por su costo y/o disponibilidad en el mercado.

16.3 Cuadrado latino

En este diseño, la restricción para controlar la variabilidad, está en dos direcciones: filas y columnas. Se lo usa cuando la gradiente de fertilidad en el suelo va en dos direcciones. Se arregla los tratamientos en bloques en dos sentidos, apareciendo cada tratamiento una sola vez en cada fila y en cada columna. El análisis de los datos puede remover del error, la variabilidad debida a diferencias entre filas y columnas.

Este diseño es usado en ganadería, comercialización, industria, agronomía, etc. Así, se puede medir la eficiencia de máquinas lavadoras de ropa y se podría asignar las máquinas como tratamientos, operadores o detergentes como filas y tiempo o temperatura de lavado como columnas. Para ello, debe existir el mismo número de tratamientos, que filas y columnas.

Este diseño resulta adecuado, cuando el número de tratamientos va de cuatro a 12; como en el caso de **BCA**, si el tamaño del bloque aumenta, el error por unidad también aumenta. Varios textos dan cuadrados de 3x3, 6x6, ..., 12x12. Cuando se usa cuadrados pequeños, existe la desventaja de disponer de pocos *g. de l.* para el error; la alternativa sería, en caso de pocos tratamientos, usar más de un cuadrado.

Para cuatro tratamientos, el arreglo podría ser,

A	D	B	C
D	C	A	B
C	B	D	A
B	A	C	D

Si la variación en el campo, o en el invernadero, es en una sola dirección, este arreglo puede cambiarse a,

BDAC	ACBD	DBCA	CADB

16.3.1 Modelo matemático para cuadrado latino

F. de V.	g. l.	S.C.	
Filas	$r-1$	$\sum_i X^2_{i\cdot} \Big/ r - FC$	$FC = \frac{X^2_{..}}{r^2}$
Columnas	$r-1$	$\sum_j X^2_{\cdot j} \Big/ r - FC$	
Tratam.	$r-1$	$\sum_t X^2_{t} \Big/ r - FC$	
Error	$(r-1)(r-2)$	Diferencia	
Total	r^2-1	$\sum_{i,j} X^2_{ij} - FC$	

Ejemplo: Supongamos que tenemos 4 variedades de trigo (rendim. en Kg/Parcela).

					Tot. filas
	B	D	A	C	
	2	3	6	8	19
	A	C	B	D	
	7	5	5	7	24
	D	B	C	A	
	5	4	9	10	28
	C	A	D	B	
	6	9	5	5	25
Tot. Col.	20	21	25	30	96

TOTAL TRATAMIENTOS

$A = 32; \quad B = 16; \quad C = 28; \quad D = 20$

$$FC = \frac{\sum X^2_{..}}{r^2} = \frac{96^2}{4^2} = 576$$

$$SC\ Tot. = \sum_{ij} X^2_{ij} - FC; 2^2 + 3^2 + ... + 5^2 - FC = 74.0$$

$$SC\ Filas = \frac{\sum X^2_{i\cdot}}{r} - FC; 19^2 + 24^2 + ... + 25^2/4 - FC = 11,0$$

$$SC\ Col. = \frac{\sum X^2_{\cdot j}}{r} - FC; 20^2 + 21^2 + ... + 30^2/4 - FC = 16.0$$

$$SC\ Tratam. = \frac{\sum X^2_{t}}{r} - FC; 32^2 + 16^2 + ... + 20^2/4 - FC = 40.0$$

ADEVA

F. de V.	*g. l.*	SC	CM	Fc	Ft	
					5%	1%
Filas	3	11,0	3,67	3,14(NS)	4,76	9,78
Col.	3	16,0	5,33	4,56(NS)		
Trat.	3	40,0	13,33	11,39**		
Error	6	7.0	1,17			
Tot.	15	74,0				

La significación de filas o columnas indican en qué dirección del terreno está la variabilidad.

$$s_{\bar{x}} = \sqrt{\frac{1.17}{4}} = 0.54; \quad s_{\bar{d}} = \sqrt{2\left(\frac{1.17}{4}\right)} = 0.76$$

$CV = \left(\frac{s}{\bar{X}}\right)100 = \frac{1.08}{6} \times 100 = 18.8\%$, que es un valor relativo alto, para C.L.

16.3.2 Parcelas Perdidas

$$X = \frac{r(F + C + T) - 2G}{(r-1)(r-2)}$$

donde, *F, C, T* y *G* corresponden al total de fila, columna y tratamiento donde está la parcela perdida; *G*, el gran total; *r = repeticiones*. Como en **BCA**, este valor sobreestima la *SC* de tratamientos y hay que restar

$$\frac{[G - R - C - (r-1)T]^2}{[(r-1)(r-2)]^2}$$

y para una parcela perdida

$$s_{\bar{d}} = \sqrt{s^2\left[\frac{2}{r} + \frac{1}{(r-1)(r-2)}\right]}$$

Supongamos que se hubiera perdido la observación de la columna 4 y fila 2 (tratamiento D).

Restamos 7 de los tres Ítems y tenemos:

$$F2 = 17;\ C_4 = 23;\ D = 13;\ G = 89$$

$$\frac{4(17 + 23 + 13) - 2(89)}{3 \times 2} = 5.6$$

16.3.3 Eficacia Relativa

Probando columnas como si fueran bloques (ignorando columnas), y como que filas fueran bloques (ignorando filas). El error de BCA sería:

$$Ee(BCA) = \frac{nfEf + (nt + ne)Ee}{nf + nt + ne}$$

donde, Ef = cuadrado medio filas;
Ee = cuadrado medio error;
nf = g. l. filas;
nt = g. l. tratam.
ne = g. l. error

Si existen menos de 20 g. de l. hay que ajustar con:

$$\frac{(n1 + 1)(n2 + 3) \times 100}{(n1 + 3)(n2 + 1)};$$

n1 = gl. error C.L
n2 = gl. error BCA.

Cuando queremos usar filas en lugar de columnas, cambiamos en la primera fórmula *nf, Ef* por *nc* y *Ec*:

$$\frac{3(5.33)+(3+6)1.17}{3+3+6}=2.21$$

Por tener menos de 20 *g. l.* ajustamos:

$$\text{Edicacia Relativa } \frac{Ec(BCA)}{Ec(CL)}\frac{(n1+1)(n2+3)}{(n1+3)(n2+1)}\times 100$$

$$=\frac{2.21}{1.17}\times\frac{7\times 12}{9\times 10}=100=175.77\%$$

Es decir que por tratarse de un suelo heterogéneo (C.V. alto) C.L. ha sido eficiente, pues se necesitan 176 repeticiones en BCA para tener igual cantidad de información que con 100 de C.L.

16.3.4 Modelo Lineal: Una observación

$$x_{ij(i)} = \mu + \alpha_i + \beta + \tau_{(i)} + \varepsilon_{ij}$$

en la que μ es la media; α_i, el efecto de columnas; β_j, el efecto de filas; τ el efecto de tratamientos, y ε_{ij} el error experimental.

ADEVA

F. de V.	*g. l.*	Modelo I	Modelo II
Filas	$r-1$	$\sigma^2 e + r\sum\alpha^2_i/r-1$	$\sigma^2 e + r\sigma^2 f$
Col.	$r-1$	$\sigma^2 e + r\sum\beta^2/r-1$	$\sigma^2 e + r\sigma^2 c$
Trat.	$r-1$	$\sigma^2 e + r\sum\tau^2/r-1$	$\sigma^2 e + r\sigma^2 t$
Error	$(r-1)(r-2)$	$\sigma^2 e$	$\sigma^2 e$

16.4 Transformaciones

Una de las suposiciones básicas para el uso del error combinado en el ADEVA, es que los errores de las variables son independientes y normalmente distribuidos, con la misma variancia.

Hay ocasiones en que esto no sucede; en Entomología, por ejemplo, las parcelas tratadas con insecticidas, van a presentar un amplio rango en la cantidad de insectos o en cuanto al daño producido por ellos. En este caso, las suposiciones de independencia y normalidad no se cumplen. Para poder usar el **ADEVA**, hay

que transformar los datos, en cuyo caso, la prueba de F es mucho más sensitiva para detectar diferencias reales entre tratamientos. En todo caso, se trata de controlar la heterogeneidad del error, llamada heterocedasticidad.

En los ejemplos mencionados, existe dependencia entre medias de tratamientos y sus variancias (distribución de Poisson). Al transformar los datos, se puede contar con poblaciones homogéneas.

Cuando el análisis estadístico de los datos originales da un coeficiente de variación muy alto, o, si se aprecia que los datos tienen un rango amplio, es conveniente graficarlos –dentro de ejes de coordenadas– a fin de tratar de establecer a qué tipo de distribución se ajustan. Si queda descartada la distribucion normal, y parece más bien que se trata de una distribución binomial o de Poisson, es conveniente usar los tres tipos de transformaciones que se mencionan seguidamente y optar por aquella que dé el coeficiente de variación más bajo.

16.4.1 Transformación a raíz cuadrada

Cuando los datos son números enteros pequeños –contaje de número de insectos o huevos de insectos en un área dada– los datos siguen generalmente la distribución de Poisson, en cuyo caso se debe transformar los datos.

El procedimiento para transformar los datos consiste en extraer la raíz cuadrada de cada observación antes de proceder con el ADEVA. Para valores muy pequeños se codifica, añadiendo a cada observación 0.5, en cuyo caso la raíz cuadrada queda como $\sqrt{X + 0.5}$

Se recomienda este tipo de transformación cuando las observaciones están expresadas como porcentajes y éstos van de cero a 20 o de 80 a 100. Si el rango de observaciones va de 30 a 70, generalmente no es necesario transformar los datos.

16.4.2 Transformación logarítmica

Los efectos multiplicativos en la escala original, se convierten en aditivos en la escala logarítmica. Para cantidades cero, se añade uno a cada observación y, en lugar de *log. X* se toma el *log.(X+1)*.

Cuando las variancias son proporcionales al cuadrado de las medias, la transformación logarítmica iguala las variancias. Se usa este tipo de transformación cuando las observaciones son números positivos que tiene un rango amplio.

16.4.3 Transformación arco-seno

Aplicable para la distribución binomial, en la que los datos están expresados como fracciones o porcentajes, en especial cuando éstos presentan un rango amplio. Después del ADEVA, se debe interpretar y comparar los datos transformados, pero no se puede reconvertir a las unidades originales.

En una publicación en la que se expongan los resultados de un ensayo, se deben presentar los datos originales, explicando que, para su análisis, han sido transformados. Puede también presentarse los datos transformados, con la desviación típica de los mismos.

Como se vio en el Cap. 7, la variancia de los resultados observados está dada por $\sqrt{Npq}$, o lo que es igual, $\sqrt{p(1-p)N}$. La transformación de arco-seno, de las proporciones observadas presenta una variancia aproximadamente constante, a menos que $Np < 0.8$; o que $N(1-p) < 0{,}8$. La transformación puede obtenerse usando la tabla de funciones de una proporción p para los valores X y N. Los valores tabulares están en radianes (*1 radián = 180°/3,1416, unidad de medida de ángulos*) y la variancia de las medias transformadas es aproximadamente $1/\left(N+\frac{1}{2}\right)$.

La fórmula para la transformación es

$$\phi = arco-seno\sqrt{\frac{X}{(N+1)}} + arco-seno\sqrt{\frac{(X+1)}{(N-1)}}, \text{ o más}$$

usualmente, ϕ = arco-seno$\sqrt{X}$, de acuerdo con la Tabla IX del Apéndice.

En este caso, las observaciones transformadas resultan aproximadamente normales.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM
SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 17

EXPERIMENTOS FACTORIALES

17.1 Factores, niveles, factoriales

La palabra "factorial" es de uso común en diseño experimental. No es raro oír la expresión "diseño factorial", la cual es incorrecta, puesto que no existe tal clase de diseño. El adjetivo "factorial" se refiere a un modo especial de formar las combinaciones de tratamientos o factores.

Si en un diseño de bloques al azar o parcelas divididas, se usan combinaciones de naturaleza factorial, la designación correcta sería "diseño de bloques al azar o parcelas divididas, en arreglo factorial".

Como se ha indicado en el capítulo precedente, el investigador estudia más de una variable independiente a la vez. Cada una de estas variables independientes, es lo que se conoce como factores. Si se van a probar cinco variedades de alfalfa y tres épocas de corte, el factor variedad tiene cinco "niveles" (v_1, v_2, v_3, v_4 y v_5) y el factor "épocas de corte", tres (e_1, e_2 y e_3). La combinación factorial de las dos variables estaría representada así:

$$
\begin{array}{llll}
 & e_1(v_1\ e_1) & & e_1(v_3\ e_1) \\
v_1 & e_2(v_1\ e_2) & v_3 & e_2(v_3\ e_2) \\
 & e_3(v_1\ e_3) & & e_3(v_3\ e_3)
\end{array}
$$

$$
\begin{array}{ll}
 & e_1(v_5\ e_1) \\
v_5 & e_2(v_5\ e_2) \\
 & e_3(v_5\ e_3)
\end{array}
$$

$$
\begin{array}{llll}
 & e_1(v_2\ e_1) & & e_1(v_4\ e_1) \\
v_2 & e_2(v_2\ e_2) & v_4 & e_2(v_4\ e_2) \\
 & e_3(v_2\ e_3) & & e_3(v_4\ e_3)
\end{array}
$$

Una repetición completa tendría, así, 15 tratamientos.

En esta forma estaríamos cumpliendo con un principio básico de diseño experimental y que se refiere, en el ejemplo descrito, a que –para poder obtener una conclusión válida– todas las cinco variedades se hallen sujetas a los mismos tratamientos de corte.

Si queremos probar tres factores, N, P y K, en dos dosis cada uno, existen cinco formas en que se puede representar las notaciones respectivas, como se indica en el Cuadro No 17.1.

CUADRO No 17.1 Notaciones para representar las combinaciones de N, P y K, con dos niveles cada uno.

Combinaciones de Tratam.	MÉTODO				
	1	2	3	4	5
1.	$n_1p_1k_1$	111	$n_0p_0k_0$	000	(1)
2.	$n_1p_1k_2$	112	$n_0p_0k_1$	001	c
3.	$n_1p_2k_1$	121	$n_0p_1k_0$	010	b
4.	$n_1p_2k_2$	122	$n_0p_1k_1$	011	bc
5.	$n_2p_1k_1$	211	$n_1p_0k_0$	100	a
6	$n_2p_1k_2$	212	$n_1p_0k_1$	101	ac
7.	$n_2p_2k_1$	221	$n_1p_1k_0$	110	ab
8.	$n_2p_2k_2$	222	$n_1p_1k_1$	111	abc

En la ilustración de las cinco variedades de alfalfa, el arreglo factorial se conoce como 5 x 3 factorial. En el último ejemplo, se trataría de un factorial 2 x 2 x 2 o 2^3.

Como se ve, un factor es una clase de tratamiento, el que puede tener varios subtratamientos, llamados "niveles".

En un ensayo de alimentación de terneros, los factores pueden ser balanceados y los niveles, cantidades o "dosis", del producto.

Se usan los experimentos factoriales en trabajos exploratorios, cuando no se conoce mucho sobre los niveles adecuados de un factor. Supongamos que se va a probar una nueva variedad de trigo en una región dada. En este caso, sería interesante probar, además, densidad y épocas de siembra; como se anotó en el Cap. 14, en situaciones similares es conveniente incluir testigos locales en el ensayo. En consecuencia, podemos planear un experimento que incluya, variedades, densidades y épocas de siembra.

Interacción. En el ejemplo propuesto, no sólo vamos a conocer la respuesta al factor "variedad", aparte de "densidad" y "épocas", sino que vamos a determinar la interacción entre los factores, es decir la consistencia de la respuesta de varios niveles de un factor, frente, o cuando son comparados con diferentes niveles de otro factor.

Cuando ya existe alguna información previa, es preferible seleccionar un número limitado de combinaciones de varios factores, usando niveles específicos (ver Sec. 14.6 y 17.5).

Podemos ver, entonces, que con el uso de factoriales el alcance del experimento y las inferencias que pueden obtenerse sobre una población, aumentan notablemente.

En cuanto a eficacia, es preferible combinar tratamientos antes que analizarlos solos, porque existe una forma inherente de repetición en los experimentos factoriales. Además, aumenta el número de grados de libertad para el error,

lo cual incrementa la precisión del ensayo, para las pruebas de significación. Generalmente, los factores se identifican con letras mayúsculas y los niveles con minúsculas.

Si tenemos cuatro variedades de un cultivo con seis repeticiones y por otro lado, dos insecticidas, con siete repeticiones, los ensayos aislados y la combinación en un factorial, darían los siguientes esquemas de **ADEVA**:

F. de V.	*g. de l.*	*F. de V.*	*g. de l.*	*F. de V.*	*g. de l.*
Total	23	Total	13	Total	31
Rep.	5	Rep.	6	Rep.	3
Varied.	3	Insect.	1	Varied.	3
Error	15	Error	6	Insect.	1
				V X 1	3
				Error	21

En el primer caso tendríamos un total de 24 parcelas, 14 en el segundo y 32 en el tercero. Es decir que necesitaríamos de 38 parcelas en dos experimentos separados para conseguir la información que nos dan 32 parcelas. Por otro lado, además de estudiar la interacción *V* x *I*, en el último caso, incrementamos el número de *g. de l.* del error a 21.

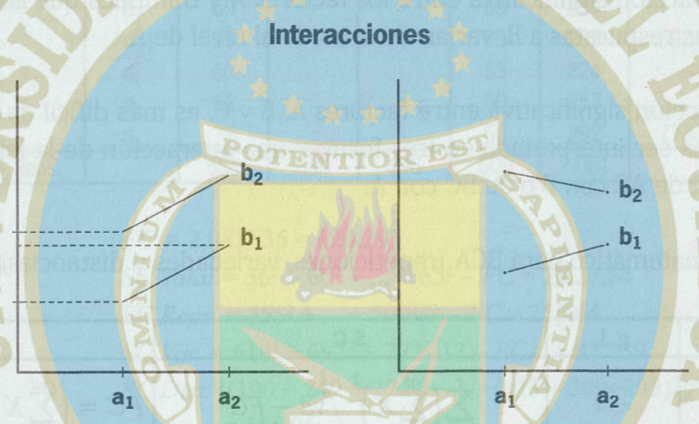
Cuando el número de tratamientos es muy alto, dentro de un experimento factorial, podríamos recurrir a un experimento en confundido, como se indica en Sec. 16.5.

A continuación se ilustra sobre los efectos principales y simples, en el caso de dos factores (A y B), cada uno con dos niveles.

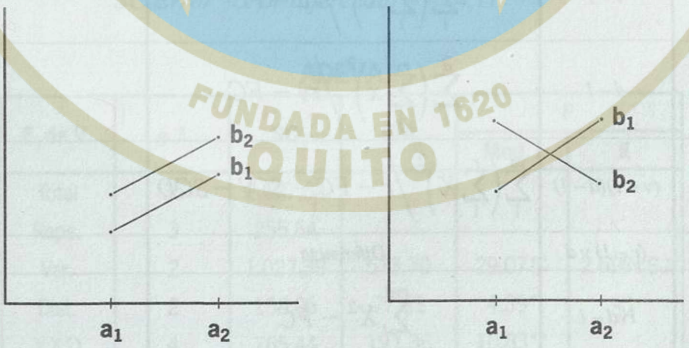
		a_0	a_1		
B	b_0	$90(a_0 b_0)$	$-(a_1 b_0)108$	18	(Efecto simple)
	b_1	$138(a_0 b_1)$	$-(a_1 b_1)144$	6	
		48	36	24	(Efecto principal)
84					

$$
\begin{aligned}
A &= 18 + 6 = 24 \\
B &= 48 + 36 = 84 \\
AB &= 6 - 18 = -12 \\
&= 36 - 48 = -12
\end{aligned}
$$

Los efectos principales son positivos y las interacciones negativas. Los tres grados de libertad y sumas de cuadrados para los cuatro tratamientos en el factorial de 22, indicado arriba, pueden ser descompuestos en grados de libertad individuales independientes al igual que las sumas de cuadrados. La interpretación es fácil y puede dar información valiosa.



La interacción puede ser una diferencia en la magnitud de la respuesta.



La interacción puede ser la diferencia en la dirección de la respuesta.

La presencia o ausencia de efectos principales no nos dice nada sobre la interacción. La interacción significativa nos indica que los factores no son independientes, sino más bien que actúan recíprocamente.

En este caso, se debe estudiar los datos con cuidado, e inclusive, profundizarlos en un estudio posterior.

La interacción de dos factores se llama interacción de primer orden y la de tres factores, interacción de segundo orden.

Una interacción significativa entre los factores A y B implica que las diferencias de las respuestas a B, varían de acuerdo al nivel de A.

La interacción significativa entre factores A, B y C, es más difícil de interpretar. Podría ser interpretada en tres formas: una interacción de la interacción AB con C; de AC con B o de BC con A.

Modelo matemático para BCA (repeticiones, variedades y distancias).

F. de V.	g. l.	S C	
Reps.	$r-1$	$\sum_{i}^{r}\left(\sum_{1}^{vd} X\right)^2 \Big/ vd - FC$	$FC = \left(\sum_{1}^{rvd} X\right)^2 \Big/ rvd$
Varied.	$v-1$	$\sum_{i}^{v}\left(\sum_{1}^{rd} X\right)^2 \Big/ rd - FC$	
Dist.	$d-1$	$\sum_{i}^{d}\left(\sum_{1}^{rv} x\right)^2 \Big/ rv - FC$	
V X D	$(v-1)(d-1)$	$\sum_{1}^{vd}\left(\sum_{1}^{r} X\right)^2 \Big/ r - FC - SCV - SCD$	
Error	$(r-1)\ v\ d$	*Diferencia*	
Total	$rvd-1$	$\sum_{1}^{rvd} X^2 - FC$	

Ejemplo:

CUADRO N° 17.2 Rendimiento de tres variedades, sembradas a tres distancias entre surcos.

REPETICIONES

Variedad	Dist.	I	II	II	IV	Trat.	Var.
V_1	d_1	56			46	190	
	d_2	60			48	203	
	d_3	66			50	223	616
V_2	d_1	65			63	249	
	d_2	60			60	234	
	d_3	53			55	209	692
V_3	d_1	60			53	224	
	d_2	62			60	257	
	d_3	73			65	292	773
∑ Reps.		555	530	496	500	2.081	

$FC = 2.081^2/36 = 120.293$

$SC\ Total = 56^2 + 60^2 + ... + 65^2 - FC = 2.627,64$

$SC\ Reps. = 555^2 + ... + 500^2/9 - FC = 255,64$

$SC\ Var. = 616^2 + 692^2 + 773^2/12 - FC = 1.027,39$

$SC\ Dist. = [(190 + 249 + 224)^2 + ... + (223 + 209 + 292)^2]/12 - FC = 155,06$

$SC\ V \times D = 190^2 + 203^2 + ... + 292^2/4 - FC = 765,44$

$SC\ Error = Por\ diferencia = 424,11$

ADEVA

F. de V.	g. l.	SC	CM	F	
				Mod. I	II
Total	35	2.627,64			
Reps.	3	255,64			
Var.	2	1.027,39	513,70	29.07**	2.69(N.S.)
Dist.	2	155,06	77,53	4.39*	<1
V X D	4	765,44	191,36	10.83**	
Resto	24	424,11	17,67		

17.2 Modelo Lineal

Con dos factores, una observación

$$X_{ijk} = \mu + \rho_i + \alpha_j + \beta_k + (\alpha + \beta)_{jk} + \varepsilon_{ijk}$$

Donde μ es el efecto de la media por observación; ρ el de repeticiones; α_j del factor A; β_k del factor B; $(\alpha + \beta)_{jk}$ de la interacción AxB y ε_{ijk} un elemento al azar (error experimental).

ADEVA

F. de V.	g. de l.	Modelo I	Modelo II
Rep.	$r-1$	$\sigma^2+vd\sigma^2/r-1$	$\sigma^2+vd\sigma^2R$
Var.	$v-1$	$\sigma^2+rd\sigma^2V/v-1$	$\sigma^2+r\sigma^2VxD+rd\sigma^2V$
Dist.	$d-1$	$\sigma^2+rv\sigma^2D/d-1$	$\sigma^2+r\sigma^2VxD+rv\sigma^2D$
VxD	$(v-1)(d-1)$	$\sigma^2+r\sigma^2VxD/(v-1)(d-1)$	$\sigma^2+r\sigma^2VxD$
Error	$(vd-1)(r-1)$	σ^2	σ^2

En el Modelo II (al azar) vemos que la interacción VxD es un componente del cuadrado medio de V y D. Por ser significativa, podemos adoptarle como término de error para V y D, sin tomar en cuenta el error común.

Del ADEVA anterior, podemos observar que el Modelo I es más liberal, puesto que V, D y VxD resultaron valores significativos. De acuerdo al Modelo II, ninguno de los tres lo fueron.

En general, puede decirse que, cuando de acuerdo al Modelo II, el valor de *F* para la interacción de dos (o más) factores es significativa, debe usarse este valor como término de error (en lugar del error común), para probar la significación de esos factores y de las interacciones de orden más bajo que la que se está usando.

Si ese valor no es significativo, se puede probar los factores contra los cuadrados medios (VxD) + E, sumando los cuadrados medios de la interacción más el del error. En esta forma se obtiene mayor precisión en las comparaciones.

17.3 Comparaciones de grados de libertad individuales

La prueba de F con más de un grado de libertad para el cuadrado medio del numerador, puede ser desdoblada en tantas comparaciones independientes como grados de libertad existen.

Si solo una de las comparaciones da una diferencia real, y esta diferencia es promediada con otras que no son reales, una prueba con ese promedio podría no detectar la diferencia verdadera. Por esta razón, planeamos comparaciones que nos den la información deseada y que sean independientes.

En muchos experimentos con más de dos tratamientos, generalmente se planea comparaciones de grados de libertad individuales entre tratamientos. Así, una suma de cuadrados de tratamientos con $t-1$ grados de libertad, puede ser desdoblada, dándonos $t-1$ comparaciones individuales, cada una con un grado de libertad. Cada una de estas comparaciones separadas se la divide para el error experimental. La suma de las sumas de cuadrados de las $t-1$ comparaciones ortogonales debe ser igual a la suma de cuadrados de tratamientos.

Si tenemos $T_1, T_2, \ldots T_t$ tratamientos, la función lineal está dada por:

$$M = \sum_i C_i T_i, \quad donde \sum_i C_i = 0$$

Las C son constantes numéricas, íntegros y es esencial que su suma sea 0.

La SC para cada comparación individual está dada por

$$SC = \frac{M^2}{K_r}, \quad donde \; K = \sum C_i^2$$

$$r = rept.$$

Vamos a usar los datos del Cuadro No 16.2, para llevar a cabo las comparaciones de grados individuales de libertad respectivas. Las variedades probadas fueron P.R. (Testigo), Carate, Cal., R.K., Max. y C.R., a las que, para facilitar la explicación, vamos a identificar con las letras A, B, C, D, E y F, respectivamente.

Por tratarse de seis variedades, tendremos cinco comparaciones, cada una con un *g. de l.*; una comparación lógica es *Testigo vs. resto,* por cuanto nos interesa conocer si existe diferencia entre la variedad local y las introducidas. Los totales de tratamientos y las comparaciones, constan en el Cuadro No 17.3.

CUADRO N: 17.3 Comparaciones ortogonales entre seis variedades de fréjol (1).

Tratamiento	A	B	C	D	E	F	M	kr	SC
Total Tratm.	**131**	**167**	**111**	**88**	**163**	**109**			
Comparación									
A vs. Resto	–5	+1	+1	+1	+1	+1	–17	30(4)	2,40
BC vs. DEF	0	–3	–3	+2	+2	+2	–114	30(4)	108,10
B vs. C	0	–1	+1	0	0	0	–56	2(4)	392,00
D vs. EF	0	0	0	–2	+1	+1	+96	6(4)	384,00
E vs.F	0	0	0	0	–1	+1	–54	2(4)	364,50
								Total	1.251,00

(1) Datos del Cuadro N° 16.2.

La suma de los coeficientes en cualquiera de las filas da cero; al igual que la suma de productos cruzados de los coeficientes de cualquier par de comparaciones.

Se puede cambiar los coeficientes de cualquier fila sin que esto cambie las sumas de cuadrados. Para la primera comparación, tenemos,

$$M = (-5)(131)+(1)(167)+(1)(111)+(1)(88)+(1)(163)+(1)(109) = (-655+638) = -17$$

El valor de $k = (-5)^2 + (1)^2 + \ldots + (1)^2 = 30$

El valor de $r = 4$ = *número de repeticiones.*

La suma de cuadrados está dada por,

$$SC = \frac{M^2}{kr}, \text{ en nuestro ejemplo,}$$

$$= \frac{(-17)^2}{30(4)} = 2.40$$

La suma total de las sumas de cuadrados de todas las comparaciones individuales debe reproducir la suma de cuadrados para tratamientos del **ADEVA**. En efecto, esas cantidades iguales son 1.251,00, como se desprende del **ADEVA** de los datos del Cuadro 16.2. Puesto que cada comparación tiene un *g. de l.* la suma de *g. de l.* de todas las comparaciones debe ser igual también al total de los *g. de l.* para tratamientos del **ADEVA**.

Para obtener los cuadrados medios, dividimos cada suma de cuadrados para uno y, como en el caso del cuadrado medio de tratamientos, los dividimos para el cuadrado medio del error del **ADEVA**, a fin de establecer que comparaciones son significativas. La comparación es, en cada caso, con 1 y 15 *g. de l.*

Para los datos analizados, tenemos,

F. de V.	*g. de l.*	S. de C.	C.M.	Fc	Ft	
					5%	1%
Variedades	5	1.251,00	250,20	13,08**	2,90	4,56
A vs. Resto	1	2,40	2,40	< 1(N.S.)	4,54	8,68
BC vs. DEF	1	108,10	108,10	5,64*		
B vs. C	1	392,00	392,00	20,48**		
D vs. EF	1	384,00	384,00	20,07**		
E vs. F	1	364,50	364,50	19,52**		
		1.251,00				
Error	15	287,00	19,13			

Para la comparación A vs. resto. Fc = 2,40/19,13 = <1, valor no significativo, puesto que no supera ni el valor tabular al 5%, que es 4,54. El valor tabular se lo encuentra en la Tabla de *F*, para un *g. de l.* en el numerador (fila al tope superior) y 15 *g. de l.* para el denominador (columna izquierda).

Como se puede ver, todas las comparaciones entre variedades dieron valores altamente significativos, excepto la comparación BC vs. DEF, que fue significativa (5%). Esto nos permite rechazar la H_0 de que no existen diferencias en el rendimiento de las variedades. La única excepción está dada por la compara-

ción *A vs. resto,* es decir *testigo vs. resto*, para la que no hubo significación; en otras palabras, la variedad testigo dio un rendimiento estadísticamente igual que algunas de las otras variedades.

Se puede apreciar que el desdoblamiento de los grados de libertad ha sido una prueba capaz de detectar diferencias individuales. De no haber aplicado el procedimiento, toda la información existente se hubiera limitado a establecer que había diferencias altamente significativas entre variedades pero sin determinar la contribución de cada una a la suma de cuadrados total para tratamientos, es decir, sin establecer el comportamiento de cada variedad.

Los valores de la variedad testigo, promediados con los rendimientos de las otras variedades, enmascaraban los resultados reales.

17.4 Respuesta superficial; grados de libertad individuales: tratamientos igualmente espaciados

Se llama respuesta superficial al experimento planeado para detectar la forma como responden los niveles de un factor, cuando son comparados con varios nivelas de otro factor. Tal sería el caso al aplicar varias dosis de un fungicida, frente a diferentes niveles de humedad ambiental, en un invernadero; o de usar niveles de N frente a dosis de P. En estos experimentos, el investigador busca el punto máximo de la respuesta al factor estudiado, puesto que ya sabemos que, en general, el tipo de respuesta en experimentación no es lineal sino curvilineal. Para esta situación, es posible aplicar el análisis de regresión, tal como se vio en el Cap. 12. Sin embargo, cuando el incremento o intervalo entre los niveles de un factor, es igual, se puede llevar a cabo el análisis de los datos en forma más sencilla, usando los coeficientes de polinomios ortogonales.

Los polinomios ortogonales son ecuaciones de regresión en las que cada una está asociada con un exponente de la variable independiente: X, X^2, X^3, y todas son independientes unas de otras, es decir, son ortogonales.

El Cuadro N° 17.4 presenta los coeficientes y valores de k ($k = \sum C_i^2$), para comparaciones ortogonales hasta cinco tratamientos igualmente espaciados.

CUADRO N° 17.4 Coeficientes y valores de k para cinco tratamientos igualmente espaciados.

N° de Tratamts.	Grado de polinomio	Totales de tratamientos					
		t_1	t_2	t_3	t_4	t_5	$k=\sum C_j^2$
2	1	−1	+1				2
3	1	−1	0	+1			2
	2	+1	−2	+1			6
4	1	−3	−1	+1	+3		20
	2	+1	−1	−1	+1		4
	3	−1	+3	−3	+1		20
5	1	−2	−1	0	+1	+2	10
	2	+2	−1	−2	−1	+2	14
	3	−1	+2	0	−2	+1	10
	4	+1	−4	+6	−4	+1	70

Estos polinomios ortogonales no se aplican para tratamientos que no son igualmente espaciados (Ver Sec. 17.3).

Cuando estudiamos cinco niveles de un factor, existen cuatro *g. de l.*, los que pueden ser desdoblados en uno asociado con la respuesta lineal; otro con la respuesta cuadrática, o respuesta de segundo grado; uno con la respuesta cúbica o de tercer grado; y otro con la respuesta cuártica o de cuarto grado. Para más de cuatro niveles, generalmente son de interés las respuestas lineal y cuadrática y, a veces, la cúbica. Como en el caso de comparaciones de *g. de l.* individuales para tratamientos desigualmente espaciados, la suma de las sumas de cuadrados de las comparaciones independientes, debe ser igual a la suma de cuadrados del factor desdoblado. Cada uno de los cuadrados medios individuales es contrastado con el del error experimental del **ADEVA** original, para encontrar significación.

Si sólo la función lineal es significativa, concluimos que la respuesta a varios niveles de un factor es constante, es decir que con cada incremento del factor, obtenemos un aumento adicional en el otro factor. La respuesta puede ser positiva o negativa, de acuerdo a si la respuesta del segundo factor aumenta o disminuye en relación con el primero.

La función cuadrática significativa, indicaría que la respuesta del segundo factor no es constante, es decir que los datos se ajustan a una parábola, antes que a la línea recta. En la siguiente ilustración, se desdobla los *g. de l.* individual, dentro de un ensayo tendiente a medir el efecto de la aplicación de cinco niveles de azufre, (s), en parcelas de trébol blanco. El Cuadro N° 17.5 presenta los datos respectivos.

CUADRO No 17.5 Respuesta del trébol blanco a la aplicación de cinco niveles de azufre (escala relat. 1-10).

REPETICIÓN

Nivel (Kg. | Ha.)	I	II	III	IV	Σ	$\overline{X}$
0	2	1	4	3	10	2,50
5	3	2	5	4	14	3.50
10	4	3	6	4	17	4,25
15	6	8	4	5	23	5,75
20	8	7	6	4	25	5,00
Σ	23	21	25	20	89	

ADEVA

F. de I.	*g. l.*	SC	CM	Fc	Ft	
					5%	1%
Total	19	71				
Rep.	3	3	1			
Trat.	4	38,75	9,68	3,98*	3,26	5,41
Error	12	29,25	2,43			

Como vemos, existen diferencias significativas para niveles de azufre. Vamos a proceder a desdoblar la suma de cuadrados, usando polinomios ortogonales. El Cuadro No 17.6 presenta los coeficientes, sumas de cuadrados y valores de F respectivos.

CUADRO No 17.6 Polinomios ortogonales para cinco tratamientos.

Función	Niveles de S y Respuesta (Totales)					M	kr	SC	F
	0 (10)	5 (14)	10 (17)	15 (23)	20 (25)				
Lineal	–2	–1	0	+1	+2	39	10(4)	38,02	15,64**
Cuadrát.	+2	–1	–2	–1	+2	–1	14(4)	0,02	<1
Cúbica	–1	+2	0	–2	+1	–3	10(4)	0,22	<1
Cuártica	+1	–4	+6	–4	+1	–21	70(4)	0,43	<1
							Total	38,69	

Por cuanto existen cinco niveles, podemos desdoblar los *g. de l.* y sumas de cuadrados, en cuatro funciones: lineal, cuadrática, cúbica y cuártica. El procedimiento es similar al descrito en la Sec. 17.3.

Para la función lineal tenemos:

$$M = -2(10) - 1(14) + 0(17) + 1(23) + 2(25) = 39$$

La suma de cuadrados para esta comparación está dada por,

$$\frac{M^2}{kr} = \frac{(39)^2}{[(-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2]4} = 38,02$$

Para la función cuadrática.

$$M = +2(10) - 1(14) - 2(17) - 1(23) + 2(25) = -1$$

la SC:

$$\frac{M^2}{kr} = \frac{(-1)^2}{[(+2)^2 + (-1)^2 + (-2)^2 + (-1)^2 + (2)^2]4} = 0.02$$

Para la función cúbica:

$$M = -1(10) + 2(14) + 0(17) - 2(23) + 1(25) = -3$$

la SC:

$$\frac{M^2}{kr} = \frac{(-3)^2}{[(-1)^2 + (2)^2 + (0)^2 - (2)^2 + (1)^2]4} = 0.22$$

y para la función cuártica:

$$M = 1(10) - 4(14) + 6(17) - 4(23) + 1(25) = -21$$

la SC:

$$\frac{M^2}{kr} = \frac{(-21)^2}{[(1)^2 + (-4)^2 + (6)^2 + (-4)^2 + (1)^2]4} = 0.43$$

Como vemos, la suma de las sumas de cuadrados de las comparaciones individuales reproduce con una pequeña diferencia por redondeo, la suma de cuadrados para tratamientos del ADEVA. Por otro lado, puede apreciarse que el 98% de la suma de cuadrados para tratamientos corresponde a la función lineal, que es la única significativa. En efecto, al dividir la suma de cuadrados que es 38,02 para 1 *(1 g. de l.)* obtenemos el cuadrado medio y este valor dividido, a su vez, para el cuadrado medio del error, nos da el valor de F.

$F = \frac{38.02}{2.43} = 15.64$ **, con 1 y 12 *g. de l.* valor altamente significativo.

De acuerdo con estos resultados, podemos concluir que para cada incremento en la cantidad de *S* aplicada, la población de trébol blanco aumenta en forma constante, hasta llegar al nivel de 15 *Kg.* de *s* por *Ha.* Sin embargo, de la Fig. 17.1 se desprende que, con la aplicación de 20 *Kg.* por *Ha.*,

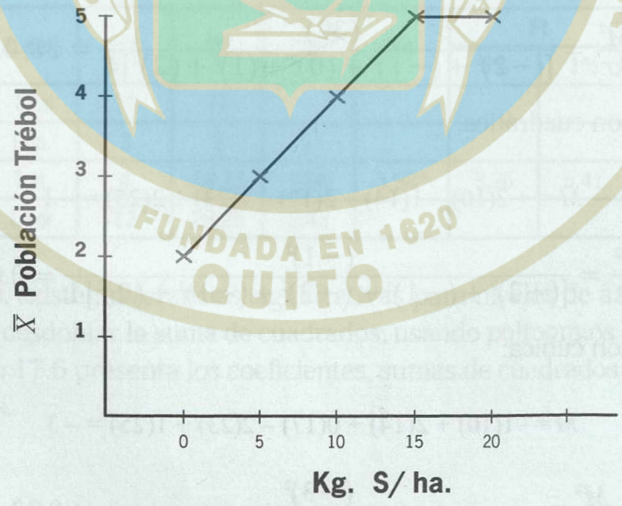


Fig. 17.1 Respuesta del t. b. a la aplicación de 5 dosis de s.

se presenta un pequeño descenso en el rendimiento, graficado por el quiebre de la línea. Esto sugeriría que es imperativo diseñar un segundo experimento, descartando quizás, los niveles O y 5 Kg. de S y aumentando dos que equivalgan a 25 y 30 Kg./Hg, con el objetivo de ver si la curva de respuesta continúa declinando a partir de 20 Kg.

17.4.1 Respuesta superficial; "diseño" de diamante doble

Cuando se prueban niveles de un factor, frente a niveles de otro factor, puede usarse el arreglo de diamante doble, que algunos autores llaman "diseño" de diamante doble. Como se ilustra en la Fig. 17.2, para que pueda llamarse arreglo de diamante doble, dos ejes de los rombos que conforman la respuesta superficial, deben tocar la abscisa y la ordenada, del sistema de coordenadas. De otro modo, sería una modificación del arreglo,la que desde luego, puede ser también usada, de acuerdo a los objetivos del experimento.

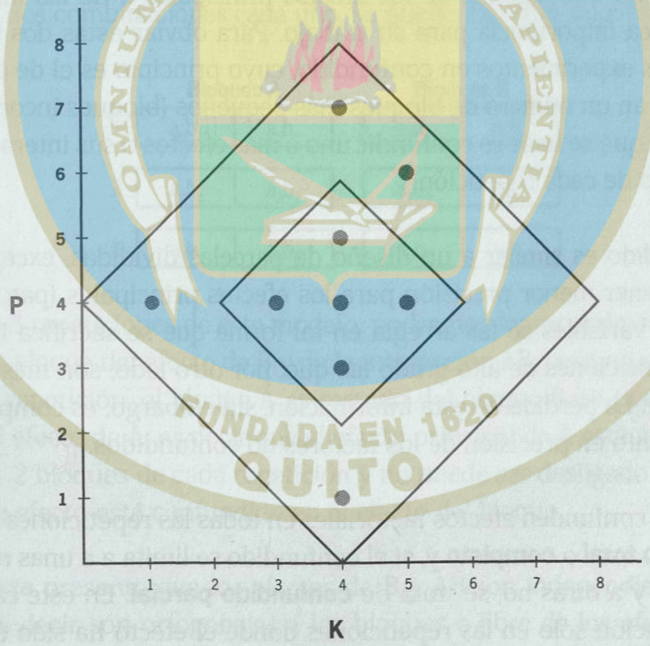


Fig. 17.2 Arreglo de "diamante doble".

El arreglo indicado proporciona un número adecuado de niveles de P y K para poder hallar el punto en que las líneas de respuesta se nivelan. También ofrece un método que permite determinar si los ejes se encuentran significativamente debajo de la línea de respuesta nivelada. Por medio de este análisis, se puede comparar los niveles 4–4 con 5–5 y 6–6. Si 5–5 y 6–6 son altamente significativos, en cuanto a rendimiento, se debe repetir el ensayo, usando niveles más altos. Dependiendo del tipo de suelo en el que se trabaje, puede usarse un nivel básico de N, por ejemplo, cuando los suelos son ricos en materia orgánica.

17.5 Experimentos en confundido

En experimentos factoriales de gran número de combinaciones (tratamientos) es difícil, por la superficie de terreno necesaria, comparar todos los efectos principales y sus respectivas interacciones. Otras veces, el investigador sabe de antemano que algunos de los efectos principales o de las interacciones, son de poca importancia para su trabajo. Para obviar estas dos situaciones, se usan los experimentos en confundido, cuyo principio es el de desdoblar la repetición en un número de bloques más pequeños (bloques incompletos). En realidad lo que se hace es confundir uno o más efectos o sus interacciones con los bloques de cada repetición.

El confundido es similar a un diseño de parcelas divididas, excepto que, en lugar de tener menor precisión para los efectos principales (parcelas principales), las variables se las arregla en tal forma que se sacrifica información sobre interacciones de alto grado las que, por otro lado, son más difíciles de interpretar. La pérdida de esta información, sin embargo, es compensada con el incremento en precisión de los factores no confundidos.

Cuando se confunden efectos factoriales en todas las repeticiones se habla de **confundido total** o **completo** y, si el confundido se limita a a unas repeticiones solamente y a otras no, se trata de **confundido parcial**. En este caso se pierde información sólo en las repeticiones donde el efecto ha sido confundido. Cuando se confunden dos interacciones AB y ABC,

$\frac{AB}{ABC}$, C también resulta confundido.

Si existe un gran número de tratamientos en un factorial, el error típico por unidad aumenta, en comparación con un experimento de un solo factor. Una de los objetivos del confundido es precisamente reducir el error típico por parcela.

Si el propósito del experimento es determinar un interacción, se puede confundir un efecto principal; pero esto no es usual. Generalmente se trata de ganar información sobre los efectos principales.

Para tener una idea más clara de lo indicado analicemos un ejemplo. Supóngase el caso de un experimento factorial 2x2, en cual deseamos confundir el efecto principal de A.

Si las 4 combinaciones de los factores en cada repetición, son dispuestos en dos bloques de dos combinaciones cada uno, se puede lograr el siguiente arreglo:

	Bloques A		Bloques B	
I	a_0b_1	a_0b_0	a_1b_1	a_1b_0
II	a_0b_0	a_0b_1	a_1b_0	a_1b_1
III	a_0b_1	a_0b_0	a_1b_1	a_1b_0

De tenerse 3 repeticiones de este modelo, podremos lograr estimaciones dentro de cada bloque del efecto de B y de la interacción AB, pero no para A; pero si en cada repetición, el bloque A es restado del bloque B, se tiene una estimación del efecto de A, es decir que el efecto principal de A está íntimamente unido a los 2 bloques de cada repetición y no puede ser desligado del bloque, es decir, su efecto está confundido en el efecto del bloque.

Debe tenerse presente que los efectos de B y AB son independientes de los bloques, es decir son ortogonales a los bloques o libre de los efectos de los bloques. De este modo logramos restar del error la variabilidad debido a bloques de 2 parcelas y no de 4 como hubiera sido el caso si hubiéramos usado en experimento factorial en bloques al azar.

El esquema del **ADEVA** para estas comparaciones dentro de bloques sería el siguiente:

ADEVA

Fuente	GL.
Bloques	5
B, AB	2
Error	4
Total	11

De las fuentes ofrecidas, quizas sólo convenga explicar que los *g. de l.* para el error considera los bloques A de las 3 repeticiones como en un experimento en bloques al azar de 2 tratamientos y 3 repeticiones; el error tendría (2–1) (3–1) = 2 grados de libertad. En la misma forma los bloques B de las 3 repeticiones darán 2 grados de libertad para el error. Estos dos grupos de 2 grados de libertad cada uno constituyen el error para el experimento completo. Es decir, que el error en experimentos factoriales confundidos está formado de interacciones entre tratamientos y bloques incompletos.

Aunque el efecto de A no puede ser medido tan precisamente como B y AB, puede tenerse una ligera idea de su efecto, al considerar el experimento en otra forma. Como uno en bloques al azar con 3 repeticiones y 2 tratamientos (Bloques). El esquema del ADEVA sería así:

	GL.
Repeticiones	2
A	1
Error para A	2
Total	5

Si el confundido ha sido eficiente, el error para A (con 2 *g. de l.*) debe ser más grande que el error del experimento (con 4 *g. de l.*).

Supongamos un ensayo con tres factores. A, B y C en cuatro repeticiones. El esquema del ADEVA ordinario y el esquema con ABC confundido, serían como sigue:

B.C.A.		ABC confundido	
F. de V.	g. de l.	F. de V.	g. de l.
Repet.	3	(ABC) Bloques	7
A	1	A	1
B	1	B	1
C	1	C	1
AB	1	AB	1
AC	1	AC	1
BC	1	BC	1
ABC	1	Error	18
Error	21	Total	31
Total	31		

Si bien BCA tiene 21 *g. de l.* para el error, contra 18 en confundido, el efecto de la variabilidad del suelo se reduce a ocho bloques pequeños, en confundido, frente a cuatro repeticiones grandes en BCA, lo cual resulta en mayor error experimental.

Vamos a ilustrar todo lo expuesto, con un ejemplo, en el cual se confunde la interacción de segundo grado NPK, al probar tres factores, N, P y K, cada uno con dos niveles. Para facilitar la explicación, se usará la siguiente notación:

$$
\begin{aligned}
n_0\ p_0\ k_0 &= (l)\\
n_1\ p_0\ k_0 &= A\\
n_0\ p_1\ k_0 &= B\\
n_0\ p_0\ k_1 &= C\\
n_1\ p_0\ k_1 &= AC\\
n_0\ p_1\ k_1 &= BC\\
n_1\ p_1\ k_0 &= AB\\
n_1\ p_1\ k_1 &= ABC
\end{aligned}
$$

en la que los factores N, P y K están representados por las letras A, B y C. El ejemplo tiene cuatro repeticiones y los rendimientos son como se indica en el Cuadro No 17.7.

CUADRO No 17.7 Arreglo de bloques y rendimientos de un factorial 2^3.

Parcela	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Tratam.	B	ABC	C	A	(1)	BC	AC	AB
Rendim.	42.	48.	46.	38.	25.	36.	40.	38.
Total Bloque	B = 174				A = 139			
I Rep. = 313								
Parcela	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Tratam.	(1)	AB	AC	BC	A	C	B	ABC
Rendim.	22.	36.	42.	34.	40.	52.	38.	35.
Total Bloque	A = 134				B = 165			
II Rep. = 299								
Parcela	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Tratam.	AC	(1)	BC	AB	A	B	C	ABC
Rendim.	38.	26.	32.	38.	36.	42.	50.	38.
Total Bloque	A = 134				B = 166			
III Rep. = 300								
Parcela	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
Tratam.	ABC	B	A	C	BC	AC	(1)	AB
Rendim.	42.	48.	28.	42.	35.	40.	28.	40.
Total Bloque	B = 160				A = 143			
IV Rep. = 303								

La disposición de los bloques provino del siguiente arreglo:

	Bloques A (–)				**Bloques B (+)**			
Rep. I	(1)	BC	AC	AB	B	ABC	C	A
II	(1)	AB	AC	BC	A	C	B	ABC
III	AC	(1)	BC	AB	A	B	C	ABC
IV	BC	AC	(1)	AB	ABC	B	A	C

CUADRO No 17.8 Rendimiento de cada tratamiento, por repetición.

Tratamiento	Repetición				Σ
	I	II	III	IV	
(1)	25	22	26	28	101
A	38	40	36	28	142
B	42	38	42	48	170
C	46	52	50	42	190
AB	38	36	38	40	152
AC	40	42	38	40	160
BC	36	34	32	35	137
ABC	48	35	38	42	163
Σ	313	299	300	303	1.215

Las sumas de cuadrados de bloques y total, se obtiene así:

$$F.C. = 1.215^2/rxt = 1.215^2/4x8 = 46.132,0$$

$$SC\ (Bloques) = 174^2 + 139^2 + 134^2 + \ldots + 160^2 + 143^2/4 - FC$$
$$= 46.585 - FC = 453,0$$

$$SC\ (Total) = 42^2 + 46^2 + 38^2 + \ldots + 28^2 + 40^2 - FC$$
$$= 47.683 - FC = 1.551,0$$

Los signos del Cuadro N° 17.9, permiten agrupar los tratamientos en los bloques A (–) y en los bloques B (+). Se pone signo (+), cuando el factor principal A, B o C, en la primera columna de la izquierda, se cruza con el mismo factor (a, b o c), de la fila al tope del cuadro. El signo de la interacción viene de multiplicar los signos de los efectos principales en el cuerpo del cuadro.

Dentro de cada repetición y de cada bloque, se sortea los tratamientos, como se ve al pie dej Cuadro N° 17.7. Los bloques A (–) y B (+) básicos, que serán sorteados para cada repetición, son:

B (+)			
A	B	C	ABC

A (–)			
(1)	AB	AC	BC

ubicación que puede ser verificada para cada repetición.

El total de los bloques B, menos el total de los bloques A, da el valor de la interacción ABC, la misma que se dice está confundida con bloques. Los demás tratamientos son ortogonales con bloques, es decir, están libres del efecto de bloques.

Para encontrar las sumas de cuadrados de los efectos principales y de las interacciones, se usa la técnica de los coeficientes de polinomios ortogonales, como se indica en el Cuadro No 17.9.

CUADRO N° 17.9 Coeficientes y sumas de cuadrados para los datos del Cuadro N° 17.7.

										Kr	SC
Trat.	**(1)**	**a**	**b**	**c**	**ab**	**ac**	**bc**	**abc**	**M**	$\Sigma c^2 r$	M^2/Kr
Total	**101**	**142**	**170**	**190**	**152**	**160**	**137**	**163**			
A	−1	+1	−1	−1	+1	+1	−1	+1	19	8(4)	11,28
B	−1	−1	+1	−1	+1	−1	+1	+1	29	8(4)	26,28
C	−1	−1	−1	+1	−1	+1	+1	+1	85	8(4)	225,78
AB	+1	−1	−1	+1	+1	−1	−1	+1	−3	8(4)	0,28
AC	+1	−1	+1	−1	−1	+1	−1	+1	−27	8(4)	22,78
BC	+1	+1	−1	−1	−1	−1	+1	+1	129	8(4)	520,03
ABC	−1	+1	+1	+1	−1	−1	−1	+1	-	8(4)	–

CUADRO No 17.10 ADEVA de los efectos principales e interacciones.

F. de V.	*g. de l.*	SC	CM	Fc	Ft	
					5%	1%
Bloques	7	453.00	64.71	3.99**	2,58	3.84
A	1	11.28	11.28	<1		
B	1	26.28	26.28	1.62	4,41	8,29
C	1	225.78	225.72	13.94**		
AB	1	0.28	0.28	<1		
AC	1	22.78	22.78	1.41		
BC	1	520.03	520.03	32.10**		
Error	18	291.57	16.20			
Total	31	1.551.00				

La interacción ABC está confundida con bloques.

Para comprobar si el confundido ha sido eficiente, efectuamos el ADEVA de bloques, a fin de comparar el tamaño del error experimental de este análisis (Cuadro No 17.11) con el del Cuadro No 17.10.

Para obtener la suma de cuadrados para repeticiones, se procede así,

$$SC\ (Rep.) = 313^2 + 299^2 + 300^2 + 303^2/8 - FC$$
$$= 369.179/8 = 46.147 - FC = 15.0$$

$$SC\ (ABC) = \underset{(Bloques\ A)}{(\ 139 + 134 + 134 + 143\ )^2} + \underset{(Bloques\ B)}{(\ 174 + 165 + 166 + 160\ )^2}/16 - F.C.$$
$$= (550)^2 + (665)^2/16 - FC$$
$$= 46.545 - FC = 413.0$$

$$SC\ (Error) = SC\ (Bloques) - SC\ (Rep.) - SC\ (ABC)$$
$$= 453.0 - 15.0 - 413.0 = 25$$

CUADRO No 17.11 ADEVA de bloques.

F. de V.	g. de l.	SC	CM	Fc	Ft 5%	Ft 1%
Rep.	3	15	5,00	<1		
ABC	1	413	413,00	49,58**	10,13	34,12
Error	3	25	8,33			
Total	7	453				

Si el confundido ha sido eficiente, el cuadrado medio del error para el ADEVA de ABC, tiene que ser más grande que el cuadrado medio del error para el ADEVA de los efectos principales e interacciones de primer orden.

En el ejemplo propuesto, se han usado datos supuestos, para facilitar el cálculo, y se ha encontrado que el tamaño del error para el Cuadro No 17.10 es mayor que en el caso del Cuadro N° 17.11. Lo que hemos pretendido es, simplemente, ilustrar el procedimiento de cálculo.


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM
SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 18

PARCELAS DIVIDIDAS

Este tipo de diseños se usa frecuentemente para experimentos factoriales. El principio general es el de que la **parcela principal**, a la que se ha asignado un nivel de un factor, se la divide en subunidades o subparcelas, que reciben varios niveles de un factor adicional. En esta forma, cada Unidad principal se convierte en un bloque con respecto a los tratamientos asignados a las subunidades.

Considérese el caso de dos factores A (métodos de riego) y B (variedades), cada uno con tres niveles; después de randomizar los tratamientos, una repetición sería como sigue:

		a1	a3	a2
Rep. I	B	$a_1 b_3$	$a_3 b_1$	$a_2 b_2$
		$a_1 b_1$	$a_3 b_2$	$a_2 b_3$
		$a_1 b_2$	$a_3 b_3$	$a_2 b_1$
			A	

Podemos decir de a_1 que es un bloque incompleto en confundido. Randomizamos primero los niveles de A en todas las parcelas principales y luego los de B dentro de la parcela principal. Cada parcela principal (factor A) puede ser

considerada como un bloque respecto a los niveles de B, pero sólo como un bloque incompleto con relación al grupo completo de tratamientos A. Por esta razón se llama a las parcelas divididas "diseño de bloques incompletos", o "experimentos factoriales con los efectos principales confundidos". Este diseño es aplicable en las siguientes circunstancias:

1. Cuando se dispone de menor cantidad de material experimental correspondiente a un factor que para otros. En un experimento de campo un factor podría ser **métodos de riego, preparación del suelo** o **aplicación de fertilizantes;** los tres usualmente requieren parcelas grandes para su aplicación. El otro factor podría ser **variedades,** las que pueden ser comparadas en **parcelas más pequeñas.**

2. Este diseño se usa cuando se va a incluir un **factor adicional** en el experimento. Si el objeto principal de un ensayo es probar **algunos insecticidas,** se podría aumentar el **alcance** del mismo probando algunas variedades. En este caso, se asignarían las **variedades** a las **parcelas principales** y los **insecticidas** a las **sub-unidades.**

3. Por experiencias previas, sabemos que se puede esperar diferencias más grandes entre los niveles de un factor que entre los de otro. Si queremos detectar diferencias significativas más pequeñas entre los niveles de un factor, éste será asignado a las subparcelas donde, por haber mayor repetición –así como mayor número de grados de libertad– ese factor será medido con más precisión que aquel asignado a las parcelas principales.

En resumen, puesto que la variación entre sub-unidades será menor que entre parcelas principales, se asigna a las primeras los factores que requieren menor cantidad de material experimental, o los que son de mayor importancia, o de los que se espera diferencias de menor tamaño, o para los que se necesita mayor precisión.

Generalmente, el error (a) es mayor que el error (b). El tamaño de parcelas divididas se lo usa frecuentemente para experimentos factoriales y puede incorporar uno o más diseños DCA, BCA y cuadrado latino.

Ejemplo: El Cuadro 18.1 presenta el rendimiento de dos variedades de fréjol, sembradas a tres distancias, en cuatro repeticiones.

CUADRO Nº 18.1 Rendimiento de dos variedades de fréjol, sembradas a tres distancias.

Variedad	Distancia	Repeticiones				Σ
		I	II	III	IV	
V1	d1	28.8	35.4	40.2	30.1	134.5
	d2	32.4	38.6	34.6	28.9	134.5
	d3	30.2	40.1	37.2	26.5	134.0
		91.4	114.1	112.0	85.5	403.0
V2	d1	42.4	38.9	39.3	40.1	160.7
	d2	44.6	37.2	40.8	42.4	165.0
	d3	38.9	41.7	38.6	38.8	158.0
		125.9	117.8	118.7	121.3	483.7
		217.3	231.9	230.7	206.8	886.7

$$FC = \frac{(886.7)^2}{2 \times 3 \times 4} = 32.759,87$$

$$S.C.\ (Total\ o\ subparcelas) = 28,8^2 + 32,42 + 30,22 + \ldots + 42,42 + 38,82 - F.C.$$
$$= 33.329,25 - FC = 569,38$$

$$SC\ (Reps.) = 217.3^2 + 231,9^2 + 230,7^2 + 206,8^2/2x3 - FC$$
$$= 32.830,94 - FC = 71,07$$

$$SC(Dist.) = (134,5 + 160,7)^2 + (134,5 + 165,0)^2 + (134,0 + 158,0)^2/2x4 - FC$$
$$= 295,2^2 + 299,5^2 + 292.Q^2/8 - FC$$
$$= 32.763,41 - FC = 3,54$$

$$SC\ (Parcelas) = 91,4^2 + 114,1^2 + 112,0^2 + \ldots + 121,3^2/3 - FC$$
$$= 33.252,68 - FC = 492,81$$

$$SC\ (Var.) = (134,5 + 134,5 + 134,0)^2 + (160,7 + 165,0 + 158,0)^2/4x3 - FC$$
$$= (403,0)^2 + (483,7)^2/12 - FC$$
$$= 33.031,22 - FC = 271,35$$

$$SC\ (Var.\ x\ Dist.) = 134{,}5^2 + 134.5^2 + \ldots + 158{,}0^2/4 - FC - SC(Var.) - SC(Dist.)$$
$$= 33.037{,}5 - FC - 271{,}35 - 3{,}54$$
$$= 2{,}74$$

$$SC\ Error\ (a) = SC\ (Parcelas) - SC\ (Rep.) - SC\ (Varied.)$$
$$= 492{,}81 - 71{,}07 - 271{,}35$$
$$= 150{,}39$$

$$SC\ Error\ (b) = SC\ (Subparce]as) - SC\ (Parcelas) - SC\ (Dist.) - SC\ (Var.\ x\ Dist.)$$
$$= 569{,}38 - 492{,}81 - 3{,}54 - 2{,}74$$
$$= 70{,}29$$

CUADRO No 18.2 ADEVA para los datos del Cuadro No 17.1.

F. de V.	g. de l.	SC	CM	Fc	Ft 5%	Ft 1%
Rep.	3	71,07				
Varied.	1	271,35	271,35	5,41 N.S.	10,13	34,12
Error (a)	3	150,39	50,13			
Dist.	2	3,54	1,77	<1	3,89	6,93
VXD	2	2,74	1,37	<1		
Error (b)	12	70,29	5,86			
Total	23	569,38				

Una vez obtenidas las sumas de cuadrados, en la forma indicada, se obtienen los cuadrados medios, dividiendo las sumas de cuadrados para los grados de libertad. Para obtener los valores de *F*, se usa el error (a), para probar variedades y el error (b), para probar distancias y la interacción (V x D).

Se contrasta *Fc* con *Ft,* de acuerdo al número de *g. de l.* así: para variedades, se ve el valor de *F*, para un *g. de l.* en el numerador (fila superior de la Tabla) y tres *g. de l.* (columna izquierda), en el denominador.

En nuestro ejemplo, vemos que 5,41 < 10,13, por lo que declaramos los resultados no significativos, es decir, no existe diferencia entre las variedades.

Para probar la significación de distancias y la interacción (V x D), vemos el valor de *F* para dos y 12 *g. de l.*

Como en el caso de variedades, no existe significación para distancias, ni para la interacción.

Los coeficientes de variación, están dados por,

$$C.V._{(a)} = \sqrt{50.13}/36.95(100) = 19\%$$

$$C.V._{(b)} = \sqrt{5.86}/36.95(100) = 6,55\%$$

18.1 Valores perdidos en parcelas divididas

Cuando, por alguna circunstancia, se ha perdido una observación *X*, ésta puede estimarse con la siguiente fórmula:

$$X = \frac{p(R) + q(M) - P}{(p-1)(q-1)}, \text{ en la que,}$$

X = *Rendimiento estimado de la parcela perdida.*
p = *N° de sub-parcelas por parcela (niveles de B).*
q = *N° de repeticiones.*
P = *Total del rendimiento del tratamiento en parcelas que contiene la sub-parcela perdida.*
M = *Total de rendimiento de la parcela específica que contiene la sub-parcela perdida.*
R = *Total del rendimiento del tratamiento en sub-parcelas que contiene la sub-parcela perdida.*

Si hay varios valores perdidos en diferentes parcelas principales, se estimará las sub-parcelas, dentro de cada parcela por medio de la fórmula indicada. Si hay más de una sub-parcela perdida dentro de la misma parcela, se usará en forma repetida la misma ecuación. Antes de proceder al análisis de variancia, se substraerá un grado de libertad del error *(b)* por cada sub-parcela perdida. La estimación del *E(b)* es insesgada; sin embargo los cuadrados medios para

tratamientos y *E(a)* pueden ser sobreestimados. Esto no es problema cuando hay pocos valores perdidos. Steel y Torrie (1960) y Cochran y Cox (1951) dan fórmulas para estimar algunas parcelas perdidas y errores típicos.

18.2 Modelo lineal aditivo

En parcelas divididas, una observación está representada así,

$$X_{ijk} = \mu + \rho_i + \alpha_j + \delta_{ij} + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}$$

en la que μ es la media por observación; ρ_i, el efecto de bloques; α_j el efecto del factor A; δ_{ij}, el componente al azar de la parcela grande; β_k, el efecto del factor B (subparcelas); $(\alpha\beta)_{jk}$, de la interacción AB y ε_{ijk}, el error experimental.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Capítulo 19

CÁLCULO DE APTITUD COMBINATORIA GENERAL Y ESPECÍFICA, EN CRUZAMIENTOS DIALÉLICOS

Dentro del trabajo de creación de nuevas variedades, el fitomejorador introduce miles de variedades, líneas o clones para, como fase final del proceso de mejoramiento, comparar el comportamiento de unas pocas de ellas, en combinaciones híbridas.

Es importante comenzar con un crecido número de selecciones, con el objeto de tener un amplio rango de variabilidad genética. De hecho, esta variabilidad es la que permite al mejorador realizar progresos, consiguiendo frecuentemente variedades más resistentes a plagas y enfermedades, de mayor rendimiento y calidad.

El procedimiento general es establecer un vivero de selección fenotípica, que incluye varios miles de plantas individuales. Después de tomar notas sobre aspectos agronómicos, que son de interés para el investigador, se selecciona 200 a 400 individuos, con los que se forma otro vivero con repeticiones. En el caso de cultivos como raigrás, pasto azul, festuca, constituyen repeticiones los clones extraídos de cada individuo: se corta la corona en cuatro a seis pedazos, cada uno con raíces, para facilitar el prendimiento. En el caso de alfalfa, puede también usarse segmentos de tallo, a los que se enraiza previamente en el invernadero, antes de transplantarlos al campo.

Se somete los clones a condiciones adversas, como ataque de enfermedades, pastoreo intensivo, etc., con el objeto de aislar el material más promisorio del vivero.

De la evaluación realizada en las cuatro-seis repeticiones, se seleccionan 25-50 individuos superiores, a fin de comparar la aptitud combinatoria de los mismos. Esta prueba se efectúa estudiando la progenie de los cruzamientos, para lo que se puede aplicar uno de los siguientes métodos: a) polinización libre, b) policruzamiento, c) topcros, d) autofecundaciones y e) cruzamientos simples, dobles, etc.

Estos métodos pueden ser usados para cualquier cultivo, al igual que los sistemas empleados para la determinación de aptitud combinatoria, como se describe más adelante.

Cuando se quiere determinar la aptitud combinatoria específica, es necesario realizar cruzamientos simples, para lo que se selecciona alrededor de 10 individuos o clones superiores y se los cruza formando todas las combinaciones posibles, llamadas cruzamientos dialélicos; los cruzamientos dialélicos se usan también para mejoramiento animal y en todos los casos que incluyan genética cuantitativa.

El número de plantas para este tipo de cruzamientos no debe ser muy alto puesto que, en ese caso, el trabajo manual puede ser largo y tedioso.

En los cruzamientos dialélicos, un número p de progenitores, da lugar a un número de p^2 combinaciones.

Las técnicas en los cruzamientos dialélicos varían dependiendo de si se incluye los progenitores, las F_1 y recíprocos. Griffing (1956), clasifica las posibles combinaciones en cuatro grupos, de acuerdo a los objetivos que se persigue:

1) Se incluye los progenitores, las F_1 y recíprocos (todas las combinaciones p^2).

2) Progenitores y F_1, pero no los recíprocos (las combinaciones son $1/2P(p + 1)$).

3) Se incluye las F_1 y recíprocos, pero no los progenitores (combinaciones $p\,(p-1)$).

4) Se incluye las F_1 pero no recíprocos ni progenitores (combinaciones 1/2 $p(p-1)$).

Para cada método se aplica una forma diferente de análisis, de acuerdo a los Modelos I y II, según se considere al material experimental como la población total, a la que se va a aplicar las inferencias (Modelo Fijo); o si la suposición es de que estamos trabajando con muestras al azar, extraídas de una población más grande (Modelo al Azar).

19.1 Modelo matemático

Para la ilustración que se va a presentar posteriormente, el autor usó el Método I, Modelo II, cuyo modelo matemático consta en el Cuadro No 19.1.

CUADRO No 19.1 ADEVA para el Método I, Modelos I y II.

F. de V.	g. de l.	SC[3]	CM	C. M. Esperados	
				Modelo I	Modelo II
A.c.g.[1]	$p-1$	Sg	Mg	$\sigma^2 + 2p\left[\frac{1}{p-1}\right]\sum g_i^2$	$\sigma^2 + \frac{2(p-1)}{p}\sigma^2 s + 2p\sigma^2 g$
A.c.e. [2]	$p(p-1)/2$	Se	Me	$\sigma^2 + \frac{2}{p(p-1)}\sum_i\sum_j s^2_{ij}$	$\sigma^2 + \frac{2(p^2-p+1)}{p^2}\sigma^2 s$
Recip.	$p(p-1)/2$	Sr	Mr	$\sigma^2 + 2\left[\frac{2}{p(p-1)}\right]^{\sum\sum r_{ij}}_{i<j}$	$\sigma^2 + 2\sigma^2 r$
Error	m	Se'	Me'	σ^2	σ^2

[1] ***Aptitud combinatoria general.***
[2] ***Aptitud combinatoria específica.***
[3] ***Donde la suma de cuadrados, se obtiene así:***

$$Sg = \frac{1}{2p}\sum(X_{i\cdot} + X_{\cdot i})^2 - \frac{2}{p^2}X^2\ldots$$

$$Se = \frac{1}{2}\sum_i\sum_j X_{ij}(X_{ij} + X_{ji}) - \frac{1}{2p}\sum_i(X_{\cdot i} + X_{i\cdot})^2 + \frac{1}{p^2}X^2$$

$$Sr = \frac{1}{2}\sum_{i<j}\sum(X_{ij} + X_{ji})^2$$

Para determinar significación, se compara los CM, así:

1) Para probar *a. c. g.*,

$$F[(p-1), m]Mg/Me', \text{ donde,}$$

(p – 1) son los g. de l. para a. c. g. y m los g. de l. para el error.

2) Para probar a. c. e.,

$$F[p\ (p-1)/2, m] = Me/Me'$$

3) Para efectos recíprocos,

$$F[p\ (p-1)/2, m] = Mr/Me'$$

19.2 El experimento

De una prueba anterior se seleccionaron ocho clones de alfalfa resistentes a la marchitez bacteriana. Las calificaciones de resistencia, en base a una escala relativa, fueron realizadas arrancando las plantas del suelo y seccionando la raíz longitudinalmente; la escala usada fue la siguiente:

0 : No infección, raíces blancas, limpias.

1 : Trazas de infección, pequeñas manchas amarillas o rayas en el tejido vascular.

2 : Ligera infección, manchas amarillentas-cafés; las manchas no cubren toda la superficie; anillos incompletos en corte seccional.

3 : Infección moderadamente severa; manchas extensas, anillos continuos en corte seccional.

4 : Infección severa, anillos continuos con manchas amarillo-cafés.

5 : Infección muy severa, tejido vascular muerto o en descomposición.

Se calificó a las plantas individualmente en cada parcela, obteniendo el porcentaje de plantas resistentes por parcela, en la siguiente forma:

$$\% \text{ de plantas resistentes: } \frac{0+1}{total}$$

Además de los clones seleccionados, se usaron seis testigos.

Los ocho clones progenitores fueron llevados al invernadero, en donde se extrajeron segmentos de tallo, los que fueron enraizados previamente en vermiculita y transplantados después individualmente a macetas de 12,5 cm. de diámetro. De estas plantas se cosechó la semilla para el trabajo que se describe.

19.3 Método de cruzamiento

Se usaron 16 horas de luz artificial, dentro del invernadero, para inducir floración. Cuando fue necesario, se cortaron las plantas y se las dejó florecer nuevamente a fin de obtener la cantidad requerida de semilla, por autofecundaciones o cruzamientos manuales.

Los ocho clones fueron cruzados formando todas las combinaciones posibles, sin emasculación. Dos plantas de un clon y dos de otro, formaron una unidad de cruzamiento.

Se utilizó un pedazo de papel secante doblado a fin de colectar el polen para efectuar los cruzamientos. Para cada unidad de cruzamiento se usó una flor de un clon y una flor de otro, en forma alternada.

Una o dos plantas formaron una unidad de autofecundación. Se empleó un palillo de dientes para permitir la "explosión" de la columna estaminal en las flores de cada uno de los clones progenitores y conseguir su autofecundación. Debido a una alta autoesterilidad de la mayoría de los clones, hubo que repetir esta operación en un gran número de ellos, a fin de reunir la cantidad necesaria de semilla.

A la madurez, la semilla fue cosechada en su totalidad de cada una de las unidades de cruzamiento, manteniéndose la semilla proveniente de los cruzamientos recíprocos, separada. Posteriormente, la semilla fue trillada a mano y escarificada. Se empleó una caja de madera con fondo de caucho para la trilla.

19.4 Semillero y proceso de infección

La semilla proveniente de todos los cruzamientos, y autofecundaciones, fue sembrada en el invernadero en arena fina. Las parcelas individuales estaban formadas por cuatro surcos de un metro de largo, con cinco centímetros de separación; se sembró una semilla cada 2,5 cm. Los cruzamientos fueron randomizados en el invernadero, al igual que en el transplante posterior en el campo. Las plántulas fueron fertilizadas periódicamente con una solución nutritiva que contenía N, P y K, regadas y mantenidas limpias de maleza todo el tiempo.

Ocho días después de la emergencia, las plántulas de cada parcela fueron extraídas de la arena, lavadas en agua y se formaron atados para ser infectadas. Las plántulas fueron transplantadas enseguida al campo, formándose parcelas con un promedio de 100 plantas cada una. Se usó un diseño de bloques al azar con tres repeticiones.

19.5 Toma de notas

Se efectuó un contaje de las plantas prendidas, una vez que estuvieron bien establecidas. Al final del quinto mes, todas las plantas fueron arrancadas y calificadas de acuerdo a la escala establecida. El Cuadro Nº 19.2 presenta el número de plantas resistentes dentro de cada repetición.

CUADRO No 19.2 Porcentaje de plantas resistentes (0 + 1/Total).

Cruzamiento N°	REPETICIÓN			Total	$\overline{X}$
	I	II	III		
1	22	17	38	77	25.66
2	62	44	80	186	62.00
3	30	59	52	141	47.00
4	32	48	53	133	44.33
5	7	8	15	30	10.00
6	30	35	70	135	45.00
7	11	42	48	101	33.66
8	1	10	4	15	3.00
9	59	88	85	232	77.33
10	44	58	56	158	52.66
11	44	57	64	165	55.00
12	3	6	17	26	8.66
13	27	60	47	134	44.66
14	22	12	44	78	26.00
15	0	0	1	1	0.33
16	73	93	93	259	86.33
17	54	88	92	234	78.00
18	80	98	98	276	92.00
19	42	84	89	215	71.66
20	82	91	88	261	87.00
21	48	54	87	189	63.00
22	40	70	79	189	63.00
23	72	78	92	242	80.66
24	72	97	88	257	85.66
25	77	36	78	191	63.66
26	64	74	77	215	71.66
27	58	62	82	202	67.33
28	44	63	72	179	80.00
30	60	42	90	192	64.00
31	40	50	44	134	44.66
32	49	72	80	201	67.00
33	34	42	73	149	49.66
34	21	32	26	79	26.33
35	54	45	68	167	55.66
36	24	48	38	110	36.66
37	53	70	74	197	65.66
38	20	62	57	139	46.33
39	15	14	38	67	22.33

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1826
QUITO

Cruzamiento Nº	REPETICIÓN I	II	III	Total	$\overline{X}$
40	66	4	84	154	51.33
41	45	56	67	168	56.00
42	86	87	84	257	85.66
43	59	41	74	174	58.00
44	14	12	21	47	15.66
45	24	38	48	110	36.66
46	50	81	60	191	63.66
47	43	86	78	207	69.00
48	20	23	21	64	21.33
49	4	15	8	27	9.00
50	38	46	29	113	37.66
51	31	28	36	95	31.66
52	1	0	0	1	0.33
53	44	42	44	130	43.33
54	10	52	41	103	34.33
55	0	0	0	0	–
56	58	80	92	230	76.66
57	44	82	66	192	64.00
58	30	36	22	88	29.33
59	56	76	70	202	67.33
60	4	29	48	81	27.00
61	42	34	66	142	47.33
62	4	31	18	53	17.66
63	11	7	4	22	7.33
64	0	21	0	21	7.00
	2.440	3.086	3.542	9.068	

ADEVA (Basado en parcelas individuales)

F. de V.	*g. de l.*	S.C.	C.M.	F
Total	191	150.727,92		
Repeticiones	2	9.581,54	4.790,00	
Cruzamientos	63	122.064,58	1.937,53	12,79**
Error	126	19.081,80	151,44	

SC Total=579.002 – *F.C.*= 150.727,92 ; *F.C.* = $(9.068)^2/192$

SC (Reps.)=28.022.760/64 – *F.C.*= 9.581,54 ; = 428.274,08

SC (Cruzam.) = 1.651.016/ 3 – *F.C.*= 122.064,58

19.6 Análisis estadístico

El análisis estadístico se hizo de acuerdo al diseño de bloques al azar, con tres repeticiones, en base a las 64 parcelas individuales, excluyendo a los testigos. Al pie del Cuadro No 19.2 se presenta el ADEVA de los datos, el que muestra un valor de F altamente significativo para cruzamientos. La suma de cuadrados para cruzamientos, fue desdoblada en los componentes de a. c. g., a. c. e. y recíprocos, de acuerdo al Método 1, Modelo II, de Griffing (1956), para todas las combinaciones inclusive autofecundaciones.

Para estos cálculos, se usó la media de las tres repeticiones.

El Cuadro No 19.3 presenta los totales y promedios de plantas resistentes *(0 + 1/Total)*, para los 28 cruzamientos simples, y el Cuadro No 19.4 el resumen de medias para los 64 cruzamientos, incluyendo autofecundaciones. Enseguida del Cuadro No 19.4 se demuestra el procedimiento detallado de cálculo, para obtener la suma de cuadrados para a.c.g., a.c.e. y recíprocos. Se ha incluido todas las operaciones con el objeto de que, a partir del Cuadro No 19.4, el lector interesado pueda saber exactamente de dónde provienen los diferentes valores.

CUADRO No 19.3 Porcentaje de plantas resistentes para los 28 cruzamientos simples (0 + 1)/Total).

Cruzamiento	Total	Promedio
1	186	62
2	141	47
3	133	44
4	30	10
5	135	45
6	101	34
7	15	5
8	234	78
9	276	92
10	215	72
11	261	87
12	189	63
13	189	63

Cruzamiento	Total	Promedio
14	192	64
15	134	45
16	201	67
17	149	50
18	79	26
19	168	56
20	257	86
21	174	58
22	47	16
23	113	38
24	95	32
25	1	0.33
26	192	64
27	88	29
28	53	18

CUADRO N° 19.4 Valores promedios de plantas resistentes, de acuerdo a los cruzamientos.

	Progenitor No									
	1	2	3	4	5	6	7	8	$\sum$	$X_{.j}$
1	9	9	64	37	37	38	32	0	226	$X_{1.}$
2	10	26	62	47	44	45	34	3	271	$X_{2.}$
3	72	77	86	78	92	87	63	63	618	$X_{3.}$
4	45	53	81	80	64	67	50	26	466	$X_{4.}$
5	56	55	86	56	51	86	58	16	464	$X_{5.}$
6	43	45	72	66	64	77	64	29	460	$X_{6.}$
7	34	26	67	46	69	67	47	18	374	$X_{7.}$
8	0	0	60	22	21	27	7	7	144	$X_{8.}$
$\sum$	269	291	578	432	442	494	355	162	3.023	
$X_{i.}$	$X_{.1}$	$X_{.2}$	$X_{.3}$	$X_{.4}$	$X_{.5}$	$X_{.6}$	$X_{.7}$	$X_{.8}$		

Cálculo

$$S.C.(a.c.g.) = \frac{1}{2p}\sum_i (X_{i.} + X_{.i}) - \frac{2}{p^2}(X..)^2, \textit{para las medias.}$$

$$= \frac{1}{2p(3)}\sum_i (X_{i.} + X_{.i})^2 - \frac{2}{3p}(X..)^2, \textit{para totales}$$

$$S.C.(a.c.e.) = \frac{1}{2}\sum_i\sum_j X_{ij}(X_{ij} + X_{ji}) - \frac{1}{2p}\sum_i (X_{i\cdot} + X_{\cdot i})^2 + \frac{1}{p^2}X^2{}_{\cdot\cdot}$$

$$SC\ (recípro\ cos) = \frac{1}{2}\sum_{i<j}\sum (X_{ij} + X_{ji})^2$$

19.6.1 Cálculo de la SC para a. c. g.

$$\frac{1}{2\times 8}\sum_i (x_{i\cdot} + x_{\cdot i})^2 - \frac{2}{8^2}(3.023)^2$$

$$\sum_i (x_{i\cdot} + x_{\cdot i})^2 = \begin{array}{rcl} (269+226)^2 & = & 495^2 \\ (291+271)^2 & = & 562^2 \\ (578+618)^2 & = & 1.196^2 \\ (432+466)^2 & = & 898^2 \\ (442+464)^2 & = & 906^2 \\ (494+460)^2 & = & 954^2 \\ (355+374)^2 & = & 792^2 \\ (162+144)^2 & = & 306^2 \\ \sum & = & 5.153.718 \end{array}$$

$$\frac{1}{2\times 8}(5.153.718) - \frac{1}{64}(3.023)^2$$

$$322.107{,}38 - 285.579{,}0 = 36.528{,}34$$

19.6.2 Cálculo de la S.C. para a.c.e.

$$\frac{1}{2}\sum_i\sum_j X_{ij}(X_{ij} + X_{ji}) - 322.107.38 + \frac{1}{64}(3.023)^2$$

$$\sum_i\sum_j X_{ij}(X_{ij} + X_{ji}) = \begin{array}{rlrcl} 1. & X_{11} & 9(9+9) & = & 9(18) \\ 2. & X_{12} & 9(9+10) & = & 9(19) \\ 3. & X_{13} & 64(64+72) & = & 64(136) \\ 4. & X_{14} & 37(37+45) & = & 37(82) \\ 5. & X_{15} & 37(37+56) & = & 37(93) \\ 6. & X_{16} & 38(38+43) & = & 38(81) \\ 7. & X_{17} & 32(32+34) & = & 32(66) \\ 8. & X_{18} & 0+0 & = & 0 \\ 9. & X_{22} & 26(26+26) & = & 26(52) \end{array}$$

10. X_{23} 62(62+77) = 62(139)
11. X_{24} 47(47+53) = 47(100)
12. X_{25} 44(44+55) = 44(99)
13. X_{26} 45(45+45) = 45(90)
14. X_{27} 34(34+26) = 34(60)
15. X_{28} 3(3+ 0) = 3(3)
16. X_{33} 86(86+86) = 86(172)
17. X_{34} 78(78+81) = 78(159)
18. X_{35} 92(92+86) = 92(178)
19. X_{36} 87(87+72) = 87(159)
20. X_{37} 63(63+67) = 63(130)
21. X_{38} 63(63+60) = 63(123)
22. X_{44} 80(80+80) = 80(160)
23. X_{45} 64(64+56) = 64(120)
24. X_{46} 67(67+66) = 67(133)
25. X_{47} 50(50+46) = 50(96)
26. X_{48} 26(26+22) = 26(48)
27. X_{55} 51(51+51) = 51(102)
28. X_{56} 86(86+64) = 86(150)
29. X_{57} 58(58+69) = 58(127)
30. X_{58} 16(16+21) = 16(37)
31. X_{66} 77(77+77) = 77(154)
32. X_{67} 64(64+67) = 64(131)
33. X_{68} 29(29+27) = 29(56)
34. X_{77} 47(47+47) = 47(94)
35. X_{78} 18(18+7) = 18(25)
36. X_{88} 7(7+7) = 7(14)
37. X_{21} 10(10+9) = 10(19)
38. X_{31} 72(72+64) = 72(136)
39. X_{32} 77(77+62) = 77(139)
40. X_{41} 45(45+37) = 45(82)
41. X_{42} 53(53+47) = 53(100)
42. X_{43} 81(81+78) = 81(159)
43. X_{51} 56(56+37) = 56(93)
44. X_{52} 55(55+44) = 55(99)

45.	X_{53}	86(86+92)	=	86(178)
46.	X_{54}	56(56+64)	=	56(120)
47.	X_{61}	43(43+38)	=	43(81)
48.	X_{62}	45(45+45)	=	45(90)
49.	X_{63}	72(72+87)	=	72(159)
50.	X_{64}	66(66+67)	=	66(133)
51.	X_{65}	64(64+86)	=	64(150)
52.	X_{71}	34(34+32)	=	34(66)
53.	X_{72}	26(26+34)	=	26(60)
54.	X_{73}	67(67+63)	=	67(130)
55.	X_{74}	46(46+50)	=	46(96)
56.	X_{75}	69(69+58)	=	69(127)
57.	X_{76}	67(67+64)	=	67(131)
58.	X_{81}	0+0	=	0
59.	X_{82}	0(0+3)	=	0(3)
60.	X_{83}	60(60+63)	=	60(123)
61.	X_{84}	22(22+26)	=	22(48)
62.	X_{85}	21(21+16)	=	21 (37)
63.	X_{86}	27(27+29)	=	27(56)
64.	X_{87}	7(7+18)	=	7(25)
		$\sum X_{ij}\,(X_{ij}+X_{ji})$	=	365.464,01

$$= \frac{1}{2}(365.464.0) - 322.107.38 + \frac{1}{64}(3.023)^2$$

$$= 182.732,0 - 322.107,38 + 142.789,52$$

$$= 3.414,14$$

19.6.3 Cálculo de la SC para recíprocos (28 cruzamientos simples)

$$\frac{1}{2}\sum_i\sum_j (X_{ij} - X_{ji})^2$$

$\sum_i\sum_j (X_{ij} - X_{ji})^2$

1.	$X_{12} - X_{21}$	$(10-9)^2$	=	$(1)^2$
2.	$X_{13} - X_{31}$	$(72-64)^2$	=	$(8)^2$
3.	$X_{14} - X_{41}$	$(45-37)^2$	=	$(8)^2$
4.	$X_{15} - X_{51}$	$(37-56)^2$	=	$(-19)^2$

5.	$X_{16} - X_{61}$	=	$(43 - 38)^2$	=	$(5)^2$
6.	$X_{17} - X_{71}$	=	$(34 - 32)^2$	=	$(2)^2$
7.	$X_{18} - X_{81}$	=	$(0 - 0)^2$	=	0
8.	$X_{23} - X_{32}$	=	$(77 - 62)^2$	=	$(15)^2$
9.	$X_{24} - X_{42}$	=	$(53 - 47)^2$	=	$(6)^2$
10.	$X_{25} - X_{52}$	=	$(55 - 44)^2$	=	$(11)^2$
11.	$X_{26} - X_{62}$	=	$(45 - 45)^2$	=	$(0)^2$
12.	$X_{27} - X_{72}$	=	$(26 - 34)^2$	=	$(-8)^2$
13.	$X_{28} - X_{82}$	=	$(0 - 3)^2$	=	$(-3)^2$
14.	$X_{34} - X_{43}$	=	$(81 - 78)^2$	=	$(3)^2$
15.	$X_{35} - X_{53}$	=	$(86 - 92)^2$	=	$(-6)^2$
16.	$X_{36} - X_{63}$	=	$(72 - 87)^2$	=	$(-15)^2$
17.	$X_{37} - X_{73}$	=	$(67 - 63)^2$	=	$(4)^2$
18.	$X_{38} - X_{83}$	=	$(60 - 63)^2$	=	$(-3)^2$
19.	$X_{45} - X_{54}$	=	$(56 - 64)^2$	=	$(-8)^2$
20.	$X_{46} - X_{64}$	=	$(66 - 67)^2$	=	$(-1)^2$
21.	$X_{47} - X_{74}$	=	$(46 - 50)^2$	=	$(-4)^2$
22.	$X_{48} - X_{84}$	=	$(22 - 26)^2$	=	$(-4)^2$
23.	$X_{56} - X_{65}$	=	$(64 - 86)^2$	=	$(-22)^2$
24.	$X_{57} - X_{75}$	=	$(69 - 58)^2$	=	$(11)^2$
25.	$X_{58} - X_{85}$	=	$(21 - 16)^2$	=	$(5)^2$
26.	$X_{67} - X_{76}$	=	$(67 - 64)^2$	=	$(3)^2$
27.	$X_{68} - X_{86}$	=	$(27 - 29)^2$	=	$(-2)^2$
28.	$X_{78} - X_{87}$	=	$(7 - 18)^2$	=	$(-11)^2$

$$\sum_i \sum_j (X_{ij} - X_{ji})^2 = 2.130.0$$

$$\frac{1}{2}\sum_i \sum_j (X_{ij} - X_{ji})^2 = \frac{1}{2}(2.130.0) = 1.065$$

19.6.4 Resumen del cálculo de las S. C. de los componentes

$$a)\ S.C.(a.c.g.) = \frac{1}{2p}\sum (X_{i.} + X_{.j})^2 - \frac{2}{p^2}(\wedge..)^2$$

$$= \frac{1}{2 \times 8}(5.153.718) - \frac{2}{64}(3.023)^2$$

$$= 322.107.38 - 285.579.0 = 36.528.34$$

$$b)\ S.C.(a.c.e.) = \frac{1}{2}\Sigma\Sigma x_{ij}(x_{ij} + x_{ji}) - SC(a.c.g.)\ no\ corregidos + \frac{1}{64}(3.023)^2$$

$$= \frac{1}{2}(365.464) - 322.107.34 + \frac{(3.023)^2}{64}$$

$$= 182.732.0 - 322.107.34 + 142.789.52$$

$$= 3.414.14$$

$$c)\ SC(recíprocos) = \frac{1}{2}\sum_i\sum_j (X_{ij} - X_{ji})^2$$

$$= \frac{1}{2}(2.130.0) = 1.065$$

ADEVA

F. de V.	g. de l.	S.C.	C.M.	Fc
Cruzamientos	63	41.007,48	650,90	
a. c. g.	7	36.528,34	5.218,30	103,37**
a. c. e.	28	3.414,14	121,93	2,40**
Recíprocos	28	1.065,00	38,04	<1
Error	126	6.360,60	50,48	

El valor de F altamente significativo para a.c.g., sugiere que la resistencia a la marchitez bacteriana, estaría dada principalmente por la acción genética aditiva. La progenie de clones que demostraron alta a.c.g., tendrían la mayor resistencia a la enfermedad.

Sin embargo, puesto que el valor de F es también altamente significativo para a.c.e., se debe tomar también en cuenta la contribución genética no aditiva, a la variancia total para resistencia a la enfermedad.

El Cuadro N° 19.5 presenta los valores estimados de a.c.g. y a.c.e. para los clones y, a continuación, se indica el método de cálculo para los efectos aditivo y no aditivo.

CUADRO No 19.5 Valores calculados de a.c.g. y a.c.e.

Aptitud Combinatoria Específica

Clon	2	3	4	5	6	7	8	a.c.g.
1	–9,33	+9,55	+1,17	+6,17	–2,83	+3,73	–2,83	–16,30
2		+6,85	+5,98	+4,98	–2,52	–3,45	–7,02	–12,11
3			–4,15	+4,85	–7,64	–8,08	+14,87	+27,52
4				–5,52	–2,02	–6,45	–4,02	+8,89
5					+5,98	+8,55	–10,02	+9,39
6						+7,55	–3,52	+12,39
7							–4,95	–1,67
8								–28,11
							Σ	0

19.6.5 Cálculo de a.c.g. para cada combinación, $(\widehat{gi})$.

$$\widehat{g}_i = \frac{1}{2p}(X_{i\cdot} + X_{\cdot j}) - \frac{1}{p^2}X_{\cdot\cdot}$$

$$\widehat{g}_1 = \frac{1}{2\times 8}(269 + 226) - \frac{1}{64}(3.023) = -16.30$$

$$\widehat{g}_2 = \frac{1}{2\times 8}(291 + 271) - \frac{1}{64}(3.023) = -12.11$$

$$\widehat{g}_3 = \frac{1}{2\times 8}(578 + 618) - \frac{1}{64}(3.023) = +27.52$$

$$\widehat{g}_4 = \frac{1}{2\times 8}(432 + 466) - \frac{1}{64}(3.023) = +8.89$$

$$\widehat{g}_5 = \frac{1}{2\times 8}(442 + 464) - \frac{1}{64}(3.023) = +9.39$$

$$\widehat{g}_6 = \frac{1}{2\times 8}(494 + 460) - \frac{1}{64}(3.023) = +12.39$$

$$\widehat{g}_7 = \frac{1}{2\times 8}(355 + 374) - \frac{1}{64}(3.023) = -1.67$$

$$\widehat{g}_8 = \frac{1}{2\times 8}(162 + 144) - \frac{1}{64}(3.023) = -28.11$$

19.6.6 Cálculo de a. c. e. para cada combinación, $\widehat{e}_{ij}$ (28 cruzamientos simples).

$$\widehat{e}_{ij} = \frac{1}{2}(X_{ij} + X_{ji}) - \frac{1}{2p}(X_{i\cdot} + X_{\cdot i} + X_{j\cdot} + X_{\cdot j}) + \frac{1}{p^2}X_{\cdot\cdot}$$

$$\frac{1}{p^2}X_{\cdot\cdot} = \frac{1}{64}(3.023) = 47.23$$

1. $e_{12} = \frac{1}{2}(9 + 10) - \frac{1}{2 \times 8}(269 + 226 + 291 + 271) + 47{,}23 = -9{,}33$

2. $e_{13} = \frac{1}{2}(72 + 64) - \frac{1}{16}(269 + 618 + 226 + 578) + 47{,}23 = +9{,}55$

3. $e_{14} = \frac{1}{2}(45 + 37) - \frac{1}{16}(269 + 466 + 226 + 432) + 47{,}23 = +1{,}17$

4. $e_{15} = \frac{1}{2}(56 + 37) - \frac{1}{16}(269 + 464 + 226 + 442) + 47{,}23 = +6{,}17$

5. $e_{16} = \frac{1}{2}(43 + 38) - \frac{1}{16}(269 + 460 + 226 + 494) + 47{,}23 = -2{,}83$

6. $e_{17} = \frac{1}{2}(34 + 32) - \frac{1}{16}(269 + 374 + 226 + 355) + 47{,}23 = +3{,}77$

7. $e_{18} = \frac{1}{2}(0 + 0) - \frac{1}{16}(269 + 144 + 226 + 162) + 47{,}23 = -2{,}83$

8. $e_{23} = \frac{1}{2}(77 + 62) - \frac{1}{16}(291 + 618 + 271 + 578) + 47{,}23 = +6{,}85$

9. $e_{24} = \frac{1}{2}(53 + 47) - \frac{1}{16}(291 + 466 + 432 + 271) + 4\ 7{,}23 = +5{,}98$

10. $e_{25} = \frac{1}{2}(55 + 44) - \frac{1}{16}(291 + 464 + 442 + 271) + 47{,}23 = +4{,}98$

11. $e_{26} = \frac{1}{2}(45 + 45) - 94{,}75 + 47{,}23 = -2{,}52$

12. $e_{27} = 30{,}0 - 80{,}68 + 47{,}23 = -3.45$

13. $e_{28} = 0 - 54{,}25 + 47{,}23 = -7.02$

14. $e_{34} = 79{,}5 - 130{,}88 + 47{,}23 = -4{,}15$

15. $e_{35} = 89{,}0 - 131{,}38 + 47{,}23 = +4{,}85$

16. $e_{36} = 79{,}5 - 134{,}37 + 47{,}23 = -7{,}64$

17. $e_{37} = 65{,}0 - 120{,}31 + 47{,}23 = -8{,}08$

18. $e_{38} = 61{,}5 - 93{,}86 + 47{,}23 = +14{,}87$

19. $e_{45} = 60{,}0 - 112{,}75 + 47{,}23 = -5{,}52$

20. $e_{46} = 66{,}5 - 115{,}75 + 47{,}23 = -2{,}02$

21. $e_{47} = 48{,}0 - 101{,}68 + 47{,}23 = -6{,}45$

22. $e_{48} = 24{,}0 - 75{,}25 + 47{,}23 = -4{,}02$

23. $e_{56} = 75{,}0 - 116{,}25 + 47{,}23 = +5{,}98$

24. $e_{57} = 63{,}5 - 102{,}18 + 47{,}23 = +8{,}55$

25. $e_{58} = 18{,}5 - 75{,}75 + 47{,}23 = -10.02$

26. $e_{67} = 65{,}5 - 105{,}18 + 47{,}23 = +7{,}55$

27. $e_{68} = 28{,}0 - 78{,}75 + 47{,}23 = -3{,}52$

28. $e_{78} = 12{,}5 - 64{,}68 + 47{,}23 = -4{,}95$

19.7 Conclusiones

Como se puede ver del Cuadro N° 19.5, algunos clones que tienen alta *a. c. g.*, presentan baja a. c. e.; si, dentro del método de multiplicación de semilla, la semilla del mejorador, la básica, registrada y certificada, van a corresponder a las generaciones *Sintético – 1, Sin – 2, Sin – 3* y *Sin – 4,* respectivamente, sería más importante basar la producción de sintéticos en *a. c. g.*

La selección basada en *a. c. e.*, daría sintéticos de alto rendimiento en *Sin – 1*, pero la disipación de los efectos no aditivos en las combinaciones respectivas, podría resultar en una declinación del rendimiento en las próximas generaciones.

La selección por a. c. e. tendría gran importancia, cuando el objetivo es producir híbridos en primera generación.

REFERENCIAS

COCHRAN, W. G. and COX, G. M. 1957. Experimental Designs. 2nd Ed. John Wiley & Sons. N. Y. 611 p.

DE LA LOMA, J. L. 1955. Experimentación Agrícola. Unión Tipográfica Editorial Hispano-Americana. México. 427 p.

DIXON, W. J. y MASSEY Jr., F. J. 1965. Introducción al Análisis Estadístico. Editora del Castillo, S. A. Madrid. 489 p.

FINNEY, D. J. 1946. Standard errors of yields adjusted for regression on an independent measurement. Biometrics. Bull. 2: 53-55.

GATES, C. E. & WILCOX, A. N. 1965. Analysis of variance of diallel crosses in spiit-plot experiments. Aust. J. Scie. 18: 837-850.

GRIFFING, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Scie. 9; 463-493.

GULLON, A. 1971. Introducción a la Estadística Aplicada. Edit. Alhambra, S. A. Madrid. 219 p.

HOEL, P. G. 1968. Estadística Elemental. Compañía Editora Continental, S. A. México. 400 p.

KEMPTHORNE, O. 1952. The Design and Analysis of Experiments. John Wiley & Sons, Inc. N. Y.

LE CLERG, E. L, LEONARD, W. H. and CLARK, A. G. 1962. Field Plot Technique. Burgess Publishng Company. Minneapolis, Minn. 373 p.

LYNCH, P. B. 1966. Conduct of Field Experiments. N9 399. New Zealand Dept. of Agric. 154 p.

OSTLE, B. 1965. Estadística Aplicada. Editorial Limusa-Wiley, S. A. México. 629 p.

PANSE, V. G. y SUKHATME. 1959. Métodos Estadísticos para investigadores Agrícolas. Fondo de Cultura Económica. México. 349 p.

PEARSON, E. S. 1936. Karl Pearson, an appreciation of some aspects of his life and work, Part II: 1906-1936. Biometrika 29: 161-248.

POEHLMAN, J. M. 1959. Breeding Field Crops. Henry Holt & Co., Inc. N. Y. p 21-23.

SNEDECOR, W. G. and COCHRAN, W. G. 1967. Statistical Methods 6th ed. The Iowa State Univ. Press Ames, Iowa. 593 p.

SPIEGEL, M. R. 1961. Theory and Problems of Statistics. Schaum Publishing Co. N. Y. 359 p.

SPRAGUE, G. F. and TATUM, L. A. 1942. General vs. specific combining ability in single crosses of corn. Jour. Amer. Soc. Agro. 34:923-932.

SRB, A. M., OWEN D. and EDGAR, R. S. 1965. General Genetics. 2nd ed. W. H. Freeman & Co. San Francisco, p 442-43.

STEEL, R. G. & TORRIE, J. H. 1960. Principies and Procedures of Statistics. McGraw-Hill Company, Inc. N. Y. 481 p.

APÉNDICE
LISTA DE TABLAS

I. NÚMEROS AL AZAR

II. CUADRADOS Y RAÍCES CUADRADAS.

III. ÁREAS BAJO LA CURVA NORMAL TÍPICA, DE 0 A Z.

IV. VALORES DE X^2.

V. VALORES DE **t**.

VI. VALORES DE **F**.

VII. VALORES DE RANGO MÍNIMO DE DUNCAN.

VIII. VALORES DE DUNNETT, PARA COMPARACIONES DE UNA COLA, DE DOS COLAS.

IX. VALORES DE TRASFORMACION ARCO-SENO.

X. VALORES DE CORRELACIÓN.

XI. COEFICIENTES DE POLINOMIOS ORTOGONALES.

XII. SÍMBOLOS DE USO COMÚN.

XIII. REFERENCIAS ADICIONALES SOBRE TEMAS ESTADÍSTICOS Y DISEÑO EXPERIMENTAL.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

TABLA No. I NÚMEROS AL AZAR

03991	10461	93716	16894	98953	73231	39528	72484	82074	25593
38555	95554	32886	59780	09958	18065	81616	18711	53342	44276
17546	73704	92052	46215	15917	06253	07586	16120	82641	22820
32643	52861	95819	06831	19640	99413	90767	04235	13574	17200
69572	68777	39510	35905	85244	35159	40188	28193	29593	88627
24122	66591	27699	06494	03152	19121	34414	82157	86887	55087
61196	30231	92962	61773	22109	78508	63439	75363	44989	16822
30532	21704	10274	12202	94205	20380	67049	09070	93399	45547
03788	97599	75867	20717	82037	10268	79495	04146	52162	90286
48228	63379	85783	47619	87481	37220	91704	30552	04737	21031
88618	19161	41290	67312	71857	15957	48545	35247	18619	13674
71299	23853	05870	01119	92784	26340	75122	11724	74627	73707
27954	58909	82444	99005	04921	73701	92904	13141	32392	19763
80863	00514	20247	81759	45197	25332	69902	63742	78464	22501
33564	60780	48460	85558	15191	18782	94972	11598	62095	36787
90899	75754	60833	25983	01291	41349	19152	00023	12302	80783
78038	70267	43529	06318	38384	74761	36024	00867	76378	41605
55986	66485	88722	56736	66164	49431	94458	74284	05041	49807
87539	08823	94813	31900	54155	83436	54158	34243	46978	35482
16818	60311	74457	90561	72848	11834	75051	93029	47665	64382
34677	58300	74910	64345	19325	81549	60365	94653	35075	33949
45305	07521	61318	31855	14413	70951	83799	42402	56623	34442
59747	67277	76503	34513	39663	77544	32960	07405	36409	83232
16520	69676	11654	99893	02181	68161	19322	53845	57620	52606
68652	27376	92852	55866	88448	03584	11220	94747	07399	37408
79375	95220	01159	63267	10622	48391	31751	57260	68980	05339
33521	26665	55823	47641	86225	31704	88492	99382	14454	04504
59589	49067	66821	41575	49767	04037	30934	47744	07481	83828
20554	91409	96277	48257	50816	97616	22888	48893	27499	98748
59404	72059	43947	51680	43852	59693	78212	16993	35902	91386
42614	29297	01918	28316	25163	01889	70014	15021	68971	11403
34994	41374	70071	14736	65251	07629	37239	33295	18477	65622
99385	41600	11133	07586	36815	43625	18637	37509	14707	93997
66497	68646	78138	66559	64397	11692	05327	82162	83745	22567
48509	23929	27482	45476	04515	25624	95096	67946	16930	33361
15470	48355	88651	22596	83761	60873	43253	84145	20368	07126
20094	98977	74843	93413	14387	06345	80854	09279	41196	37480
73788	06533	28597	20405	51321	92246	80088	77074	66919	31678
60530	45128	74022	84617	72072	00008	80890	18002	35352	54131
44372	15486	65741	14014	05466	55306	93128	18464	79982	68416
18611	19241	66083	24653	84609	58232	41849	84547	46850	52326
58319	15997	08355	60860	29735	47762	46352	33049	69248	93460
61199	67940	55121	29281	59076	07936	11087	96294	14013	31792
18627	90872	00911	98936	76355	93779	52701	08337	56303	87315
00441	58997	14060	40619	29549	69616	57275	36898	81304	48585
32624	68691	14845	46672	61958	77100	20857	73156	70284	24326
65961	73488	41839	55382	17267	70943	15633	84924	90415	93614
20288	34060	39685	23309	10061	68829	92694	48297	39904	02115
59362	95938	74416	53166	35208	33374	77613	19019	88152	00080
99782	93478	53152	67433	35663	52972	38688	32486	45134	63545

TABLA No. I NÚMEROS AL AZAR (Continuación 2)

27767	43584	85301	88977	29490	69714	94015	64874	32444	48277
13025	14338	54066	15243	47724	66733	74108	88222	88570	74015
80217	36292	98525	24335	24432	24896	62880	87873	95160	59221
10875	62004	90391	61105	57411	06368	11748	12102	80580	41867
54127	57326	26629	19087	24472	88779	17944	05600	60478	03343
60311	42824	37301	42678	45990	43242	66067	42792	95043	52680
49739	71484	92003	98086	76668	73209	54244	91030	45547	70818
78626	51594	16453	94614	39014	97066	30945	57589	31732	57260
66692	13986	99837	00582	81232	44987	69170	37403	86995	90307
44071	28091	07362	97703	76447	42537	08345	88975	35841	85771
59820	96163	78851	16499	87064	13075	73035	41207	74699	09310
25704	91035	26313	77463	55387	72681	47431	43905	31048	56699
22304	90314	78438	66276	18396	73538	43277	58874	11466	16082
17710	59621	15292	76139	59526	52113	53856	30743	08670	84741
25852	58905	55018	56374	35824	71708	30540	27886	61732	75454
46780	56487	75211	10271	36633	68424	17374	52003	70707	70214
59849	96169	87195	46092	26787	60939	59202	11973	02902	33250
47670	07654	30342	40277	11049	72049	83012	09832	25571	77628
94304	71803	73465	09819	58869	35220	09504	96412	90193	79568
08105	59987	21437	36786	49226	77837	98524	97831	65704	09514
64281	61826	18555	64937	64654	25843	41145	42820	14924	39650
66847	70495	32350	02985	01755	14750	48968	38603	70312	05682
72461	33230	21529	53424	72877	17334	39283	04149	90850	64618
21032	91050	13058	16218	06554	07850	73950	79552	24781	89683
95362	67011	06651	16136	57216	39618	49856	99326	40902	05069
49712	97380	10404	55452	09971	59481	37006	22186	72682	07385
58275	61764	97586	54716	61459	21647	87417	17198	21443	41808
89514	11788	68224	23417	46376	25366	94746	49580	01176	28838
15472	50669	48139	36732	26825	05511	12459	91314	80582	71944
12120	86124	51247	44302	87112	21476	14713	71181	13177	55292
95294	00556	70481	06905	21785	51101	49386	54480	23604	23554
66986	34099	74474	20740	47458	64809	06312	88940	15995	69321
80620	51790	11436	38072	40405	68032	60942	00307	11897	92674
55411	85667	77535	99892	71209	92061	92329	98932	78284	46347
95083	06783	28102	57816	85561	29671	77936	63574	31384	51924
90726	57166	98884	08583	95889	57067	38101	77756	11657	13897
68984	83620	89747	98882	92613	89719	39641	69457	91339	22502
36421	16489	18059	51061	67667	60631	84054	40455	99396	63680
92638	40333	67054	16067	24700	71594	47468	03577	57649	63266
21036	82808	77501	97427	76479	68562	43321	31370	28977	23896
13173	33365	41468	85149	49554	17994	91178	10174	29420	90438
86716	38746	94559	37559	49678	53119	98189	81851	29651	84215
92581	02262	41615	70360	64114	58660	96717	54244	10701	41393
12470	56500	50273	93113	41794	86861	39448	93136	25722	08564
01016	00857	41396	80504	90670	08289	58137	17820	22751	36518
34030	60726	25807	24260	71529	78920	47648	13885	70669	93406
50269	46345	06170	97965	88302	98041	11947	56203	19324	20504
73959	76145	60808	54444	74412	81105	69181	96845	38525	11600
46874	37088	80940	44893	10408	36222	14004	23153	69249	05747
60883	52109	19516	90120	46759	71643	62342	07589	08899	05985

EMPLEO DE LA TABLA DE RAÍCES CUADRADAS

Para obtener la raíz cuadrada de un número, separe las cifras en períodos de dos en dos a partir del punto decimal. Si el número es mayor que uno y el último período de la izquierda contiene solamente una cifra, use la columna $\sqrt{N}$; si el período contiene dos cifras, use $\sqrt{10N}$. Si el número es menor que uno y la primera cifra significativa a la derecha del punto decimal está precedida por un cero dentro de su período, use la columna de $\sqrt{N}$, si no, use $\sqrt{10N}$. Cada período de cifras en el número corresponde a una sola cifra en la respuesta. Ejemplos:

$\sqrt{241.67} = \sqrt{2|41.|67}$; entonces se emplea la columna $\sqrt{N}$ para $\sqrt{242.} = 15.6$

$\sqrt{.0024167} = \sqrt{00|24|16|7}$, entonces se emplea la columna $\sqrt{10N}$ para $\sqrt{.00242} = .0492$

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

TABLA No. II CUADRADOS Y RAÍCES CUADRADAS

N	N²	√N	√10N
1.00	**1.0000**	**1.00000**	**3.16228**
1.01	1.0201	1.00499	3.17805
1.02	1.0404	1.00995	3.19374
1.03	1.0609	1.01489	3.20936
1.04	1.0816	1.01980	3.22490
1.05	1.1025	1.02470	3.24037
1.06	1.1236	1.02956	3.25576
1.07	1.1449	1.03441	3.27109
1.08	1.1664	1.03923	3.28634
1.09	1.1881	1.04403	3.30151
1.10	**1.2100**	**1.04881**	**3.31662**
1.11	1.2321	1.05357	3.33167
1.12	1.2544	1.05830	3.34664
1.13	1.2769	1.06301	3.36155
1.14	1.2996	1.06771	3.37639
1.15	1.3225	1.07238	3.39116
1.16	1 .3456	1.07703	3.40588
1 .17	1.3689	1.08167	3.42053
1.18	1.3924	1.08628	3.43511
1.19	1.4161	1.09087	3.44964
1.20	**1.4400**	**1.09545**	**3.46410**
1.21	1.4641	1.10000	3.47851
1.22	1.4884	1.10454	3.49285
1.23	1.5129	1.10905	3.50714
1.24	1.5376	1.11355	3.52136
1.25	1.5625	1.11803	3.53553
1.26	1.5876	1.12250	3.54965
1.27	1.6129	1.12694	3.56371
1.28	1.6384	1.13137	3.57771
1.29	1.6641	1.13578	3.59166
1.30	**1.6900**	**1.14018**	**3.60555**
1.31	1.7161	1.14455	3.61939
1.32	1.7424	1.14891	3.63318
1.33	1.7689	1.15326	3.64692
1.34	1.7956	1.15758	3.66060
1.35	1.8225	1.16190	3.67423
1.36	1.8496	1.16619	3.68782
1.37	1.8796	1.17047	3.70135
1.38	1.9044.	1.17473	3.71484
1.39	1.9321	1.17898	3.72827
N	N²	√N	√10N

N	N²	√N	√10N
1.40	**1.9600**	**1.18322**	**3.74166**
1.41	1.9881	1.18743	3.75500
1.42	2.0164	1.19164	3.76829
1.43	2.0449	1.19583	3.78153
1 44	2.0736	1.20000	3.79473
1.45	2.1025	1.20416	3.80789
1.46	2.1316	1.20830	3.82099
1.47	2.1609	1.21244	3.83406
1.48	2.1904	1.21655	3.84708
1.49	2.2201	1.22066	3.86005
1.50	**2.2500**	**1.22474**	**3.87298**
1.51	2.2801	1.22882	3.88587
1.52	2.3104	1.23288	3.89872
1.53	2.3409	1.23693	3.91152
1.54	2.3716	1.24097	3.92428
1.55	2.4025	1.24499	3.93700
1.56	2.4336	1.24900	3.94968
1.57	2.4649	1.25300	3.96232
1.58	2.4964	1.25698	3.97492
1.59	2.5281	1.26095	3.98748
1.60	**2.5600**	**1.26491**	**4.00000**
1.61	2.5921	1.26886	4.01248
1.62	2.6244	1 .27279	4.02492
1.63	2.6569	1.27671	4.03733
1.64	2.6896	1.28062	4.04969
1.65	2.7225	1.28452	4.06202
1.66	2.7556	1.28841	4.07431
1.67	2.7889	1.29228	4.08656
1.68	2.8224	1.29615	4.09878
1.69	2.8561	1.30000	4.11096
1.70	**2.8900**	**1.30384**	**4.12311**
1.71	2.9241	1.30767	4.13521
1.72	2.9584	1.31149	4.14729
1.73	2.9929	1.31529	4.15933
1.74	3.0276	1.31909	4.17133
1.75	3.0625	1.32288	4.18330
1.76	3.0976	1.32665	4.19524
1.77	3.1329	1.33041	4.20714
1.78	3.1684	1.33417	4.21900
1.79	3.2041	1.33791	4.23084
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 2)

N	N²	√N	√10N
1.80	**3.2400**	**1.34164**	**4.24264**
1.81	3.2761	1.34536	4.25441
1.82	3.3124	1.34907	4.26615
1.83	3.3489	1.35277	4.27785
1.84	3.3856	1.35647	4.28952
1.85	3.4225	1.36015	4.30116
1.86	3.4596	1.36382	4.31277
1.87	3.4969	1.36748	4.32435
1.88	3.5344	1.37113	4.33590
1.89	3.5721	1.37477	4.34741
1.90	**3.6100**	**1.37840**	**4.35890**
1.91	3.6481	1.38203	4.37035
1.92	3.6864	1.38564	4.38178
1.93	3.7249	1.38924	4.39318
1.94	3.7636	1.39284	4.40454
1.95	3.8025	1.39642	4.41588
1.96	3.8416	1.40000	4.42719
1.97	3.8809	1.40357	4.43847
1.98	3.9204	1.40712	4.44972
1.99	3.9601	1.41067	4.46094
2.00	**4.0000**	**1.41421**	**4.47214**
2.01	4 0401	1.41774	4.48330
2.02	4.0804	1.42127	4.49444
2.03	4.1209	1.42478	4.50555
2.04	4.1616	1.42829	4.51664
2.05	4.2025	1.43178	4.52769
2.06	4.2436	1.43527	4.53872
2.07	4.2849	1.43875	4.54973
2.08	4.3264	1.44222	4.56070
2.09	4.3681	1.44568	4.57165
2.10	**4.4100**	**1.44914**	**4.58258**
2.11	4.4521	1.45258	4.59347
2.12	4.4944	1.45602	4.60435
2 13	4.5369	1.45945	4.61519
2.14	4.5796	1.46287	4.62601
2.15	4.6225	1.46629	4.63681
2 16	4.6656	1.46969	4.64758
2.17	4.7089	1.47309	4.65833
2.18	4.7524	1.47648	4.66905
2.19	4.7961	1.47986	4.67974
N	N²	√N	√10N

N	N²	√N	√10N
2.20	**4.8400**	**1.48324**	**4.69042**
2.21	4.8841	1.48661	4.70106
2.22	4.9284	1.48997	4.71169
2.23	4.9729	1.49332	4.72229
2.24	5.0176	1.49666	4.73286
2 25	5.0625	1.50000	4.74342
2.26	5.1076	1.50333	4.75395
2.27	5.1529	1.50665	4.76445
2.28	5.1984	1.50997	4.77493
2.29	5.2441	1.51327	4.78539
2.30	**5.2900**	**1.51658**	**4.79583**
2.31	5.3361	1.51987	4.80625
2.32	5.3824	1.52315	4.81664
2.33	5.4289	1.52643	4.82701
2.34	5.4756	1.52971	4.83735
2.35	5.5225	1.53297	4.84768
2.36	5.5696	1.53623	4.85798
2.37	5.6169	1.53948	4.86826
2.38	5.6644	1.54272	4.87852
2.39	5.7121	1.54596	4.88876
2.40	**5.7600**	**1.54919**	**4.89898**
2.41	5.8081	1.55242	4.90918
2.42	5.8564	1.55563	4.91935
2.43	5.9049	1.55885	4.92950
2.44	5.9536	1.56205	4.93964
2.45	6.0025	1.56525	4.94975
2.46	6.0516	1.56844	4.95984
2.47	6.1609	1.57162	4.96991
2.48	6.1504	1.57480	4.97996
2.49	6.2001	1.57797	4.98999
2.50	**6.2500**	**1.58114**	**5.00000**
2.51	6.3001	1.58430	5.00999
2.52	6.3504	1.58745	5.01996
2.53	6.4009	1.59060	5.02991
2.54	6.4516	1.59374	5.03984
2.55	6.5025	1.59687	5.04975
2.56	6.5336	1.60000	5.05964
2.57	6.6049	1.60312	5.06952
2.58	6.6564	1.60624	5.07937
2.59	6.7081	1.60935	5.08920
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 3)

N	N²	√N	√10N
2.60	**6.7600**	**1.61245**	**5.09902**
2.61	6.8121	1.61555	5.10882
2.62	6.8644	1.61864	5.11859
2.63	6.9169	1.62173	5.12835
2.64	6.9696	1.62481	5.13809
2 65	7.0225	1.62788	5.14782
2.66	7.0756	1.63095	5.15752
2.67	7.1289	1.63401	5.16720
2.68	7.1824	1.63707	5.17687
2.69	7.2361	1.64012	5.18652
2.70	**7.2900**	**1.64317**	**5.19615**
2 71	7.3441	1.64621	5.20577
2.72	7.3984	1.64924	5.21536
2.73	7.4529	1.65227	5.22494
2.74	7.5076	1.65529	5.23450
2.75	7.5625	1.65831	5.24404
2.76	7.6176	1.66132	5.25357
2.77	7.6729	1.66433	5.26308
2.78	7.7284	1.66733	5.27257
2.79	7.7841	1.67033	5.28205
2.80	**7.8400**	**1.67332**	**5.29150**
2.81	7.8961	1.6/631	5.30094
2.82	7.9524	1.67929	5.31437
2.83	8.0089	1.68226	5.31977
2.84	8.0656	1.68523	5.32917
2.85	8.1225	1.68819	5.33854
2.86	8.1796	1.69115	5.34790
2.87	8.2369	1.69411	5.35724
2.88	8.2944	1.69706	5.36656
2.89	8.3521	1.70000	5.37587
2.90	**8.4100**	**1.70294**	**5.38516**
2.91	8.4681	1.70587	5.39444
2.92	8.5264	1.70880	5.40370
2.93	8.5849	1.71172	5.41295
2.94	8.6436	1.71464	5.42218
2.95	8.7025	1.71756	5.43139
2.96	8.7616	1.72047	5.44059
2.97	8.8209	1.72337	5.44977
2.98	8.8804	1.72627	5.45894
2.99	8.9401	1.72916	5.46809
N	N²	√N	√10N

N	N²	√N	√10N
3.00	**9.0000**	**1.73205**	**5.47723**
3.01	9.0601	1.73494	5.48635
3.02	9.1204	1.73781	5.49545
3.03	9.1809	1.74069	5.50454
3.04	9.2416	1.74356	5.51362
3.05	9.3025	1.74642	5.52268
3.06	9.3636	1.74929	5.53173
3.07	9.4249	1.75214	5.54076
3.08	9.4864	1.75499	5.54977
3.09	9.5481	1.75784	5.55878
3.10	**9.6100**	**1.76068**	**5.56776**
3.11	9.6721	1.76352	5.57674
3.12	9.7344	1.76635	5.58570
3.13	9.7969	1.76918	5.59464
3.14	9.8596	1.77200	5.60357
3.15	9.9225	1.77482	5.61249
3.16	9.9856	1.77764	5.62139
3.17	10.0489	1.78045	5.53028
3.18	10.1124	1.78326	5.63915
3.19	10.1761	1.78606	5.64801
3.20	**10.2400**	**1.78885**	**5.65685**
3.21	10.3041	1.79165	5.66569
3.22	10.3684	1.79444	5.67450
3.23	10.4329	1.79722	5.68331
3.24	10.4976	1.80000	5.69210
3.25	10.5625	1.80278	5.70088
3.26	10.6276	1.80555	5.70964
3.27	10.6929	1.80831	5.71839
3.28	10.7584	1.81108	5.72713
3.29	10.8241	1.81384	5.73585
3.30	**10.8900**	**1.81659**	**5.74456**
3.31	10.9561	1.81934	5.75326
3.32	11.0224	1.82209	5.76194
3.33	11.0889	1.82483	5.77062
3.34	11.1556	1.82757	5.77927
3.35	11.2225	1.83030	5.78792
3.36	11.2896	1.83303	5.79655
3.37	11.3569	1.83576	5.80517
3.38	11.4244	1.83848	5.81378
3.39	11.4921	1.84120	5.82237
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 4)

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
3.40	**11.5600**	**1.84391**	**5.83095**
3.41	11.6281	1.84662	5.83952
3.42	11.6964	1.84932	5.84808
3.43	11.7649	1.85203	5.85662
3.44	11.8336	1.85472	5.86515
3.45	11.9025	1.85742	5.87367
3.46	11.9716	1.86011	5.88218
3.47	12.0409	1.86279	5.89067
3.48	12.1104	1.86815	5.89915
3.49	12.1801	1.86815	5.90762
3.50	**12.2500**	**1.87083**	**5.91608**
3.51	12.3201	1.87350	5.92453
3.52	12.3904	1.87617	5.93296
3.53	12.4609	1.87883	5.94138
3.54	12.5316	1.88149	5.94979
3.55	12.6025	1.88414	5.95819
3.56	12.6736	1.88680	5.96657
3.57	12.7449	1.88944	5.97495
3.58	12.8164	1.89209	5.98331
3.59	12.8881	1.89473	5.99166
3.60	**12.9600**	**1.89737**	**6.00000**
3.61	13.0321	1.90000	6.00833
3.62	13.1044	1.90263	6.01664
3.63	13.1769	1.90526	6.02495
3.64	13.2496	1.90788	6.03324
3.65	13.3225	1.91050	6.04152
3.66	13.3956	1.91311	6.04979
3.67	13.4689	1.91572	6.05805
3.68	13.5424	1.91833	6.06630
3.69	13.6161	1.92094	6.07454
3.70	**13.6900**	**1.92354**	**6.08276**
3.71	13.7641	1.92614	6.09098
3.72	13.8384	1.92873	6.09918
3.73	13.9129	1.93132	6.10737
3.74	13.9876	1.93391	6.11555
3.75	14.0625	1.93649	6.12372
3.76	14.1376	1.93907	6.13188
3.77	14.2129	1.94165	6.14003
3.78	14.2884	1.94422	6.14817
3.79	14.3641	1.94679	6.15630
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
3.80	**14.4400**	**1.94936**	**6.16441**
3.81	14.5161	1.95192	6.17252
3.82	14.5924	1.95448	6.18061
3.83	14.6689	1.95704	6.18870
3.84	14.7456	1.95959	6.19677
3.85	14.8225	1.96214	6.20484
3.86	14.8996	1.96469	6.21289
3.87	14.9769	1.96723	6.22093
3.88	15.0544	1.96977	6.22896
3.89	15.1321	1.97231	6.23699
3.90	**15.2100**	**1.97484**	**6.24500**
3.91	15.2881	1.97737	6.25300
3.92	15.3664	1.97990	6.26099
3.93	15.4449	1.98242	6.26897
3.94	15.5236	1.98494	6.27694
3.95	15.6025	1.98746	6.28490
3.96	15.6816	1.98997	6.29285
3.97	15.7609	1.99249	6.30079
3.98	15.8404	1.99499	6.30872
3.99	15.9201	1.99750	6.31664
4.00	**16.0000**	**2.00000**	**6.32456**
4.01	16.0801	2.00250	6.33246
4.02	16.1604	2.00499	6.34035
4.03	16.2409	2.00749	6.34823
4.04	16.3216	2.00998	6.35610
4.05	16.4025	2.01246	6.36396
4.06	16.4836	2.01494	6.37181
4.07	16.5649	2.01742	6.37966
4.08	16.6464	2.01990	6.38749
4.09	16.7281	2.02237	6.39531
4.10	**16.8100**	**2.02485**	**6.40312**
4.11	16.8921	2.02731	6.41093
4.12	16.9744	2.02978	6.41872
4.13	17.0569	2.03224	6.42651
4.14	17.1396	2.03470	6.43428
4.15	17.2225	2.03715	6.44205
4.16	17.3056	2.03961	6.44981
4.17	17.3889	2.04206	6.45755
4.18	17.4724	2.04450	6.46529
4.19	17.5561	2.04695	6.47302
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

CUADRADOS Y RAÍCES CUADRADAS (Continuación 5)

N	N²	√N	√10N
4.20	**17.6400**	**2.04939**	**6.48074**
4.21	17.7241	2.05183	6.48845
4.22	17.8084	2.05426	6.49615
4.23	17.8929	2.05670	6.50384
4.24	17.9776	2.05913	6.51153
4.25	18.0625	2.06155	6.51920
4.26	18.1476	2.06398	6.52687
4.27	18.2329	2.06640	6.53452
4.28	18.3184	2.06882	6.54217
4.29	18.4041	2.07123	6.54981
4.30	**18.4900**	**2.07364**	**6.55744**
4.31	18.5761	2.07605	6.56506
4.32	18.6624	2.07846	6.57267
4.33	18.7489	2.08087	6.58027
4.34	18.8356	2.08327	6.58787
4.35	18.9225	2.08567	6.59545
4.36	19.0096	2.08806	6.60303
4.37	19.0969	2.09045	6.61060
4.38	19.1844	2.09284	6.61816
4.39	19.2721	2.09523	6.62571
4.40	**19.3600**	**2.09762**	**6.63325**
4.41	19.4481	2.10000	6.64078
4.42	19.5364	2.10238	6.64831
4.43	19.6249	2.10476	6.65582
4.44	19.7136	2.10713	6.66333
4.45	19.8025	2.10950	6.67083
4.46	19.8916	2.11187	6.67832
4.47	19.9809	2.11424	6.68581
4.48	20.0704	2.11660	6.69328
4.49	20.1601	2.11896	6.70075
4.50	**20.2500**	**2.12132**	**6.70820**
4.51	20.3401	2.12368	6.71565
4.52	20.4304	2.12603	6.72309
4.53	20.5209	2,12838	6.73053
4.54	20.6116	2.13073	6.73795
4.55	20.7025	2.13307	6.74537
4.56	20.7936	2.13542	6.75278
4.57	20.8849	2.13776	6.76018
4.58	20.9764	2.14009	6.76757
4.59	21.0681	2.14243	6.77495
N	N²	√N	√10N

N	N²	√N	√10N
4.60	**21.1600**	**2.14476**	**6.78233**
4.61	21.2521	2.14709	6.78970
4.62	21.3444	2.14942	6.79706
4.63	21.4369	2.15174	6.80441
4.64	21.5296	2.15407	6.81175
4.65	21.6225	2.15639	6.81909
4.66	21.7156	2.15870	6.82642
4.67	21.8089	2.16102	6.83374
4.68	21.9024	2.16333	6.84105
4.69	21.9961	2.16564	6.84836
4.70	**22.0900**	**2.16795**	**6.85565**
4.71	22.1841	2.17025	6.86294
4.72	22.2784	2.17256	6.87023
4.73	22,3729	2.17486	6.87750
4.74	22.4676	2.17715	6.88477
4,75	22.5625	2.17945	6.89202
4.76	22.6576	2.18174	6.89928
4.77	22,7529	2.18403	6.90652
4.78	22.8484	2.18632	6.91375
4.79	22.9441	2.18861	6.92098
4 80	**23.0400**	**2.19089**	**6.92820**
4.81	23,1361	2.19317	6.93542
4.82	23.2324	2.19545	6,94262
4.83	23.3289	2.19773	6.94982
4.84	23.4256	2.20000	6.95701
4.85	23.5225	2.20227	6.96419
4.86	23,6196	2.20454	6.87137
4.87	23.7169	2.20681	6.97854
4.88	23.8144	2.20907	6.98570
4.89	23.9121	2.21133	6.99285
4.90	**24.0100**	**2.21359**	**7.00000**
4.21	24.1081	2.21585	7.00714
4.92	24.2064	2.21811	7.01427
4.93	24.3049	2.22036	7.02140
4.94	24.4036	2.22261	7.02851
4.95	24.5025	2.22486	7.03562
4.96	24.6016	2.22711	7.04273
4.97	24.7009	2.22935	7.04982
4.98	24 8004	2.23159	7.05691
4.99	24.9001	2.23383	7.06399
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 6)

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
5.00	**25.0000**	**2.23607**	**7.07107**
5.01	25.1001	2.23830	7.07814
5.02	25.2004	2.24054	7.08520
5.03	25.3009	2.24277	7.09225
5.04	25.4016	2.24499	7.09930
5.05	25.5025	2.24722	7.10634
5.06	25.6036	2.24944	7.11337
5.07	25.7049	2.25167	7.12039
5.08	25.8064	2.25389	7.12741
5.09	25.9081	2.25610	7.13442
5.10	**26.0100**	**2.25832**	**7.14143**
5.11	26.1121	2.26053	7.14843
5.12	26.2144	2.26274	7.15542
5.13	26.3169	2.26495	7.16240
5.14	26.4196	2.26716	7.16938
5.15	26.5225	2.26936	7.17635
5.16	26.6256	2.27156	7.18331
5.17	26.7289	2.27376	7.19027
5.18	26.8324	2.27596	7.19722
5.19	26.9361	2.27816	7.20417
5.20	**27.0400**	**2.28035**	**7.21110**
5.21	27.1441	2.28254	7.21803
5.22	27.2484	2.28473	7.22496
5.23	27.3529	2.28692	7.23187
5.24	27.4576	2.28910	7.23878
5.25	27.5625	2.29129	7.24569
5.26	27.6676	2.29347	7.25259
5.27	27.7729	2.29565	7.25948
5 28	27.8784	2.29783	7.26636
5.29	27.9841	2.30000	7.27324
5.30	**28.0900**	**2.30217**	**7.28011**
5.31	28.1961	2.30434	7.28697
5.32	28.3024	2.30651	7.29383
5.33	28.4089	2.30868	7.30068
5.34	28.5156	2.31084	7.30753
5.35	28.6225	2.31301	7.31437
5.36	28.7296	2.31517 l	7.32120
5.37	28.8369	2.31733 1	7.32803
5.38	28.9444	2.31948	7.33485
5.39	29.0521	2.32164	7.34166
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
5.40	**29.1600**	**2.32379**	**7.34847**
5.41	29.2681	2.32594	7.35527
5.42	29.3764	2.32809	7.36206
5.43	29.4849	2.33024	7.36885
5.44	29.5936	2.33238	7.37564
5.45	29.7025	2.33452	7.38241
5.46	29.8116	2.33666	7.38918
5.47	29.9209	2.33880	7.39594
5.48	30.0304	2.34094	7.40270
5.49	30.1401	2.34307	7.40945
5.50	**30.2500**	**2.34521**	**7.41620**
5.51	30.3601	2.34734	7.42294
5.52	30.4704	2.34947	7.42967
5.53	30.5809	2,35160	7.43640
5.54	30.6915	2.35372	7.44312
5.55	30.8025	2.35584	7.44983
5.56	30.9136	2.35797	7.45654
5.57	31.0249	2.36008	7.46324
5.58	31.1364	2.36220	7.46994
5.59	31.2481	2.36432	7.47663
5.60	**31.3600**	**2.36643**	**7.48331**
5.61	31.4721	2.36854	7.48999
5.62	31.5844	2.37065	7.49667
5.63	31.6969	2.37276	7.50333
5.64	31.8096	2.37487	7.50999
5.65	31.9225	2.37597	7.51665
5.66	32.0356	2.37908	7.52330
5.67	32.1489	2.38118	7.52994
5.68	32.2624	2.38328	7.53658
5.69	32.3761	2.38537	7.54321
5.70	**32.4900**	**2.38747**	**7.54983**
5.71	32.6041	2.38956	7.55645
5.72	32.7184	2.39165	7.56307
5.73	32.8329	2.39374	7.56968
5.74	32.9476	2.39583	7.57628
5.75	33.0625	2.39792	7.58288
5.76	33.1776	2.40000	7.58947
5.77	33.2929	2.40208	7.59605
5.78	33.4084	2.40416	7.60263
5.79	33.5241	2.40624	7.60920
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

CUADRADOS Y RAÍCES CUADRADAS (Continuación 7)

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
5.80	**33.6400**	**2.40832**	**7.61577**
5.81	33.7561	2.41039	7.62234
5.82	33.8724	2.41247	7.62889
5.83	33.9889	2.41454	7.63544
5.84	34.1056	2.41661	7.64199
5.85	34.2225	2.41868	7.64853
5.86	34.3396	2.42074	7.65506
5.87	34.4569	2.42281	7.66159
5.88	34.5744	2.42487	7.66812
5.89	34.6921	2.42693	7.67463
5.90	**34.8100**	**2.42899**	**7.68115**
5.91	34.9281	2.43105	7.68765
5.92	35.0464	2.43311	7.69415
5.93	35.1649	2.43516	7.70065
5.94	35.2836	2.43721	7.70714
5.95	35.4025	2.43926	7.71362
5.96	35.5216	2.44131	7.72010
5.97	35.6409	2.44336	7.72658
5.98	35.7604	2.44540	7.73305
5.99	35.8801	2.44745	7.73951
6.00	**36.0000**	**2.44949**	**7.74597**
6.01	36.1201	2.45153	7.75242
6.02	36.2404	2.45357	7.75887
6.03	36.3609	2.45561	7.76531
6.04	36.4816	2.45764	7.77174
6.05	36.6025	2.45967	7.77817
6.06	36.7236	2.46171	7.78460
6.07	36.8449	2.46374	7.79102
6.08	36.9664	2.46577	7.79744
6.09	37.0881	2.46779	7.80385
6.10	**37.2100**	**2.46982**	**7.81025**
6.11	37.3321	2.47184	7.81665
6.12	37.4544	2.47386	7.82304
6.13	37.5769	2.47588	7.82943
6.14	37.6996	2.47790	7.83582
6.15	37.8225	2.47992	7.84219
6.16	37.9456	2.48193	7.84857
6.17	38.0689	2.48395	7.85493
6.18	38.1924	2.48596	7.86130
6.19	38.3161	2.48797	7.86766
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

N	N^2	$\sqrt{N}$	$\sqrt{10N}$
6.20	**38.4400**	**2.48998**	**7.87401**
6.21	38.5641	2.49199	7.88036
6.22	38.6884	2.49399	7.88670
6.23	38.8129	2.49600	7.89303
6.24	38.9376	2.49800	7.89937
6.25	39.0625	2.50000	7.90569
6.26	39.1876	2.50200	7.91202
6.27	39.3129	2.50400	7.91833
6.28	39.4384	2.50599	7.92465
6.29	39.5641	2.50799	7.93095
6.30	**39.6900**	**2.50998**	**7.93725**
6.31	39.8161	2.51197	7.94355
6.32	39.9424	2.51396	7.94984
6.33	40.0689	2.51595	7.95613
6.34	40.1956	2.51794	7.96241
6.35	40.3225	2.51992	7.96869
6.36	40.4496	2.52190	7.97496
6.37	40.5769	2.52389	7.98123
6.38	40.7044	2.52587	7.98749
6.39	40.8321	2.52784	7.99375
6.40	**40.9600**	**2.52982**	**8.00000**
6.31	41.0881	2.53180	8.00625
6.42	41.2164	2.53377	8.01249
6.43	41.3449	2.53574	8.01873
6.44	41.4736	2.53772	8.02496
6.45	41.6025	2.53969	8.03119
6.46	41.7316	2.54165	8.03741
6.47	41.8609	2.54362	8.04363
6.48	41.9904	2.54558	8.04984
6.49	42.1201	2.54755	8.05605
6.50	**42.2500**	**2.54951**	**8.06226**
6.51	36.1201	2.45153	7.75242
6.02	36.2404	2.45357	7.75887
6.03	36.3609	2.45561	7.76531
6.04	36.4816	2.45764	7.77174
6.05	36.6025	2.45967	7.77817
6.06	36.7236	2.46171	7.78460
6.07	36.8449	2.46374	7.79102
6.08	36.9664	2.46577	7.79744
6.09	37.0881	2.46779	7.80385
N	N^2	$\sqrt{N}$	$\sqrt{10N}$

CUADRADOS Y RAÍCES CUADRADAS (Continuación 9)

N	N²	√N	√10N
6.90	**47.6100**	**2.62679**	**8.30662**
6.91	47.7481	2.62869	8.31264
6.92	47.8864	2.63059	8.31865
6.93	48.0249	2.63249	8.32466
6.94	48.1636	2.63439	8.33067
6.95	48.3025	2.63629	8.33667
6.96	48.4416	2.63818	8.34266
6.97	48.5809	2.64008	8.34865
6.98	48.7204	2.64197	8.35464
6.99	48.8601	2.64386	8.36062
7.00	**49.0000**	**2.64575**	**8.36660**
7.01	49.1401	2.64764	8.37257
7.02	49.2804	2.64953	8.37854
7.03	49.4209	2.65141	8.38451
7.04	49.5616	2.65330	8.39047
7.05	49.7025	2.65518	8.39643
7.06	49.8436	2.65707	8.40238
7.07	49.9849	2.65895	8.40833
7.08	50.1264	2.66083	8.41427
7.09	50.2681	2.66271	8.42021
7.10	**50.4100**	**2.66458**	**8.42615**
7.11	50.5521	2.66646	8.43208
7.12	50.6944	2.66833	8.43801
7.13	50.8369	2.67021	8.44393
7.14	50.9796	2.67208	8.44985
7.15	51.1225	2.67395	8.45577
7.16	51.2656	2.67582	8.46168
7.17	51.4089	2.67769	8.46759
7.18	51.5524	2.67955	8.47349
7.19	51.6961	2.68142	8.47939
7.20	**51.8400**	**2.68328**	**8.48528**
7.21	51.9841	2.68514	8.49117
7.22	52.1284	2.68701	8.49706
7.23	52.2729	2.68887	8.50294
7.24	52.4176	2.69072	8.50882
7.25	53.5625	2.69258	8.51469
7.26	52.7076	2.69444	8.52056
7.27	52.8529	2.69629	8.52643
7.28	52.9984	2.69815	8.53229
7.29	53.1441	2.70000	8.53815
N	N²	√N	√10N

N	N²	√N	√10N
7.30	**53.2900**	**2.70185**	**8.54400**
7.31	53.4361	2.70370	8.54985
7.32	53.5824	2.70555	8.55570
7.33	53.7289	2.70740	8.56154
7.34	53.8756	2.70924	8.56738
7.35	54.0225	2.71109	8.57321
7.36	54.1696	2.71293	8.57904
7.37	54.3169	2.71477	8.58487
7.38	54.4644	2.71662	8.59069
7.39	54.6121	2.71846	8.59651
7.40	**54.7600**	**2.72029**	**8.60233**
7.41	54.9081	2.72213	8.60814
7.42	55.0564	2.72397	8.61394
7.43	55.2049	2.72580	8.61974
7.44	55.3536	2.72764	8.62554
7.45	55.5025	2.72947	8.63134
7.46	55.6516	2.73130	8.63713
7.47	55.8009	2.73313	8.64292
7.48	55.9504	2.73496	8.64870
7.49	56.1001	2.73679	8.65448
7.50	**56.2500**	**2.73861**	**8.66025**
7.51	56.4001	2.74044	8.66603
7.52	56.5504	2.74226	8.67179
7.53	56.7009	2.74408	8.67756
7.54	56.8516	2.74591	8.68332
7.55	57.0025	2.74773	8.68907
7.56	57.1536	2.74955	8.69483
7.57	57.3049	2.75136	8.70057
7.58	57.4564	2.75318	8.70632
7.59	57.6081	2.75500	8.71206
7.60	**57.7600**	**2.75681**	**8.71780**
7.61	57.9121	2.75862	8.72353
7.62	58.0644	2.76043	8.72926
7.63	58.2169	2.76225	8.73499
7.64	58.3696	2.76405	8.74071
7.65	58.5225	2.76586	8.74643
7.66	58.6756	2.76767	8.75214
7.67	58.8289	2.76948	8.75785
7.68	58.9824	2.77128	8.76356
7.69	59.1361	2.77308	8.76926
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 10)

N	N²	√N	√10N
7.70	**59.2900**	**2.77489**	**8.77496**
7.71	59.4441	2.77669	8.78066
7.72	59.5984	2.77849	8.78635
7.73	59.7529	2.78029	8.79204
7.74	59.9076	2.78209	8.79773
7.75	60.0625	2.78388	8.80341
7.76	60.2176	2.78568	8.80909
7.77	60.3729	2.78747	8.81476
7.78	60.5284	2.78927	8.82043
7.79	60.6841	2.79106	8.82610
7.80	**60.8400**	**2.79285**	**8.83176**
7.81	60.9961	2.79464	8.83742
7.82	61.1524	2.79643	8,84308
7.83	61.3089	2.79821	8.84873
7.84	61.4656	2.80000	8.85438
7.85	61.6225	2.80179	8.86002
7.86	61.7796	2.80357	8.86566
7.87	61.9369	2.80535	8.87130
7.88	62.0944	2.80713	8.87694
7.89	62.2521	2.80891	8.88257
7.90	**62.4100**	**2.81069**	**8.88819**
7.91	62.5681	2.81247	8.89382
7.92	62.7264	2.81425	8.89944
7.93	62.8849	2.81603	8.90505
7.94	63.0436	2.81780	8.91067
7.95	63.2025	2.811957	8.91628
7.96	63.3616	2.82135	8.92188
7.97	63.5209	2.82312	8,92749
7.98	63.6804	2.82489	8.93308
7.99	63.8401	2.82666	8.93868
8.00	**64.0000**	**2.82843**	**8.94427**
8.01	64.1601	2.83019	8.94986
8.02	64.3204	2.83196	8.95546
8.03	64.4809	2.83373	8.96103
8.04	64.6416	2.83549	8.96660
8.05	64.8025	2.83725	8.97218
8.06	64.9636	2.83901	8.97775
8.07	65.1249	2.84077	8.98332
8.08	65.2864	2.84253	8.98888
8.09	64.4481	2.84429	8.99444
N	N²	√N	√10N

N	N²	√N	√10N
8.10	**65.6100**	**2.84605**	**9.00000**
8.11	65.7721	2.84781	9.00555
8.12	65.9344	2.84956	9.01110
8.13	66.0969	2.85132	9.01665
8.14	66.2596	2.85307	9.02219
8.15	66.4225	2.85482	9.02774
8.16	66.5856	2.85657	9.03327
8.17	66.7489	2.85832	9.03881
8.18	66.9124	2.86007	9.04434
8.19	67.0761	2.86182	9.04986
8.20	**67.2400**	**2.86356**	**9.05539**
8.21	67.4041	2.86531	9.06091
8.22	67.5684	2.86705	9.06642
8.23	67.7329	2.86880	9.07193
8.24	67.8976	2.87054	9.07744
8.25	68.0625	2.87228	9.08295
8.26	68.2276	2.87402	9.08845
8.27	68.3929	2.87576	9.09395
8.28	68.5584	2.87750	9.09945
8.29	68.7241	2.87924	9.10494
8.30	**68.8900**	**2.88097**	**9.11043**
8.31	69.0561	2.88271	9.11592
8.32	69.2224	2.88444	9.12140
8.33	69.3889	2.88617	9.12688
8.34	69.5556	2.88791	9.13236
8.35	69.7225	2.88964	9.13783
8.36	69.8896	2.89137	9.14330
8.37	70.0569	2.89310	9.14877
8.38	70.2244	2.89482	9.15423
8.39	70.3921	2.89655	9.15969
8.40	**70.5600**	**2.89828**	**9.16515**
8.41	70.7281	2.90000	9.17061
8.42	70.8964	2.90172	9.17606
8.43	71.0649	2.90345	9.18150
8.44	71.2336	2.90517	9.18695
8.45	71.4025	2.90689	9.19239
8.46	71.5716	2.90861	9.19783
8.47	71.7409	2.91033	9.20326
8.48	71.9104	2.91204	9.20869
8.49	72.0801	2.91376	9.21412
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 12)

N	N²	√N	√10N
8.50	**72.2500**	**2.91548**	**9.21954**
8.51	72.4201	2.91719	9.22497
8.52	72.5904	2.91890	9.23038
8.53	72.7609	2.92062	9.23580
8.54	72.9316	2.92233	9.24121
8.55	73.1025	2.92404	9.24662
8.56	73.2736	2.92575	9.25203
8.57	73.4449	2.92746	9.25743
8.58	73.6164	2.92916	9.26283
8.59	73.7881	2.93087	9.26823
8.60	**73.9600**	**2.93258**	**9.27362**
8.61	74.1321	2.93428	9.27901
8.62	74.3044	2.93598	9.28440
8.63	74.4769	2.93769	9.28978
8.64	74.6496	2.93939	9.29516
8.65	74.8225	2.94109	9.30054
8.66	74.9956	2.94279	9.30591
8.67	75.1687	2.94449	9.31128
8.68	75.3424	2.94618	9.31665
8.69	75.5161	2.94788	9.32202
8.70	**75.6900**	**2.94958**	**9.32738**
8.71	75.8641'	2.95127	9.33274
8.72	76.0384	2.95296	9.33809
8.73	76.2129	2.95466	9.34345
8.74	76.3876	2.95635	9.34880
8.75	76.5625	2.95804	9.35414
8.76	76.7376	2.95973	9.35949
8.77	76.9129	2.96142	9.36483
8.78	77.0884	2.96311	9.37017
8.79	77.2641	2.96479	9.37550
8.80	**77.4400**	**2.96648**	**9.38083**
8.81	77.6161	2.96816	9.38616
8.82	77.7924	2.96985	9.39149
8.83	77.9689	2.97153	9.39681
8.84	78.1456	2.97321	9.40213
8.85	78.3225	2.97489	9.40744
8.86	78.4996	2.97658	9.41276
8.87'	78.6769	2.97825	9.41807
8.88	78.8544	2.97993	9.42338
8.89	79.0321	2.98161	9.42868
N	N²	√N	√10N

N	N²	√N	√10N
8.90	**79.2100**	**2.98329**	**9.43398**
8.91	79.3881	2.98496	9.43928
8.92	79.5664	?.98664	9.44458
8.93	79.7449	2.98831	9.44987
8.94	79.9236	2.98998	9.45516
8.95	80.1025	2.99166	9.46044
8.96	80.2816	2.99333	9.46573
8.97	80.4609	2.99500	9.4710T
8.98	80.6404	2.99666	9.47629
8.99	80.8201	2.99833	9.48156
9.00	**81.0000**	**3.00000**	**9.48683**
9.01	81.1801	3.00167	9.49210
9.02	81.3604	3.00333	9.49737
9.03	81.5409	3.00500	9.50263
9.04	81.7216	3.00666	9.50789
9.05	81.9025	3.00832	9.51315
9.06	82.0836	3.00998	9.51840
9.07	82.2649	3.01164	9.52365
9.08	82.4464	3.01330	9.52890
9.09	82.6281	3.01496	9.53415
9.10	**82.8100**	**3.01662**	**9.53939**
9.11	82.9921	3.01828	9.54463
9.12	83.1744	3.01993	9.54987
9.13	83.3569	3.02159	9.55510
9.14	83.5396	3.02324	9.56033
9.15	83.7225	3.02490	9.56556
9.16	83.9056	3.02655	9.57079
9.17	84.0889	3.02820	9.57601
9.18	84.2724	3.02985	9.58123
9.19	84.4561	3.03150	9.58645
9.20	**84.6400**	**3.03315**	**9.59166**
9.21	84.8241	3.03480	9.59687
9.22	85.0084	3.03645	9.60208
9.23	85.1929	3.03809	9.60729
9.24	85.3776	3.03974	9.61249
9.25	85 5625	3.04138	9.61769
9.26	85.7476	3.04302	9.62289
9.27	85.9329	3.04467	9.62808
9.28	86.1184	3.04631	9.63328
9.29	86.3041	3.04795	9.63846
N	N²	√N	√10N

CUADRADOS Y RAÍCES CUADRADAS (Continuación 13)

N	N²	√N	√10N
9.30	**86.4900**	**3.04959**	**9.64365**
9.31	86.6761	3.05123	9.64883
9.32	86.8624	3.05287	9.65401
9.33	87.0489	3.05450	9.65919
9.34	87.2356	3.05614	9.66437
9.35	87.4225	3.05778	9.66954
9.36	87.6096	3.05941	9.67471
9.37	87.7969	3.06105	9.67988
9.38	87.9844	3.06268	9.68504
9.39	88.1721	3.06431	9.69020
9.40	**88.3600**	**3.06594**	**9.69536**
9.41	88.5481	3.06757	9.70052
9.42	88.7364	3.06920	9.70567
9.43	88.9249	3.07083	9.71082
9.44	89.1136	3.07246	9.71597
9.45	89.3025	3.07409	9.72111
9.46	89.4916	3.07571	9.72625
9.47	89.6809	3.07734	9.73139
9.48	89.8704	3.07896	9.73653
9.49	90.0601	3.08058	9.74166
9.50	**90.2500**	**3.08221**	**9.74679**
9.51	90.4401	3.08383	9.75192
9.52	90.6304	3.08545	9.75705
9.53	90.8209	3.08707	9.76217
9.54	91.0116	3.08869	9.76729
9.55	91.2025	3.09031	9.77241
9.56	91.3936	3.09192	9.77753
9.57	91.5849	3.09354	9.78264
9.58	91.7764	3.09516	9.78775
9.59	91.9681	3.09677	9.79285
9.60	**92.1600**	**3.09839**	**9.79796**
9.61	92.3521	3.10000	9.80306
9.62	92.5444	3.10161	9.80816
9.63	92.7369	3.10322	9.81326
9.64	02.9296	3.10483	9.81835
9.65	93.1225	3.10644	9.82344
9.66	93.3156	3.10805	9.82853
9.67	93.5089	3.10966	9.83362
9.68	93.7024	3.11127	9.83870
9.69	93.8961	3.11288	9.84378
N	N²	√N	√10N

N	N²	√N	√10N
9.70	**94.0900**	**3.11448**	**9.84886**
9.71	94.2841	3.11609	9.85393
9.72	94.4784	3.11769	9.85901
9.73	94.6729	3.11929	9.86408
9.74	94.8676	3.12090	9.86914
9.75	95.0625	3.12250	9.87421
9.76	95.2576	3.12410	9.87927
9.77	95.4529	3.12570	9.88433
9.78	95.6484	3.12730	9.88939
9.79	95.8441	3.12890	9.89444
9.80	**96.0400**	**3.13050**	**9.89949**
9.81	96.2361	3.13209	9.90454
9.82	96.4324	3.13369	9.90959
9.83	96.6289	3.13528	9.91464
9.84	96.8256	3.13688	9.91968
9.85	97.0225	3.13847	9.92472
9.86	97.2196	3.14006	9.92975
9.87	97.4169	3.14166	9.93479
9.88	97.6144	3.14325	9.93982
9.89	97.8121	3.14484	9.94485
9.90	**98.0100**	**3.14643**	**9.94987**
9.91	98.2081	3.14802	9.95490
9.92	98.4064	3.14960	9.95992
9.93	98.6049	3.15119	9.96494
9.94	98.8036	3.15278	9.96995
9.95	99.0025	3.15436	9.97497
9.96	99.2016	3.15595	9.97998
9.97	99.4009	3.15753	9.98499
9.98	99.6004	3.15911	9.98999
9.99	99.8001	3.16070	9.99500
10.00	100.0000	3.16228	10.00000
N	N²	√N	√10N

TABLA N° III ÁREAS BAJO LA CURVA NORMAL TÍPICA DE O a Z

Z	00	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

TABLA Nº IV VALORES DE X^2

g. l.	PROBABILIDAD DE VALORES MAYORES DE X^2												
	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005
1	.04393	.03157	.03982	.02393	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	.207	.297	.484	.711	1.06	1.82	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.53	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.44	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0

TABLA No. IV VALORES DE X^2 (Continuación 2)

g. l.	PROBABILIDAD DE VALORES MAYORES DE X^2												
	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620 QUITO
OMNIUM PROGENITOR ET ALMA MATER

TABLA No. V. VALORES DE t

g. l.	SIGNO IGNORADO (DOS COLAS)								
	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.941
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.998	5.405
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
Signo considerado (una cola)									

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
OMNIUM POTENTIOR EST SCIENTIA

TABLA N° V. VALORES DE t (Continuación 2)

g. l.	SIGNO IGNORADO (DOS COLAS)								
	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001
21	.686	.859	1.063	1.323	1.721	2.080	2.518	2.931	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	.854	1.055	1,310	1.697	2.042	2.457	2.750	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	.677	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	.842	1.936	1.282	1.645	1.960	2.326	2.576	3.291
g. l.	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005.	0.0005
	Signo considerado (una cola)								

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDACIÓN 620 QUITO
OMNIUM POTENTIOR EST SCIENTIA

No. VI. VALORES DE F

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
	.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	.010	4052.0	4999.5	5403.0	5625.0	5764.0	5859.0	5928.0	5982.0	6022.0
	.005	16211.0	20000.0	21615.0	22500.0	23056.0	23437.0	23715.0	23952.0	24091.0
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	.005	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.025	17.44	15.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	.005	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14

TABLA No. VI. VALORES DE F (Continuación 2)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	.005	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
7	.100	3.59	3.26	3.07	2.26	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.025	8.07	6.54	5.89	5.52	5.29	5:12	4.99	4.90	4.82
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.57	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34

TABLA No. VI. VALORES DE F (Continuación 3)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
60.19	60.71	61.22	61.74	62.00	60.26	62.53	62.79	63.06	63.33
241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
968.6	976.7	984.9	993.1	997.2	1001.0	1006.0	1010.0	1014.0	1018.0
6056.0	6106.0	6157.0	6209.0	6235.0	6261.0	6287.0	6313.0	6339.0	6366.0
24224.0	24426.0	24630.0	24836.0	24940.0	25044.0	25148.0	25253.0	25359.0	25465.0
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
99.40	99.42	99.43	99.45	99.46	99w.47	99.47	99.48	99.49	99.50
199.4	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5
5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83
3.92	3.90	3.87	3.84	3.83.	3.82	3.80	3.79	3.78	3.76
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620 QUITO
OMNIUM POTENTIOR EST SAPIENTIA

TABLA No. VI. VALORES DE F (Continuación 4)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
3.30	3.27	3.24	3.21	2.19	3.17	3.16	3.14	3.12	3.10
4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14
2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88
2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SCIENTIA
FUNDADA EN 1826 QUITO

TABLA No VI. VALORES DE F (Continuación 5)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.12	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.005	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.20	5.20	5.06	4.94
	.005	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.005	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.005	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.005	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94

TABLA No. VI. VALORES DE F (Continuación 6)

g. l: del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.005	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.005	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.005	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	.005	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25

TABLA No. VI. VALORES DE F (Continuación 7)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19
2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23
2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65

TABLA No. VI. VALORES DE F (Continuación 8)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44
2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26
2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11
2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
2.92	2.82	2.72	2.62	3.56	2.50	2.44	2.38	2.32	2.25
3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SCIENTIA
FUNDADA EN 1826 QUITO

TABLA No. VI. VALORES DE F (Continuación 9)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.35	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.005	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.20	4.14
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.52	3.52	3.13	2.90	3.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.005	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.005	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96
21	.100	2.96	2.57	2.36	2.23	3.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.005	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81

TABLA No. VI. VALORES DE F (Continuación 10)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.005	9.63	6.73	5.58	4.95	4.54	3.26	4.05	3.88	3.75
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60

TABLA No. VI. VALORES DE F (Continuación 11)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87
1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78
1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
3.37	2.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69
1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
2.70	2.60	2.50	2.30	2.33	2.27	2.21	2.14	2.08	2.00
3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
r	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55

TABLA No. VI. VALORES DE F (Continuación 12)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.85	1.81	1.76
2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43
1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	7.71
2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38
1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33

TABLA No VI. VALORES DE F (Continuación 13)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.005	0.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.005	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620 QUITO
OMNIUM POTENTIOR EST SCIENTIA

TABLA N9 VI. VALORES DE F (Continuación 14)

g. l. del denominador	Nivel de significancia	g. l. del Numerador								
		1	2	3	4	5	6	7	8	9
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.26	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.39	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

TABLA No. VI. VALORES DE F (Continuación 15)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.29
1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25
1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21
1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.67	1.62
2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SPIENTIA
QUITO 1600

TABLA No. VI. VALORES DE F (Continuación 16)

g. l. del Numerador									
10	12	15	20	24	30	40	60	120	∞
1.76	7.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93
1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
1.99	1.92.	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69
1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43
1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00
1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00
2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00
2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00
2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00

TABLA No. VII. VALORES DE RANGO MÍNIMO DE DUNCAN

g.l. del Error	Nivel de significancia	Número de medias consideradas													
		2	3	4	5	6	7	8	9	10	12	14	16	18	20
1	.05	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
	.01	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	.05	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
	.01	14.0	14.0	14.0	14.0	14.0	14.0.	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	.05	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
	.01	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.0	9.1	9.2	9.3	9.3
4	.05	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
	.01	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.3	7.4	7.4	7.5	7.5
5	.05	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
	.01	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.6	6.6	6.7	6.7	6.8
6	.05	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
	.01	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.1	6.2	6.2	6.3	6.3
7	.05	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61
	.01	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	5.8	5.9	5.9	6.0	6.0
8	.05	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56
	.01	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.6	5.7	5.7	5.8	5.8
9	.05	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52
	.01	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.5	5.5	5.6	5.7	5.7
10	.05	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.48
	.01	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.36	5.42	5.48	5.54	5.55

TABLA No. VII VALORES DE RANGO MÍNIMO DE DUNCAN (Continuación 2)

g.l. del Error	Nivel de significancia	Número de medias consideradas													
		2	3	4	5	6	7	8	9	10	12	14	16	18	20
11	.05	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.46	3.46	3.46	3.47	3.48
	.01	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.24	5.28	5.34	5.38	5.39
12	.05	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.46	3.46	3.46	3.47	3.48
	.01	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.13	5.17	5.22	5.24	5.26
13	.05	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47
	.01	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.04	5.08	5.13	5.14	5.15
14	.05	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.45	3.46	3.46	3.47	3.47
	.01	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	4.96	5.00	5.04	5.06	5.07
15	.05	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.47	3.47
	.01	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	4.90	4.94	4.97	4.99	5.00
16	.05	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.44	3.45	3.46	3.47	3.47
	.01	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.84	4.88	4.91	4.93	4.94
17	.05	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.44	3.45	3.46	3.47	3.47
	.01	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.72	4.75	4.80	4.83	4.86	4.88	4.89
18	.05	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.43	3.45	3.46	3.47	3.47
	.01	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.76	4.79	4.82	4.84	4.85
19	.05	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.43	3.44	3.46	3.47	3.47
	.01	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.72	4.76	4.79	4.81	4.82
20	.05	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.43	3.44	3.46	3.46	3.47
	.01	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.69	4.73	4.76	4.78	4.79

TABLA No. VII. VALORES DE RANGO MÍNIMO DE DUNCAN (Continuación 3)

g.l. del Error	Nivel de significancia	Número de medias consideradas													
		2	3	4	5	6	7	8	9	10	12	14	16	18	20
22	.05	2.93	3.03	3.17	3.24	3.29	3.32	3.35	3.37	3.39	3.42	3.44	3.45	3.46	3.47
	.01	3.99	4.17	4.28	4.36	4.42	4.48	4.53	4.57	4.60	4.65	4.68	4.71	4.74	4.75
24	.05	2.92	3.07	3.15	3.22	3.28	3.31	3.34	3.37	3.38	3.41	3.44	3.45	3.46	3.47
	.01	3.96	4.14	4.24	4.33	4.39	4.44	4.49	4.53	4.57	4.62	4.64	4.67	4.70	4.72
26	.05	2.91	3.06	3.14	3.21	3.27	3.30	3.34	3.36	3.38	3.41	3.43	3.45	3.46	3.47
	.01	3.93	4.11	4.21	4.30	4.36	4.41	4.46	4.50	4.53	4.58	4.62	4.65	4.67	4.69
28	.05	2.90	3.04	3.13	3.20	3.26	3.30	3.33	3.35	3.37	3.40	3.43	3.45	3.46	3.47
	.01	3.91	4.08	4.18	4.28	4.34	4.39	4.43	4.47	4.51	4.56	4.60	4.62	4.65	4.67
30	.05	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.40	3.43	3.44	3.46	3.47
	.01	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.54	4.58	4.61	4.63	4.65
40	.05	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.39	3.42	3.44	3.46	3.47
	.01	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.46	4.51	4.54	4.57	4.59
60	.05	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.37	3.40	3.43	3.45	3.47
	.01	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.39	4.44	4.47	4.50	4.53
100	.05	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.36	3.40	3.42	3.45	3.47
	.01	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.35	4.38	4.42	4.45	4.48
∞	.05	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.34	3.38	3.41	3.44	3.47
	.01	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.26	4.31	4.34	4.38	4.41

TABLA No. VIII. VALORES DE DUNNETT, PARA COMPARACIONES DE UNA COLA

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
5	.05	2.02	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30
	.01	3.37	3.90	4.21	4.43	4.60	4.73	4.85	4.94	5.03
6	.05	1.94	2.34	2.56	2.71	2.83	2.92	3.00	3.07	3.12
	.01	3.14	3.61	3.88	4.07	4.21	4.33	4.43	4.51	4.59
7	.05	1.89	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.01
	.01	3.00	3.42	3.66	3.83	3.96	4.07	4.15	4.23	4.30
8	.05	1.86	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92
	.01	2.90	3.29	3.51	3.67	3.79	3.88	3.96	4.03	4.09
9	.05	1.83	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86
	.01	2.82	3.19	3.40	3.55	3.66	3.75	3.82	3.89	3.94
10	.05	1.81	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81
	.01	2.76	3.11	3.31	3.45	3.56	3.64	3.71	3.78	3.83
11	.05	1.80	2.13	2.31	2.44	2.53	2.60	2.67	2.72	2.77
	.01	2.72	3.06	3.25	3.38	3.48	3.56	3.63	3.69	3.74
12	.05	1.78	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74
	.01	2.68	3.01	3.19	3.32	3.42	3.50	3.56	3.62	3.67
13	.05	1.77	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71
	.01	2.65	2.97	3.15	3.27	3.37	3.44	3.51	3.56	3.61

TABLA No. VIII. VALORES DE DUNNETT, PARA COMPARACIONES DE UNA COLA (Continuación 2)

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
14	.05	1.76	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69
	.01	2.62	2.94	3.11	3.23	3.32	3.40	3.46	3.51	3.56
15	.05	1.75	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67
	.01	2.60	2.91	3.08	3.20	3.29	3.36	3.42	3.47	3.52
16	.05	1.75	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65
	.01	2.58	2.88	3.05	3.17	3.26	3.33	3.39	3.44	3.48
17	.05	1.74	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64
	.01	2.57	2.86	3.03	3.14	3.23	3.30	3.36	3.41	3.45
18	.05	1.73	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62
	.01	2.55	2.84	3.01	3.12	3.21	3.27	3.33	3.38	3.42
19	.05	1.73	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61
	.01	2.54	2.83	2.99	3.10	3.18	3.25	3.31	3.36	3.40
20	.05	1.72	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60
	.01	2.53	2.81	2.97	3.08	3.17	3.23	3.29	3.34	3.38
24	.05	1.71	2.01	2.17	2.28	2.36	2.43	2.48	2.53	2.57
	.01	2.49	2.77	2.92	3.03	3.11	3.17	3.22	3.27	3.31

TABLA No. VIII. VALORES DE DUNNETT, PARA COMPARACIONES DE UNA COLA (Continuación 3)

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
30	.05	1.70	1.99	2.15	2.25	2.33	2.40	2.45	2.50	2.54
	.01	2.46	2.72	2.87	2.97	3.05	3.11	3.16	3.21	3.24
40	.05	1.68	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51
	.01	2.42	2.68	2.82	2.92	2.99	3.05	3.10	3.14	3.18
60	.05	1.67	1.95	2.10	2.21	2.28	2.35	2.39	2.44	2.48
	.01	2.39	2.64	2.78	2.87	2.94	3.00	3.04	3.08	3.12
120	.05	1.66	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45
	.01	2.36	2.60	2.73	2.82	2.89	2.94	2.99	3.03	3.06
∞	.05	1.64	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42
	.01	2.33	2.56	2.68	2.77	2.84	2.89	2.93	2.97	3.00

TABLA No. VIII VALORES DE DUNNETT, PARA COMPARACIONES DE DOS COLAS

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
5	.05	2.57	3.03	3.39	3.66	3.88	4.06	4.22	4.36	4.49
	.01	4.03	4.63	5.09	5.44	5.73	5.97	6.18	6.36	6.53
6	.05	2.45	2.86	3.18	3.41	3.60	3.75	3.88	4.00	4.11
	.01	3.71	4.22	4.60	4.88	5.11	5.30	5.47	5.61	5.74
7	.05	2.36	2.75	3.04	3.24	3.41	3.54	3.66	3.76	3.86
	.01	3.50	3.95	4.28	4.52	4.71	4.87	5.01	5.13	5.24
8	.05	2.31	2.67	2.94	3.13	3.28	3.40	3.51	3.60	3.68
	.01	3.36	3.77	4.06	4.27	4.44	4.58	4.70	4.81	4.90
9	.05	2.26	2.61	2.86	3.04	3.18	3.29	3.39	3.48	3.55
	.01	3.25	3.63	3.90	4.09	4.24	3.47	4.48	4.57	4.65
10	.05	2.23	2.57	2.81	2.97	3.11	3.21	3.31	3.39	3.46
	.01	3.17	3.53	3.73	3.95	4.10	4.21	4.31	4.40	4.47
11	.05	2.20	2.53	2.76	2.92	3.05	3.15	3.24	3.31	3.38
	.01	3.11	3.45	3.68	3.85	3.98	4.09	4.18	4.26	4.33
12	.05	2.18	2.50	2.72	2.88	3.00	3.10	3.18	3.25	3.32
	.01	3.05	3.39	3.61	3.76	3.89	3.99	4.08	4.15	4.22

TABLA No. VIII. VALORES DE DUNNETT, PARA COMPARACIONES DE DOS COLAS (Continuación 2)

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
13	.05	2.16	2.48	2.69	2.84	2.96	3.06	3.14	3.21	3.27
	.01	3.01	3.33	3.54	3.69	3.81	3.91	3.99	4.06	4.13
14	.05	2.14	2.46	2.67	2.81	2.93	3.02	3.10	3.17	3.23
	.01	2.98	3.29	3.49	3.64	3.75	3.84	3.92	3.99	4.05
15	.05	2.13	2.44	2.64	2.79	2.90	2.99	3.07	3.13	3.19
	.01	2.95	3.25	3.45	3.59	3.70	3.79	3.86	3.93	3.99
16	.05	2.12	2.42	2.63	2.77	2.88	2.96	3.04	3.10	3.16
	.01	2.92	3.22	3.41	3.55	3.65	3.74	3.82	2.88	3.93
17	.05	2.11	2.41	2.61	2.75	2.85	2.94	3.01	3.08	3.13
	.01	2.90	3.19	3.38	3.51	3.62	3.70	3.77	3.83	3.89
18	.05	2.10	2.40	2.59	2.73	2.84	2.92	2.99	3.05	3.11
	.01	2.88	3.17	3.35	3.48	3.58	3.67	3.74	3.80	3.85
19	.05	2.09	2.39	2.58	2.72	2.82	2.90	2.97	3.04	3.09
	.01	2.86	3.15	3.33	3.46	3.55	3.64	3.70	3.76	3.81

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
OMNIUM SCIENTIARUM PONTIFEX ET SPECULUM

TABLA No. VIII VALORES DE DUNNETT, PARA COMPARACIONES DE DOS COLAS (Continuación 3)

g. l. del Error	Nivel de significancia	Número de medias menos el testigo								
		1	2	3	4	5	6	7	8	9
20	.05	2.09	2.38	2.57	2.70	2.81	2.89	2.96	3.02	3.07
	.01	2.85	3.13	3.31	3.43	3.53	3.61	3.67	3.73	3.78
24	.05	2.06	2.35	2.53	2.66	2.76	2.84	2.91	2.96	3.01
	.01	2.80	3.07	3.24	3.36	3.45	3.52	3.58	3.64	3.69
30	.05	2.04	2.32	2.50	2.62	2.72	2.79	2.86	2.91	2.96
	.01	2.75	3.01	3.17	3.28	3.37	3.44	3.50	3.55	3.59
40	.05	2.02	2.29	2.47	2.58	2.67	2.75	2.81	2.86	2.90
	.01	2.70	2.95	3.10	3.21	3.29	3.36	3.41	3.46	3.50
60	.05	2.00	2.27	2.43	2.55	2.63	2.70	2.76	2.81	2.85
	.01	2.66	2.90	3.04	3.14	3.22	3.28	3.33	3.38	3.42
120	.05	1.98	2.24	2.40	2.51	2.59	2.66	2.71	2.76	2.80
	.01	2.62	2.84	2.98	3.08	3.15	3.21	3.25	3.30	3.33
∞	.05	1.96	2.21	2.37	2.47	2.55	2.62	2.67	2.71	2.75
	.01	2.58	2.79	2.92	3.01	3.08	3.14	3.18	3.22	3.25

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTIORI EST SAPIENTIA
FUNDADA EN 1620

TABLA No. IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE

%	0	1	2	3	4	5	6	7	8	9
0.0	0	0.57	0.81	0.99	1.15–	1.28	1.40	1.52	1.62	1.72
0.1	1.81	1.90	1.99	2.07	2.14	2.22	2.29	2.36	2.43	2.50
0.2	2.56	2.63	2.69	2.75–	2.81	2.87	2.92	2.89	3.03	3.09
0.3	3.14	3.19	3.24	3.29	3.34	3.39	3.44	3.49	3.53	3.58
0.4	3.63	3.67	3.72	3.76	3.80	3.85–	3.89	3.93	3.97	4.01
0.5	4.05+	4.09	4.13	4.17	4.21	4.25+	4.29	4.33	4.37	4.40
0.6	4.44	4.48	4.52	4.55+	4.59	4.62	4.66	4.69	4.73	4.76
0.7	4.80	4.83	4.87	4.90	4.93	4.97	5.00	5.03	5.07	5.10
0.8	5.13	5.16	5.20	5.23	5.26	5.29	5.32	5.35+	5.38	5.41
0.9	5.44	5.47	5.50	5.53	5.56	5.59	5.62	5.65+	5.68	5.71
1	5.14	6.02	6.29	6.55–	6.80	7.04	7.27	7.49	7.71	7.92
2	8.13	8.33	8.53	8.72	8.91	9.10	9.28	9.46	9.63	9.81
3	9.98	10.14	10.31	10.47	10.63	10.78	10.94	11.09	11.24	11.39
4	11.54	11.68	11.83	11.97	12.11	12.25-	12.39	12.52	12.66	12.79
5	12.92	13.05+	13.18	13.31	13.44	13.56	13.69	13.81	13.94	14.06
6	14.18	14.30	14.12	14.54	14.65+	14.77	14.89	15.00	15.12	15.23
7	15.34	15.45+	15.56	15.68	15.79	15.89	16.00	16.11	16.22	16.32
8	16.43	16.54	16.64	16.74	16.85–	16.95+	17.05+	17.16	17.26	17.36
9	17.46	17.56	17.66	17.76	17.85+	17.95+	18.05–	18.15–	18.24	18.34
10	18.44	18.53	18.63	18.72	18.81	18.91	19.00	19.09	19.19	19.28
11	19.37	19.46	19.55+	19.64	19.73	19.82	19.91	20.00	20.09	20.18
12	20.27	20.36	20.44	20.53	20.62	20.70	20.79	20.88	20.96	21.05–
13	21.13	21.22	21.30	21.39	21.47	21.56	21.64	21.72	21.81	21.89
14	21.97	22.06	22.14	22.22	22.30	22.38	22.46	22.55–	22.63	22.71

TABLA No. IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE (Continuación 2)

%	0	1	2	3	4	5	6	7	8	9
15	22.79	22.87	22.95–	23.03	23.11	23.19	23.26	23.34	23.42	23.50
16	23.58	23.66	23.73	23.81	23.89	23.97	24.04	24.12	24.20	24.27
17	24.35+	24.43	24.50	24.58	24.65+	24.73	24.80	24.88	24.95+	25.03
18	25.10	25.18	25.25+	25.33	25.40	25.48	25.55–	25.62	25.70	25.77
19	25.84	25.92	25.99	26.06	26.13	26.21	26.28	26.35-	26.42	26.49
20	26.56	26.64	26.71	26.78	26.85+	26.92	26.99	27.06	27.13	27.20
21	27.28	27.35–	27.42	27.49	27.56	27.63	27.69	27.76	27.83	27.90
22	27.97	28.04	28.11	28.18	28.25-	28.32	28.38	28.45+	28.52	28.59
23	28.66	28.73	28.79	28.86	28.93	29.00	29.06	29.13	29.20	29.27
24	29.33	29.40	29.47	29.53	29.60	29.67	29.73	29.80	29.87	29.93
25	30.00	30.07	30.13	30.20	30.26	30.33	30.40	30.46	30.53	30.59
26	30.66	30.72	30.79	30.85+	30.92	30.98	31.05–	31.11	31.18	31.24
27	31.31	31.37	31.44	31.50	31.56	31.63	31.69	31.76	31.82	31.88
28	31.95–	32.01	32.08	32.14	32.20	32.27	32.33	32.39	32.46	32.52
29	32.58	32.65–	32.71	32.77	32.83	32.90	32.96	33.02	33.09	33.15–
30	33.21	33.27	33.34	33.40	33.46	33.52	33.58	33.65–	33.71	33.77
31	33.83	33.89	33.96	4.03	34.08	34.14	34.20	34.27	34.33	34.39
32	34.45–	34.51	34.57	34.63	34.70	34.76	34.82	34.88	34.94	35.00
33	35.06	35.12	35.18	35.24	35.30	35.36	35.43	35.49	35.55–	35.61
34	35.67	35.73	35.79	35.85-	35.91	35.97	36.03	36.09	36.15+	36.21

TABLA No. IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE (Continuación 3)

%	0	1	2	3	4	5	6	7	8	9
35	36.27	36.33	36.39	36.45+	36.51	36.57	36.63	36.69	36.75+	36.81
36	36.87	36.93	36.99	37.05–	37.11	37.17	37.23	37.29	37.35–	37.41
37	37.47	37.52	37.58	37.64	37.70	37.76	37.82	37.88	37.94	38.00
38	38.06	38.12	38.17	38.23	38.29	38.35+	38.41	38.47	38.53	38.59
39	38.65-	38.70	38.76	38.82	38.88	38.94	39.00	39.06	39.11	39.17
40	39.23	39.29	39.35-	39.41	39.47	39.52	39.58	39.64	39.70	39.76
41	39.82	39.87	39.93	39.99	40.05–	40.11	40.16	40.22	40.28	40.34
42	40.40	40.46	40.51	40.57	40.63	40.69	40.74	40.80	40.86	40.92
43	40.98	41.03	41.09	41.15–	41.21	41.27	41.32	41.38	41.44	41.50
44	41.55+	41.61	41.67	41.73	41.78	41.84	41.90	41.96	42.02	42.07
45	42.13	42.19	42.25–	42.30	42.36	42.42	42.48	42.53	42.59	42.65–
46	42.71	42.76	42.82	42.88	42.94	42.99	43.05–	43.11	43.17	43.22
47	43.28	43.34	43.39	43.45+	43.51	43.57	43.62	43.68	42.74	43.80
48	43.85+	43.91	43.97	44.03	44.08	44.14	44.20	44.25+	44.31	44.37
49	44.43	44.48	44.54	44.60	44.66	44.71	44.77	44.83	44.89	44.94
50	45.00	45.06	45.11	45.17	45.23	45.29	45.34	45.40	45.46	45.52
51	45.57	45.63	45.69	45.75–	45.80	45.86	45.92	45.97	46.03	46.09
52	46.15–	46.20	46.26	46.32	46.38	46.43	46.49	46.55–	46.61	46.66
53	46.72	46.78	46.83	46.89	46.95+	47.01	47.06	47.12	47.18	47.24
54	47.29	47.35+	47.41	47.47	47.52	47.58	47.64	47.70	47.75+	47.81
55	47.87	47.93	47.98	48.04	48.10	48.16	48.22	48.27	48.33	48.39
56	48.45–	48.50	48.56	48.62	48.68	48.73	48.79	48.85+	48.91	48.97
57	49.02	49.08	49.14	49.20	49.26	49.31	49.37	49.43	49.94	49.54
58	49.60	49.66	49.72	49.78	49.84	49.89	49.95+	50.01	50.07	50.13
59	50.18	50.24	50.30	50.36	50.42	50.48	50.53	50.59	50.65+	50.71

TABLA No. IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE (Continuación 4)

%	0	1	2	3	4	5	6	7	8	9
60	50.77	50.83	50.89	50.94	51.00	51.06	51.12	51.18	51.24	51.30
61	51.35+	51.41	51.47	51.53	51.59	51.65–	51.71	51.77	51.83	51.88
62	51.94	52.00	52.06	52.12	52 18	52.24	52.30	52.36	52.42	52.48
63	52.53	52.59	52.65+	52.71	52.77	52.83	52.89	52.95+	53.01	53.07
64	53.13	53.19	53.25–	53.31	53.37	53.43	53.49	53.55–	53.61	53.67
65	53.73	53.79	53.85–	53.91	53.97	54.03	54.09	54.15+	54.21	54.27
66	54.33	54.39	54.45+	54.51	54.57	54.63	54.70	54.76	54.82	54.88
67	54.94	55.00	55.06	55.12	55.18	55.24	55.30	55.37	55.43	55.49
68	55.55+	55.61	55.67	55.73	55.80	55.86	55.92	55.98	56.04	56.11
69	56.17	56.23	56.29	56.35	56.42	56.48	56.54	56.60	56.66	56.73
70	56.79	56.85+	56.91	56.98	57.04	57.10	57.17	57.23	57.29	57.35+
71	57.42	57.48	57.54	57.61	57.67	57.73	57.80	57.86	57.92	57.99
72	58.05+	58.12	58.18	58.24	58.31	58.37	58.44	58.44	58.56	58.63
73	58.69	58.76	58.82	58.89	58.95+	59.02	59.08	59.15–	59.21	59.28
74	59.34	59.41	59.47	59.54	59.60	59.67	59.74	59.80	59.87	59.93
75	60.00	60.07	60.13	60.20	60.27	60.33	60.40	60.47	60.53	60.60
76	60.67	60.73	60. SO	60.87	60.94	61.00	61.07	61.14	61.21	61.27
77	61.34	61.41	61.48	61.55–	61.62	61.68	61.75+	61.82	61.89	61.96
78	62.03	62.10	62.17	62.24	62.31	62.37	62.44	62.51	62.58	62.65+
79	62.72	62.80	62.87	62.94	63.01	63.08	63.15–	63.22	63.29	63.36

TABLA N° IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE (Continuación 5)

%	0	1	2	3	4	5	6	7	8	9
80	63.44	63.51	63.58	63.65+	63.72	63.79	63.87	63.94	64.01	64.08
81	64.16	64.23	64.30	64.38	64.45+	64.52	64.60	64.67	64.75–	64.82
82	64.90	64.97	65.05–	65.12	65.20	65.27	65.35–	65.42	65.50	65.57
83	65.65–	65.73	65.80	65.88	65.96	66.03	66.11	66.19	66.27	66.34
84	66.42	66.50	66.58	66.66	66.74	66.81	66.89	66.97	67.05+	67.13
85	67.21	67.29	67.37	67.45+	67.54	67.62	67.70	67.78	67.86	67.94
86	68.03	68.11	68.19	68.28	68.36	68.44	68.53	68.61	68.70	68.78
87	68.87	68.95–	69.04	69.12	69.21	69.30	69.38	69.47	69.56	69.64
88	69.73	69.82	69.91	70.00	70.09	70.18	70.27	70.36	70.45–	70.54
89	70.63	70.72	70.81	70.91	71.00	71.09	71.19	71.28	71.37	71.47
90	71.56	71.66	71.76	71.85+	71.95+	72.05	72.15–	72.24	72.34	72.44
91	72.54	72.64	72.74	72.84	72.95–	73.05–	73.15+	73.26	73.36	73.46
92	73.57	73.68	73.78	73.89	74.00	74.11	74.21	74.32	74.44	74.55–
93	74.66	74.77	74.88	75.00	75.11	75.23	75.35–	75.46	75.58	75.70
94	75.82	75.94	76.06	76.19	76.31	76.44	76.56	76.69	76.82	76.95–

TABLA No. IX. VALORES DE TRANSFORMACIÓN ARCO-SENO, PARA DATOS EN PORCENTAJE (Continuación 6)

%	0	1	2	3	4	5	6	7	8	9
95	77.08	77.21	77.34	77.48	77.61	77.75+	77.89	78.03	78.17	78.32
96	78.46	78.61	78.76	78.91	79.06	79.22	79.37	79.53	79.69	79.86
97	80.02	80.19	80.37	80.54	80.72	80.90	81.09	81.28	81.47	81.67
98	81.87	82.08	82.29	82.51	82.73	82.96	83.20	83.45+	83.71	83.98
99.0	84.26	84.29	84.32	84.35–	84.38	84.41	84.44	84.47	84.50	84.53
99.1	84.56	84.59	84.62	84.65–	84.68	84.71	84.74	84.77	84.80	84.84
99.2	87.87	84.90	84.93	84.97	85.00	85.03	85.07	85.10	85.13	85.17
99.3	85.20	85.24	85.27	85.31	85.34	85.38	85.41	85.45–	85.48	85.52
99.4	85.56	85.60	85.63	85.67	85.71	85.75–	85.79	85.83	85.87	85.91
99.5	85.95–	85.99	86.03	86.07	86.11	86.15–	86.20	86.24	86.28	86.33
99.6	86.37	86.42	86.47	86.51	86.56	86.61	86.66	86.71	86.76	86.81
99.7	86.86	86.91	86.97	87.02	87.08	87.13	87.19	87.25+	87.31	87.37
99.8	87.44	87.50	87.57	87.64	87.71	87.78	87.86	87.93	88.01	88.10
99.9	88.19	88.28	88.38	88.48	88.60	88.72	88.85+	89.01	89.19	89.43
100.0	90.0									

TABLA No. X. VALORES DE CORRELACIÓN

g.l. del error	Nivel de significancia	Variables Independientes							
		1	2	3	4	1	2	3	4
1	.05	.997	.999	.999	.999	20 .423	.509	.563	.604
	.01	1.000	1.000	1.000	1.000	.537	.608	.652	.685
2	.05	.950	.975	.983	.987	21 .413	.498	.522	.592
	.01	.990	.995	.997	.998	.526	.596	.641	.674
3	.05	.878	.930	.950	.961	22 .404	.488	.542	.582
	.01	.959	.976	.983	.987	.515	.585	.630	.663
4	.05	.811	.881	.912	.930	23 .396	.479	.532	.572
	.01	.917	.949	.962	.970	.505	.574	.619	.652
5	.05	.754	.836	.874	.898	24 .388	.470	.523	.562
	.01	.874	.917	.937	.949	.496	.565	.609	.642
6	.05	.707	.795	.839	.867	25 .381	.462	.514	.553
	.01	.834	.886	.911	.927	.487	.555	.600	.633
7	.05	.666	.758	.807	.838	26 .374	.454	.506	.545
	.01	.798	.855	.885	.904	.478	.546	.590	.624
8	.05	.632	.726	.777	.811	27 .367	.446	.498	.536
	.01	.765	.827	.860	.882	.470	.538	.582	.615
9	.05	.602	.697	.750	.786	28 .361	.439	.490	.529
	.01	.735	.800	.836	.861	.463	.530	.573	.606
10	.05	.576	.671	.726	.763	29 .355	.432	.482	.521
	.01	.708	.776	.814	.840	.456	.522	.565	.598

TABLA No. X. VALORES DE CORRELACIÓN (Continuación 2)

g. l. del error	Nivel de significancia	Variables independientes							
		1	2	3	4	1	2	3	4
11	.05	.553	.648	.703	.741	30 .349	.426	.476	.514
	.01	.684	.753	.793	.821	.449	.514	.558	.591
12	.05	.532	.627	.683	.722	35 .325	.397	.445	.482
	.01	.661	.732	.773	.802	.418	.481	.523	.556
13	.05	.514	.608	.664	.703	40 .304	.373	.419	.455
	.01	.641	.712	.755	.785	.393	.454	.494	.526
14	.05	.497	.590	.646	.686	45 .288	.353	.397	.432
	.01	.623	.694	.737	.768	.372	.430	.470	.501
15	.05	.482	.574	.630	.670	50 .273	.336	.379	.412
	.01	.606	.677	.721	.752	.354	.410	.449	.479
16	.05	.468	.559	.615	.655	60 .250	.308	.348	.380
	.01	.590	.662	.706	.738	.325	.377	.414	.442
17	.05	.456	.545	.601	.641	70 .232	.286	.324	.354
	.01	.575	.647	.691	.724	.302	.351	.386	.413
18	.05	.444	.532	.587	.628	80 .217	.269	.304	.332
	.01	.561	.633	.678	.710	.283	.330	.362	.389
19	.05	.433	.520	.575	.615	90 .205	.254	.288	.315
	.01	.549	.620	.665	.698	.267	.312	.343	.368

TABLA No. X. VALORES DE CORRELACIÓN (Continuación 3)

g. l. del error	Nivel de significancia	Variables Independientes			
		1	2	3	4
100	.05	.195	.241	.274	.300
	.01	.254	.297	.327	.351
125	.05	.174	.216	.246	.269
	.01	.228	.266	.294	.316
150	.05	.159	.198	.225	.247
	.01	.208	.244	.270	.290
200	.05	.138	.172	.196	.215
	.01	.181	.212	.234	.253
300	.05	.113	.141	.160	.176
	.01	.148	.174	.192	.208
400	.05	.098	.122	.139	.153
	.01	.128	.151	.167	.180
500	.05	.088	.109	.124	.137
	.01	.115	.135	.150	.162
1.000	.05	.062	.077	.088	.097
	.01	.081	.096	.106	.115

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

TABLA No. XI. COEFICIENTE DE POLINOMIOS ORTOGONALES

n:	n = 2	n = 3		n = 4			n = 5				n = 6				
ξ	1	1	2	1	2	3	1	2	3	4	1	2	3	4	5
Ci	−1	−1	+1	−3	+1	−1	−2	+2	−1	+1	−5	+5	−5	+1	−1
	+1	0	−2	−1	−1	+3	−1	1	+2	−4	−3	−1	+7	−3	+5
		+1	+1	+1	−1	−3	0	−2	0	+6	−1	−4	+4	+2	−10
				+3	+1	+1	+1	−1	2	−4	+1	−4	−4	+2	+10
							+2	+2	+1	+1	+3	−1	−7	−3	−5
											+5	+5	+5	+1	+1
Divisor:	2	2	6	20	4	20	10	14	10	70	70	84	180	28	252

n:	n = 7					n = 8					n = 9				
ξ	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Ci	−3	+5	−1	+3	−1	−7	+7	−7	+7	−7	−4	+28	−14	+14	−4
	−2	0	+1	−7	+4	−5	+1	+5	−13	+23	−3	+7	+7	−21	+11
	−1	−3	+1	+1	−5	−3	−3	+7	−3	−17	−2	−8	+13	−11	−4
	0	−4	0	+6	0	−1	−5	+3	+9	−15	−1	−17	+9	+9	−9
	+1	−3	−1	+1	+5	+1	−5	−3	+9	+15	0	−20	0	+18	0
	+2	0	−1	−7	−4	+3	−3	−7	−3	+17	+1	−17	−9	+9	+9
	+3	+5	+1	+3	+1	+5	+1	−5	−13	−23	+2	−8	−13	−11	+4
						+7	+7	+7	+7	+7	+3	+7	−7	−21	−11
											+4	+28	+14	+14	+4
Divisor.	28	84	6	154	84	168	168	264	616	2184	60	2772	990	2002	468

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR SAPIENTIA
FUNDADA EN QUITO 1620

TABLA No. XI. COEFICIENTE DE POLINOMIOS ORTOGONALES (Continuación 2)

n:	n = 10					n = 11					n = 12				
ξ	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Ci	−9	+6	−42	+18	−6	−5	+15	−30	+6	−3	−11	+55	−33	+33	−33
	−7	+2	+14	−22	+14	−4	+6	+6	−6	+6	−9	+25	+3	−27	+57
	−5	−1	+35	−17	−1	−3	−1	+22	−6	+1	−7	+1	+21	−33	+21
	−3	−3	+31	+3	−11	−2	−6	+23	−1	−4	−5	−17	+25	−13	−29
	−1	−4	+12	+18	−6	1	−9	+14	+4	−4	−3	−29	+19	+12	−44
	+1	−4	−12	+18	+6	0	10	0	+6	0	−1	−35	+7	+28	−20
	+3	−3	−31	+3	+11	+1	−9	−14	+4	+4	+1	35	−7	+28	+20
	+5	−1	−35	−17	+1	+2	−6	−23	−1	+4	+3	−29	−19	+12	+44
	+7	+2	−14	−22	−14	+3	−1	−22	−6	−1	+5	−17	−25	−13	+29
	+9	+6	+42	+18	+6	+4	+6	−6	−6	−6	+7	+1	−21	−33	−21
						+5	+15	+30	+6	+3	+9	+25	−3	−27	−57
											+11	+55	+33	+33	+33
Divisor	330	132	8580	2860	780	110	858	4290	286	156	572	12012	5148	8008	15912

TABLA No XII. Símbolos de Uso Común.

Símbolo	Significado
X, Y, Z	Variables
$X = Y$	X igual a Y
$X \neq Y$	X no igual a Y
$X \cong Y$	X aproximadamente igual a Y
$X < Y$	X menor que Y
$X \leq Y$	X igual o menor que Y
$X > Y$	X mayor que Y
$X \geq Y$	X igual o mayor que Y
$\lvert\ \rvert$	Valor absoluto; $\lvert -18 \rvert = 18$
$\sum$	Sigma mayúscula, la suma de
σ	Sigma minúscula, desviación típica de la población
σ^2	Sigma minúscula, al cuadrado, variancia de la población
s^2	Variancia de la muestra
s	Desviación típica de la muestra
$s^2_{\bar{X}}$	Variancia de las medias de las muestras
$s_{\bar{X}}$	Desviación típica de las medias de las muestras
$s_{\bar{d}}$	Desviación típica de la diferencia
μ	Miu, media población
$\overline{X}$	Equis barra, media de la muestra
a, b, c	Constantes
X_i, X_{ij}	Observación individual
$X_{i\cdot}, X_{\cdot i}, X_{\cdot j}, X_{j\cdot}$	Subtotal de un juego de observaciones X_{ij}
$X_{\cdot\cdot}$	Total general de observaciones
$\sum X^2_i$	La suma de cada X al cuadrado
$(\sum X_i)^2$	La suma total de las X al cuadrado
$\wedge$	Sombrero, indica un valor estimado
H_0	Hipótesis nula
H_i o H_1	Hipótesis alternativa
$\cdots$	Indica una serie de datos que, obviamente, están faltando
$N!$	N factorial
ε	Epsilon minúscula, error experimental
α	Alfa, probabilidad de cometer un error tipo I
β	Beta, probabilidad de cometer un error tipo II
r, ρ	Coeficientes de correlación(muestra y población)
b, β	Coeficientes de regresión (muestra y población)
δ	Delta, error del muestreo
J_i^2, X^2	J_i cuadrado
τ_i	*To*, tratamientos

L. C.	Límites de confianza
g. de l.	Grados de libertad
F. C.	Factor de correlación
S. C.	Suma de cuadrados
C. M.	Cuadrado medio
Ea, Eb	Errores a y b en parcelas divididas
N. S.	No significativo
*	Significativo
**	Altamente significativo
D. M. S.	Diferencia mínima significativa
C. V.	Coeficiente de variación
∞	Infinito

ALFABETO GRIEGO

Nombre	minúscula	mayúscula	
Alfa	α	A	a
Beta	β	B	b
Gamma	γ	Γ	entre g y j
Delta	δ	Δ	d
Epsilon	ε , ϵ	E	e corta
Zeta	ζ	Z	
Eta	η	H	e larga
Theta	θ , ϑ	Θ	
Iota	ι	I	i
Kappa	κ	K	k
Lambda	λ	Λ	
Mu	μ	M	m
Nu	ν	N	n
Xi	ξ	Ξ	ks
Omicron	o	O	
Pi	π , ϖ	Π	p
Rho	ρ , ϱ	P	r
Sigma	σ	Σ	
Tau	τ	T	t
Ypsilon	υ	Y	ü alemana
Phi	φ , ϕ	Φ	f
Ji	χ	X	
Psi	ψ	Ψ	ps
Omega	ω	Ω	o larga

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SCIENTIA
FUNDADA EN 1620
QUITO

TABLA No XIII. Referencias adicionales sobre Métodos Estadísticos y Diseño Experimental

AITKEN, A. C. 1933. On Fitting polynomials to weighted data by least squares. Proc. Roy. Soc. Edin. 54:1, and: On fitting polynomials to data with weighted and correlated errors. Ibid. 54:12.

_____.1933. On the graduation of data by the orthogonal polynomials of least. squares. Proc. Roy Soc. Edin. 53:54.

_____.1935. On least squares and linear combination of observations. Proc. Roy. Soc. Edin. 55:42.

_____and SIVERSTONE, H. 1942. On the estimation of statistical parameters. Proc. Roy. Soc. Edin. 61:186.

ALLAN, F. E. 1930. The general form of the orthogonal polynomials for simple series with proofs of their simple properties. Proc. Roy. Soc. Edin. 50:310.

_____and WISHART, J. 1930. A method of estimating the yield of a missing plot in field experimentation work. J. Agr. Sci. 20:399.

ALLEN, H. V. 1938. A theorem concerning the linearity of regression. Stat. Res. Men. 2:60.

ALLEN, R. G. D. 1939. The assumptions of linear regression. Económica, 6:191.

ANDERSON, O. 1937. On the logic of the decomposition of statistical series into separate components. J. R. S. S. 90:548.

ANDERSON, W. 1932. Researches into the theory of regression. Medd. Lunds Astr. Obs., Series 2, No 64.

_____.1934. On a new method of computing non-linear regression curves. Ann. Math. Stats. 5:81.

AROIAN, L. A. 1941. A study of R. A. Risher's z-distribution and the related R-distribution. Ann. Math. Stats. 12:429.

_____.1943. A new approximation to the leveis of significance of the chi square dis-tribution. Ann. Math. Stats. 14:93.

BACON, H. M. 1938. Note on a formula for the múltiple correlation coefficient. Ann. Math Stats. 9:227.

BA1LEY, A. L. 1931. The analysis of covariance. J. Am. Stat. Ass. 26:424.

BAKER, F. A. 1930. Random samples from non-homogeneous populations. Metron. 8(3);67.

BAKER, G. A. 1931. The relation between the means and variances, means, squared and variances in samples from combinations of normal populations. Ann. Math Stats. 2:333.

_____.1932. Distribution of the means divided by the standard deviations of samples from non-homogeneous populations. Ann. Math. Stats. 3:1.

_____.1934. Transformation of non-normal frequency-distributions into normal distributions. Ann. Math. Stats. 5:113.

_____.1936. The probability that the mean of a second sample will differ from the mean of a first sample by less than a certain múltiple of the standard deviation of the first sample. Ann. Math. Stats. 7:197.

_____.1938. The probability that the standard deviation of a second sample will differ from the standard deviation of a first sample by a certain múltiple of the first sample. Metron. 13(3):49.
_____.1941. Tests of homogeneity for normal populations. Ann. Math. Stats. 12:233.

BARBACKI, S., and FISHER, R. A. 1936. A test of the supposed precisión of arrange ments. Ann. Eug. Lond. 7:189.

BARLOWS TABLES. 1935. E. & F. M. Spon Ltd. 57 Hay Barket, S. W. 1 London 3rd Ed.

BARNARD, M. M. 1936. An enumeration of the confounded arrangements in the 2 x 2 x 2 factorial desing. Supp. J.R.S.S. 3:195.

BARTLETT, M. S. 1933. On the theory of statistical regression. Proc. Roy. Seo. Edin. 53:260.
_____.1933. Probability and chance ¡n the theory of statistics. Proc. Roy. Soc., A. 141:518.
_____.1934. The problem in statistics of testing several variances. Proc. Camb. Phil. Soc. 30:164.
_____.1935. An examination of the value of covariance in dairy cow nutrition experiments. Jour. Agr. Sci. 25:238-244.
_____.1935. The effect of non-normality on the t-distribution. Proc. Camb. Phil. Soc. 31:233.
_____.1936. Square root transformations in the analysis of variance. J. Roy. Stat. Soc. Suppl. 3:68-78.
_____and GREENHILL, A. W. 1936. The relative importance of plot variation and of field and laboratory sampling errors in small plot pasture and productivity experiments. Jour. Agr. Sci. 26:258-262.

BARTLETT, M. S. 1936. The information available in small samples. Proc. Camb. Phil. Soc. 32:560.
_____.1936. A note on the analysis of covariance. J. Agr. Sci. 26:488.
_____.1937. Properties of sufficiency and statistical tests. Proc. Roy. Soc. Londons Ser. A. 160:273.
_____.1937. Some examples of statistical methods of research in agriculture and applied biology. Jour. Roy. Stat. Soc. Supp. 4:137-183.
_____.1938. The approximate recovery of information from replicated experiments with large blocks. J. Agr. Sci. 28:418.
_____.1938. Further aspects of the theory of múltiple regression. Proc. Camb. Phil. Soc. 34:33.

BATEN, W. D. 1938. Elementary mathematical statistical statistics. John Wiley & Sons New York.
_____.1939. Formulas for finding estimates for two and three missing plots in randomized block layout. Mich. State Coll. Tech. Bul. 165.

BEALL, G. 1939. Methods of estimating the population of insects in a field. Biom 30:422.
_____.1942. The transformation of data from entomological field experiments so that analysis of variance becomes applicable. Biom. 32:243.

BERKSON, J. 1938. Some difficulties of interpretation encountered in the application of the chi-square test, J. Am. Stat. Ass. 33:526.

BERNSTEIN, F. 1937. Regression and correlations evaluated by a method of partial sums. Ann. Math. Stats. 8:77.

BILHAM, E. G. 1926. Correlation coefficients. Q. J. Roy. Met. Soc. 52:172.

BINGHAM. M. D. 1941. A new method for obtaining the inverse matrix. J. Am. Stat Ass. 36: 530.

BISHOP, D. J. 1939. On a comprehensive test for the homogeneity of variances and covariances in multivariate problems. Biom. 31:31.

BISHOP. D. J. and NAIR, U. S. 1939. A note on certain methods of testing for the homogenity of a set of estimated variances. Supp. J. R. S. S. 6:89.

BLISS, E. I. 1938. The transformation of percentages for use in the analysis of variance. Ohio J. Sci. 38:9.

BRANDT, A. E. 1937. Factorial design. Jour. Amer. Soc. Agron. 29:658-677.
_____.1938. Tests of significance in reversed or switchback trials, Iowa Exp. Sta. Res. Tul. 234.

BRYAN, A. A. 1933. Factores affecting experimental error in field plot tests with corn. La Agr. Coll. Res. Bul. 163.

BOSE, R. C. 1939. On the construction of balanced incomplete block desings. Ann. Eug. Lond. 9:353.
_____and NAIR, K. R. 1939. Partially balanced incomplete blocks. Sankhya. 4:337.
_____.1942. On complete sets of latin squares. Sankhya. 5:361.

BOSE, S. S., and MAHALANOBIS, P. C. 1938. On estimating individual yields in the case of mixed–up yields of more plots in field experiments. Sankhya. 4:103.

BOWLEY, A. L. 1912. The measurement of the accuracy of an avarage. J. R. S. S. 75:77.

BRADY, J. 1935. A biological application of the analysis of covariance. Supp J. R. S. S. 2:99.

BRODERICK, P. S. 1937. On some symbolic formulae in probability theory Proc. Roy. Irish Acad. A. 44:19.

BROWNLEE, J. 1910. The significance of the correlation coefficient when applied to Mendelian distributions. Proc. Roy. Soc. Edin. 30:473.
_____.1924. Experiments to test the theory of goodness of fit J. R. S. S. 87.76.

BURNSIDE, W. 1928. Theory of probability. Cambridge University Press.

CAMP, B. H. 1934. Methods of obtaining probability distributions. Ann. Math. Stats. 8:90.

CAMP, B. H. 1938. Further interpretations of the chi-square test. J. Am. Stat. Ass. 33:537.

CAMBELL, N. 1935. The statistical theory of errors. Proc. Phys. Soc. 47:800.

CARSLAW, H. S. 1930. Introduction to the theory of Fournier's series and integrals. MacMillan, London.

CHEMICAL RUBBER Co. 1936. Mathematical tables from Handbook of Chemistry and Physics. Chemical Rubber Publishing Co., Cleveland 5th Ed.

CHRISTIDIS, B. G. 1931. The importance of the shape of plot in field experimentatíon J. Agr. Sci. 21:14.

CLAPHRAN, W. G. 1931. Studies in sampling technique; cereal experiments, I field Technique, II Results & Discussions. Jour. Agr. Sci. 21:366-371 and 376-390.

CLARK, A., and LEONARD, W. H. 1939. The analysis of variance with special reference to data expressed as percentages. J. Am. Soc. Agron. 31:55.

COCHRAN, W. G. 1934. The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Proc. Camb. Phil. Soc. 30:718.
_____.1935. A note on the influence of rainfall on the yield of cereals in relatíon lo manurial treatments J. Agr. Sci. 25:510.
_____.1936. Statistical analysis of field counts of diseased plants. Jour Roy. Stat. Soc. Supp. 3:49-67.
_____.1937. Problems arising in the analysis of a series of similar experiments. Jour. Roy. Stat. Soc. Supp. 4:102-118.
_____. 1937. Catalogue of uniformity trial data 1. Jour. Roy. Stat. Soc. Supp. IV. 223.
_____.1937. Some difficulties in the statistical analysis of replicated experiments. Emp. Jour. Exper. Agrie. 6:157-175.
_____.1938. The omission or addition of an independent variate in múltiple linear regression. Jour Roy. Stat. Soc. Supp. 5:171-176.
_____.1938. Recent advance in mathematical statistics. Jour. Roy. Soc. C. 1. Part. 2:434.449.
_____.1938. The omission or addition of an independent variate in multiple linear regression. Supp. J. R. S. S. 5:171.

COCHRAN, W. G. 1938. Some difficulties in the statistical analysis of replicated experiments. Emp. J. Exp. 6:157.
_____.1939. Long term agricultural experiments. Jour Roy. Stat. Soc. Supp. 6:104-148.
_____.1939. The use of the analysis of variance in enumeration by sampling Jour. Amer. Stat. Soc. 34:492-510.
_____. 1939. Long-term agricultural experiments. Supp. J. R. S. S. 6:104.
_____.1940. The analysis of variance when experimental errors follow the Poisson or binomial laws. Ann. of Math. Statistics. 9:335-347.
_____.1940. A survey of experimental design. Grad. School, V. S. Dept of Agrie.
_____.1942. The X correction for continuity. lowa State College J. Sci. 61:421.
_____.1942. Sampling theory when the sampling units are of unequal sizes. J. Am. Stat. Ass. 37:199.

COLEMAN, J. B. 1932. A coeficient of linear correlation based on the method of the least squares and the line of best fit. Ann. Math. Stats. 3:79.

CORNISH, E. A. 1936. Non-replicated factorial experiments. J. Aus. Inst. Agr. Sci. 2:79.
_____.1940. The estimation of missing values in incomplete randomised block experiments. Ann. Eug. Lond. 10:112, The estimation of missing values in quasificatorial designs. Ibid., 10:137; The analysis of covariance in quasifactorial desings. Ibid, 10:269.

COX, G. M. and SNEDECOR, G. W. 1936. Covariance used to analyze the relation between corn yield and average. J. Farm. Econ. 18:597.

_____.1940. Enumeration and construction of balance incomplete block configurations. Ann. Math. Stats. 11:72.
_____, ECHHARDT, R. C., and COCHRAN, W. C. 1940. The analysis of lattice and triple lattice experiments in corn varietal lowa Exp. Sta. Res. Bul. 281.

CRAIG, A. T. 1932. The simultaneous distribution of mean and standard deviation in small samples. Ann. Math. Stats. 3:126.

CRAIG, C. C. 1931. Sampling in the case of correlated observations. Ann. Math. Stats. 2:324.

CRAIG, A. T. 1933. On the correlation between certain averges for small samples Ann. Math. Stats. 4:127.
_____.1938. On the independence of certain estimates of variance. Ann. Math. Stats. 9:48.

CRAMPTON, E. W. and HOPKINS, J. W. 1934. Of the method of partial regression in the analysis of comparative feeding trial data. Jour. Nutrition 8:113-123; 329-339.

CURTISS, J. H. 1943. On transformations used in the analysis of variance. Ann. Math. Stats. 14:107.

DAVENPORT, C. D. and EKANS, M. P. 1936. Statistical methods in biology, medicine and psychology. John Wiley & Sons New York.

DAVIES, O. L and PEARSON, E. S. 1934. Methods of estimating from samples, the population standard deviation. Supp. J. R. S. S. 1:76.

DAVIES, F. L. 1934. Study of the deviations of yields from duplícate pot cultures. Jour. Amer. Spc. Agrom. 26:831-838.

DAVIS, F. N. 1938. Tables of the correlation coefficient. Cambridge University Press.

DAVIS, H. T. 1933. Polynomial approximation by the method of least squares. Ann. Math. Stats. 4:155.

DEMING, W. E. 1937. On the significant figure of least squares and correlations, Science, 85:451.

DOOB, J. L. 1934. Probability and statistics. Trans, Am. Math. Soc. 36:759.

DWYER, P. S. 1937. The simultaneous computation of groups of regression equations and associated múltiple regression coeficientes. Ann. Math. Stats. 8:224.
_____.1941. The solution of simultaneous equations. Psychometrika. 6:101.
_____.1941. The Doolittie technique. Ann. Math. Stats. 12:449.
_____.1942. Recent developments in correlation technique. J. Am. Stat. Ass. 37:441.

EDEN, T. 1935. The development of field experiments in agricultural research. Trop. Agriculturist 84:63-69, 131-149, 188-195.

ENGELHART, M. S. 1936. The technique of path coefficients. Psychometrika. 1:287.

EISENHART, C. 1939. The interpretation of certain regression methods and their use in biological and industrial research. Ann. Math. Stat. 10:162-186.

EZEKIEL, M. 1930. Methods of correlation analysis. John Wiley & Sons New York.
_____.1926. The determination of curvilinear regression surfaces in the presence of other variables. J. Amer. Statistical. Assoc. 21:310-320.

EZEKIEL, M. 1930. Methods of correlation analysis. John Wiley and S. Sons, New York. (Chapman and Hall, London).
_____.1930. The sampling variability of linear and curvilinear regression. Ann. Math. 1:275.

FERTIG, J. W. 1936. On a Method of testing the hypotheses that an observed sample of size N has been drawn from a specified population of the same number of variables. Ann. Matha. Stats. 7:13.
_____and PROEHL, E. A. 1937. A test of a sample variance based on both tail ends of the distribution. Ann. Math. 8:193.

FISHER, R. A. 1921. On the mathematical foundations of theoretical statistics. Phil. Trans. Roy. Soc. A. 222:309.
_____.1921. On the probable error of a coefficient of correlation deduced from a small sample. Metron, 1 (4):1.
_____.1922. On the interpretation of X^2 from contingency tables and the calculation of P. J. R. S. S. 85:87.
_____.1922. The mathematical theory of probabilities and its applications to frequency -curves and statistical methods. 2nd end. Macmillan, New York.
_____.1922. The goodness of fit of regression formulae and the distribution of regression coefficients. J. R. S. S., 85. 597.
_____ and MACKENZIE, W. A. 1923. The manurial response of different potatoe varieties. Jour. Agr. Sci. 13:311.
_____.1924. On a distribution yielding the error, functions of several well known statistics. Proc. Int. Math. Cong. Toronto: 805-13.
_____.1924. The influence of rainfall on the yield of wheat at Rothamsted Phil. Trans. Roy. Soc. B. 213:89.
_____.1925, 1944. Statistical methods for research workers. (Ist edn. 1925, 2nd. edn. 1944). Oliver and Boyd, Edinburgh.
_____.1926. Applications of 'Studenfs' distribution. Metron, 5, N9 3, 90; and: Expansión of 'Studenfs' integral in powers of n-1. Metron, 5:(3):109.
_____.1926. The arrangement of field experiments. Jour. Min. of Agr. 33:503-513.
_____and WISHART, J. The arrangement of field experiments and the Statistical reduction of the results. Imp. Bur. Soil. Sci. Tech. Com. N9 10.
_____.1930. Inverse probability. Proc. Camb. Phil. Soc. 26:528.
_____.1933. The concepts of inverse probability and of fiducial probability refering to unknown parameters. Proc. Roy. Soc. A. 139:343.
_____.1935, 1942. The design of experiments 1 st edn. 1935, 3rd edn. (1942) Oliver and Boyd, Edinburgh.
_____and YATES, R. 1928, 1942. Statistical tables for use in Biological, Agricultural and Medical Research. 2nd. 1942. Oliver and Boyd, Edinburgh.
_____.1939. The comparison of samples with possibly unequal variances. Ann. Aug. Lond. 9:174.

_____.1940. An examination of the different possible solutions of a problem in complete blocks. Ann. Eug. Lond. 10:52.
_____.1946. Statistical methods for research workers, Oliver and Boyd, London, 5th ed.

FORESTER, H. C. 1937. Design of agronomic experiments for plots differentiated in fertility by past treatments. lowa Agr. Exp. Sta. Res. Bul. 226.

FRIDEMAN, M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. 32:675.

GARBER, R. J., McILCAINE, T. C. and HOOVAR, M. M. 1926. A study of soil heterogeneity in experimental plots. Jour. Agr. Res. 33:255-268.
_____and P1ERRE, W. H. 1933. Variations of yields obtamed in small artificially constructed field plots. Jour. Amer. Soc. Agron. 25:98-105.

GARMER, F. H., GRANTHAM, J. and SANDERS, H. G. 1934. The value of covariance in analyzing field experimental data. Jour. Agr. Sci. 24:250-259.

GOULDEN. C. H. 1937. Efficiency in field trials of pseudo-factorial and incomplete randomised block methods. Canadian J. Res. 15:231-241.
_____.1938. Modern methods for a large number of varieties. Dom. Canadá Dept. Agr. Tech. Bull. 9.
_____.1939. Methods of statistical analysis. John Wiley and Sons, New York (Chapman and Hall, London).

HALDANE, J. B. S. 1937. The exact valué of the moments of the distribution of X^2, used as a test of goodness of fit, when expectations are small. Biom. 29:133.

HALL, P. 1927. Múltiple and partial correlation coefficients in the case of an n-fold variate system. Biom. 19:100.

HARRIS, J. A. 1920. Practical universality of field heterogeneity as a factor influencing plot yields. Jour. Agr. Res. 19:279-314.

HAYES, H. K., WILSON, H. K. and ANSEMUS, E. R. 1932. An experiment study of the rodrow method with spring wheat. Jour. Amer. Soc. Agron. 24:950-960.

HOEL, P. G. 1937. A significance test for component analysis. Ann. Stats. 8:149.
_____.1941. On methods of solving normal equations. Ann Math. States. 12:354.

HOLLWELL, E. A. and HEUSINKVELD, D. 1933. Border effect studies of red clover and alfalfa. J. A. S. A. 25:779-788.

HOTELLING, H. 1931. The generalisation of 'Studenft ratio. Ann. Math. Stats. 2:360.
_____and PABST, M. R. 1936. Rank correlation and test of significance involving no assumption of normality. Ann. Math. Stats. 7:29.
_____.1943. Some methods in matrix calculation. Ann. Math. Stats. 14:1 and 440. HSU, P. L. 1940. On generalised analysis of variance Biom. 31:221.
_____.1943. Some simple facts abot the separation of degrees of freedom in factorial experiments. Sankya. 6:253.

HULBURT, H. W., MICHELS, C. A. and BURKHART, F. L. 1931. Border effect in variety tests of small grains. Agr. Exp. Sta. Univ. of Idaho. Res Bull. 9.

IMMER, F. R. 1932. A study of sampling technique with sugar beets Jour. Agr. Res 44. 633-648.

IMMER, F. R. and RELEIGH, S. M. 1932. Futher studies of size shape of plot in relation to field experiments with sugar beets. Jour. Agr. Res. 47:581-598, 649-668.
_____. HAYES, H. K. and POWERS, L. 1934. Statistic determination of barley varietal adaptation. Jour. Amer. Soc. Agron. 26:403-419.
_____. 1937. Correlation between means and standard deviations in field experiments. J. Am. Stat. Ass. 32:525.

IRWIN, J. 0. 1934. On the independence of the constituet items in the analysis of variance, Jour. Roy. Stat. Soc. Supp. 1:236-251.
_____.1939. Note on the X^2 test for goodness of fit, J. R. S. S. 92:274.
_____.1931. Mathematical theorems involved in the analysis of variance. J. R. S. S. 94:284.

ISSERLIS, L. 1936. Inverse probability. J. R. S. S. 99:130.

JACKSON, D. 1934. Series of orthogonal polynomials. Ann. Math. 34, 527. and: The Sumation of series of orthogonal polynomials. Bull. Am. Math. Soc. 40:743.
_____.1938. Orthogonal polynomials in three variables. Duke Math. J. 3:228.

JEFFREYS, H. 1938. The use of minimum X as an approximation to the method of maximum likelihood. Proc. Camb. Phil. Soc. 34:156.
_____.1938. Máximum likelihodd, inverse probability and the method of moments Ann. Eug. Lond. 8:146.
_____.1939. The theory of probability. Cambridge Univ. Press.

KALAMKAR, R. J. 1932. A study in sampling technique in wheat. Jour. Agr. Sci. 22: 783-796.

KELLEY. T. L. 1923. Statistical Methods. Macmillan, New York.

KEMP. W. B. 1938. Separating a generalized interaction into components Jour. Amer. Soc. Agron. 30:419-424.

KENDALL, M. G. 1938. A new measure of rank correlation. Biom. 30:81.
_____and BABINGTON-SMITH, B. 1939. Tables of random sampling numbers, tracts for computers, N? 24, Cambridge Univ. Press.

KENDALL, M. G. and YULE 1939. Introduction to the theory of statistics C. Griffin & Co., London.
_____and BABINGTON-SMITH, B. 1940. On the method of paired comparisons. Biom. 31:324.
_____.1924. Partial rank correlation. Biom. 32:277.
_____.1946. The advanced theory of statistics Vol. I and II. Charles Griffin & Co., Ltd. London.

KENNEY, J. F. 1939. Mathematics, of Statistics, Part II D. Van Nostrand Co., New York.

KIESSELBACH, T. A. 1923. Competition as a source of error in comparative corn yields Jour. Am. Soc. Agron. 15:199-215.

_____and WECHIN, R. AA. 1933. Effect of stand irregularities upon the acre yield and plant variability of corn. Jour. Agr. Res. 47:399-416.

KISHEN, K. 1940. On a simplified method of expressing the components of the second order interaction in a 3 x 3 factorial design. Sankhya. 4:577.
_____.1942. Symmetrical unequal block arrangements. Sankya. 5:329.

KLAGES, K. H. W. 1933. The rehability of nursery test as shown by correlated yields from nursery row and field plots. Jour. Amer. Soc. Agrom. 25:464-472.

KULLBACK, S. 1936. A note on the múltiple correlation coefficient. AAetron. 12(4):67.

LECLERG, E. L. 1935. Factors affecting experimental error in grenhouse pot tests with sugar beets. Phytopath, 25:1019-1025.

LEDERMANN, W. 1938. The orthogonal transformation of a factorial matrix into itself Psychometrika. 3:181.

LEGGAT, C. W. 1939. Statistical aspects of seed analysis. Bot. Rev. 5:505-529.

LEONARD, W. H. and CLARK, A. G. 1939. Field plot technique. Burgess Publishing. Co. Minneapolis.

LIDSTONE, G. T. 1939. Notes on orthogonal polynomials, etc. J. Inst. Act. 64:128.

LIVERMORE, J. R. 1940. Report of the comittee on standardization of field plot technique. Am. Pot. Jour. 17:114-123.

LOESELL, C. N. 1936. Size of plot and number of replications necessary for varietal triáis with white pea beans. Jour. Amer., Soc. Agron. 28:534-548.
_____.1936. Are uniformity trials useful? Jour. Amer., Soc. Agron. 28:2234-245.

LOVE, H. H. 1939. Application of statistical methods to agricultural research. Commercial Press, Lta., Shanghai.
_____.1943. Experimental methods in agricultural research. The Puerto Rico Agricultural Experiments Station.

MAHONEY, C. H. and BATEN, W. R. 1939. The use of the analyses of covariance and its limitation in the adjusment of yields based upon stand irregularities. Jour. Agr. Res. 58:317-328.

MAAN H. B. 1943. On the construction of sets of orthogonal Latin Squares. Ann. Math. Stats. 14:401.

MICHELS, C. A. and SCHWENDERMAN, J. 1934. Determining yields on experimental plots by the square yard method. Jour. Amer Soc. Agron. 26:993-1001.

MINER. J. 1922. Tables of 1-r2. John Hopkins Press, Balt.

MOSAK, J. L. 1937. On the simultaneous determination of the elementary regressions and their standard erros in subsets of unvariables. Jour. Amer. Stat. Soc. 32: 517-524.

NAGEL, E. 1936. The meaning of probability. J. Am. Stat. Ass. 31:10.

NAIR, K. R. and BANERJEE. K. S. 1943. A note on fitting straight lines if both variables are subject to error. Samkhya. 6:255.

NEYMAN, J. 1937. Lectures and conferences in mathematical statistics. Grad. School U. S. Dept. of Agriculture.
_____and PEARSON, E. S. 1938. Note on some points in 'Students' paper on 'Comparison between balanced and randon arrangemets of field plots'. Biom., 29, 380.

PANSE, V. G. 1939. Preliminary studies on sampling field experiments. Sankhya. 4:139.

PATERSON, D. D. 1939. Statistical technique in agricultural research. McGraw Hill, New York.

PEARL, R. 1930. Introduction to medical biometry and statistics. Saunders and Co. Philadelphia and London.

PEARSON, E. S. 1929. Some notes on sampling tests with two variables. Biom. 21:337.
_____.1937, 1938. Some aspects of the problem of randomisation. Biom., 29, 53, II. An illustration of 'Student's inquiry into the effect of balancing in agricultural experiments. Biom. 30:159.

PEARSON, K. Table for Statisticians and Biometricians. Cambridge Univ. Press. London Part. 1.
_____Tables for Statisticians and Biometricians, Part. II. Camb. Univ. Pross London.
_____.1920. Notes on the history of correlation-. Biom. 13:25.
_____.1924. Historical note on the origin of the normal curve. Biom. 16:402.
_____.1933. On a method of determining whether a sample of given size supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. Biom. 25:379.
_____.1934. On a new method of determining 'goodness of fit'. Biom. 26:425. RIDER, P. E. 1941. Statistical Methods. John Wiley & Sons, New York.

SALMÓN, S. C. 1941. The use of modern statistical methods in field experiments. Jour. Am. Soc. Agrom. 32:308-320.

SAUNDERS, A. R. 1939. Statistical methods with special reference to field experiments. Dept. of Agr. Union of S. Africa. Sci. Bul. 200.

SHEWART, W. A. 1939. Statistical methods from the viewpoint of quality control. Grad. School, U. S. Dept. Of. Agrie.

SCHOHAT, J. 1935. On the development of functions in series of orthogonal-polynomials. Bull. Am. Math. Soc. 4:41-49.

SMITH, H. FAIRFIELD. 1936. A discriminant function for plant selection. Ann. Eug. Lond. 7:240.

SMITH, H. F. 1936. Comparison of agricultural and nursery plots in variety experiments Journ. Of the Counc. for Scient. and Ind. Res. 9:207-210.

SNEDECOR, G. W., and IRWIN, M. R. 1933. On the chi-square test for homogeneity. lowa State College J. Sci. 8:75.
_____.1934. Analysis of variance. Collegiate Press Inc., Ames. lowa.
_____.1934. Calculation and interpretation of analysis of variance and covariance. Collegiate Press, Ames, lowa.
_____and COX, G. M. 1935. Disproportionate subclass numbers in tables of múltiple classification. La. Agr. Exp. Sta. Res. Bul. 80:235-272.
_____.1936. The improvement of statistical techniques in biolog. Jour. An. Stat. Soc. 31:690-701.
_____and COX, G. M. 1937. Analysis of covariance of yield and time of first silks in maize. Jour. Agr. Res. 54:449-57.
_____.1939. Desing of sampling experiments in the social sciences. Jour. of Farm Econ. 21;.846-855.

SPRAGUE, V. G. and MEYERS, W. AA. 1945. A comparative study of methods for determining yields of Kentucky bluegrass and wild white clover when grown in association J. Am. Soc. Agron. 37:370-376.

STADLER, L. J. 1921. Experiments in field plot technique for the preliminary determination of comparative yields in small grains. Univ. Of Mo. Coll. of Agr. Res. Bul. 49.

STEECE, H. W., IMMER, R. R., KIESSELBALCK, T. A. and McCLURE, J. T. 1935. Bibliography of field experiments. Jour. Amer. Soc. Agr. 27:1013-1018.

STERNE, T. E. 1934. The accuracy of least-square solutions. Proc. Nat. Acad. Sci., 20, 365 and 601.

STOCK, J. S. and FRANKEL, L. P. 1939. The allocation of sampling among several strata Ann. Math. Stats. 10:288.

STRINFIELD G. H. 1927. Intervarietal competition among small grains Jour. Amer. Soc. Agron. 19:971-983.

'STUDENT' 1908. On the probable error of a mean. Biom. 6:1.
_____.1931. On the z-test. Biom. 23:407.
_____.1938. Comparison between balanced and random arrangements of field plots. Biom. 29:361.

SUMMERBY, R. 1934. The value of preliminary uniformity trials in increasing precision of field experiments. Techn. Bul. MacDonald Agr. Coll. (15):64.

TANG, Y. 1938. Certain statistical problems arising in plant breeding. Biom. 30, 29.

TEDIN, O. 1931. The influence of systematic plot arrangements upon the estímate of error in fieid experiments. J. Agr. Sci. 21:191.

TINNEY, F. W. AAMODT, O. A. and ABLGREN, H. L. 1937. Preliminary report of a study on methods used in botanical analysis of pasture swards. Jour. Amer. Soc. Agron. 29:835-840.

TIPPETT, L. H. C. 1927. Random sampling numbers. Cambridge Univ. Press. London.
_____.1931. The methods of statistics. Williams and Norgate, London. 2nd edn. 1937.
_____.1937. The method of statistics. William & Norgate Ltd., London, Ed 2.

TODHUNTER, I. 1865. A history of the mathematical theory of probability from the time of Pascal to that of Laplace. Macmillan, London.

TRELOAR, A. E. 1935. Outline of biometric analysis. Burgess Publishing Co., Minneapolis.
_____.1939. Elements of statistical reasoning. John Willey & Sons, New York.

TYSDAL, H. M. and KIESSELBACH, T. A. 1939. Alfalfa nursery technic. Jour. Am. Soc. Agron. 31:83-98.

USPONSKY, J. V. 1937. Introduction to mathematical probability. McGraw Hill, New York and London.

VON MISES, R. 1941. On the foundations of probability and statistics Ann. Math. Stats. 12:191.

WALDLEIGH, C. H. and THORP, W. H. 1940. Factorial design in plant nutrition experiments in the greenhouse. Ark. Agr Exp. Sta. Bul. 401.

WALLACE, H. A. & SNEDECOR, G. W. 1931. Correlation and machine calculations. Collegiate Press, Ames. Iowa.

WALD, A. 1943. On the efficient desing of statistical investigations. Ann. Math. Stats. 14:134.

WIESS, M. G. and COX, G. M. 1939. Balanced incomplete block and lattice square desings for testings yield differences among large numbers of soya bean varieties. Iowa Agr. Exp. Sta. Res. Bull. 257-289.

WELCH, B. L. 1938. On test for homogeneity. Biom. 30:149.
_____.1938. The significance of the difference between two means, when the population variances are unequal. Biom. 29:350.

WIEBE, C. A. 1935. Variation and correlation in grain yield among 1500 wheat nursery plots. Jour. Agr. Res. 50:331-357.

WILKS, S. S. 1932. Certain generalisations in the analysis of variance. Biom. 24:471.

WINDSOR, C. P. and CLARKE, G. L. 1940. A statistical study of variation in the catch of planorton nets. Jour. Marine Res. 31:34.

WISHART, J. 1931. The mean and second-moment coefficient of the multiple correlation coefficient in samples from a normal population Biom. 22:353.
_____.1933. The theory of orthogonal polynomial fitting J. R. S. S. 96:487.
_____.1934. Bibliography of agricultural statistics. Supp. J. R. S. S. 1:95.
_____.1934. Statistics in agricultural research. Jour, Roy. Stat. Soc. (Suppl). 1:26-61.
_____and SANDERS, H. G. 1935. Principies and practice of field experimentation Empire Cotton Growing Crop. Kings Building, Crop. Kings Building, Dean Stanley St., London. S. W. 1.

_____.1936. Test of significance in analysis of covariance. Supp. J. R. S. S. 3:79.
_____.1938. Field experiments of factorial design. J. Agr. Sci. 28:299.
_____.1940. Field trials, their lay-out and statistical analysis. Imp. Bur. of Pl. Breed. and Gen. Cambridge.

WONG, Y. K. 1937. On the elimination of variables in multiple correlation J. Am. Stat. Ass. 32:357.

WOODBURY, M. A. 1940. Rank correlation when there are equal variates. Ann. Math. Stats. 11:358.

WRIGHT, S. 1934. The method of path coefficients. Ann. Math. Stats. 5:161.

YATES, F. 1933. The formation of latin squares for use in fielt experimentals. Emp. Agr. 1:235-244.
_____.1933. The analysis of replicated experiments when the field results are incomplete. Emp. Jour. Exp. Agr. 1:129-142.
_____.1933. The principies of orthogonality and confounding in replicated expements. J. Agr. Sci. 23:108.
_____.1933. The formation of Latin squares for use in field experiments. Emp. J. Exp. Agr. 1:235.
_____and WATSON, D. J. 1934. Observer's bias in sampling. Observation on wheat. Exp. Jour. Exp. Agr. 2:174-177.
_____.1934. The analysis of múltiple classifications with unequal numbers in the different classes. J. Am. Stat. Ass. 29:51.
_____.1934. Contingency tables involving small numbers and the X^2 test. Supp. J. R. S. S. 1:27.
_____.1935. Some examples of biased sampling. Ann. Eug. Lond. 6:202.
_____.1935. Complex experiments. Supp. J. R. S. S. 2:181.
_____and ZACOPANAY, I. 1935. The estimation of the efficiency of sampling, with special reference to sampling for yield in cereal experiments. J. Agr. Sci. 25:545.
_____.1936. Incomplete latin squares. J. Agr. Sci. 26:301.
_____.1936. Incomplete randomised block. Ann. Eug. Lond. 121.
_____.1936. Applications of the sampling technique to crop estimation and fore-casting. Trans Manchester Stat. Soc. 103.

YATES, F. 1936. A new method of arranging variety trials involving a large number of varieties. J. Agr. Sci. 26:424.
_____.1937. A further note on the arrangement of variety trials Quasi-Latin squares. Ann. Eug. Lond. 7:319.
_____.1937. The design and analysis of factorial experiments. Imp. Bur. Soil Sci. Tech. Comm. No 35.
_____and COCHRAN, W. G. 1938. The analysis of groups of experiments Jour. Agr. Sci. 27:556-580.
_____.1938. The gain in efficiency resulting from the use of balanced designs Supp. J. R. S. S. 5:70.
_____and COCHRAN, W. G. 1938. The analysis of groups of experiments J. Agr. Sci. 28:556.
_____.1938. Orthogonal functions and test of significance in the analysis of variance. Supp. J. R. S. S. 5:177.
_____.1939. The recovery of interblock information in variety trials arranged in the dimensional lattices. Ann. Eug. Lond. 9:136.
_____.1939. Test of significance of differences between regression coefficientes derived from two sets of correlated variates, Proc. Roy. Soc. Edin. 59:184.

_____.1939. The comparative of systematic and randomised arrangements in the designs. Ann. Eug. Lond. 10:317.
_____.1939. An apparent inconsistency arising from test of significance based on fiducial distributions of unknown parameters. Proc. Camb. Phil. Soc. 35:579.
_____and HALE, R. W. 1939. The analysis of latin squares when two or more rows, columns or treatments are missing. Supp. J. R. S. S. 6:67.
_____.1940. The recovery of inter-block information in balanced incomplete block desings. Ann. Eug. Lond. 10:317.
_____.1940. Modern experimental desgin and its function in plant selection. Emp. Jour. Exp. Agr. 8:223-230.

YOUDEN. W. J. and BEALE, H. P. 1934. A Statistical study of the local lesion method for estimating tobaceo mosaic. Contr, Boyce Thompson Inst. 6:437-454.
_____.1937. Use of incomplete block replications in estimation tobaceo mosaic virus. Contr. Boyde Thompson mst. 9:41-48.
_____.1937. Selection of efficient methods for soil sampling Contr. Boyce Thompson Inst. 9:59-70.
_____.1940. Experimental design to increase accuracay of geenhouse studies Boyce Tohmpson Inst. 11:219-228.

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO


UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST
OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

Índice Alfabético

Páginas

Análisis, combinatorio 84
de covariancia 149
de variancia 185
ordenación unilateral 185
con igual repetición 191
con repetición desigual 189
ordenación múltiple 205
con sub-muestras 200
Aptitud combinatoria 257
análisis estadístico 265
específica, cálculo de 267, 272
general, cálculo de 267, 272
modelo matemático 259
recíprocos 269
Atributos 21
Azar 30
Binomial, distribución 89
Bloques, completos al azar, diseño en 206
disposición de los 176
Coeficiente
binomial 95
de correlación 135
de determinación 136
de regresión 143
de variación 63
intelectual 33
Combinaciones 86
Comparaciones de grados de libertad individuales 233
entre medias de tratamientos 193
Confundido, experimentos en 242
Correlación 133
coeficiente de 135
propiedad del 136
Covariancia 149
ajuste de medias 158
ganancia en precisión por 159
ilustración 155
modelo matemático 153
suposiciones y modelo lineal 148
usos de 151
Cuadrado, latino, diseño de 217
modelo matemático 218
medio 56, 57
mínimos 143
Cuadros, presentación de 35, 37
de frecuencia 69

Cualitativos 21
caracteres 21
variables 29
Cuantitativos 21
caracteres 21
variables 29
Curva, de frecuencia, ... 72
normal 96
Datos, presentación de 34
Darwin, Ch. 24
Deducción 25
Desviación típica 59
de las medias 62
Dialélicos cruzamientos 22, 257
Diferencia mínima significativa 193
Dispersión (ver Medidas de 55
Dispersograma 134
Diseño, experimental 19
completamente al azar 188
con igual repetición 191
con repetición desigual 189
modelo lineal 199
bloques completos al azar 206
modelo lineal 210
cálculo de parcelas perdidas 211
grado de eficiencia vs. diseño completamente al azar 213
Cuadrado latino 217
Parcelas divididas 251
Distribución 32
binomial 89
de frecuencias 89
de Ji cuadrado 115
de Poisson 32, 222
de t de Student 117
de z 98
normal 22, 96
Duncan, prueba de rango múltiple de 195
Ecuación, de regresión 141, 143
Error, experimental 171, 177
control del 178
tipo I 110
tipo II 110
Estadística 21, 51
campos en los que se aplica la 21
definición 20
descriptiva 20
historia de la 23
la, y el método científico 24
Eventos (ver sucesos) 77

Éxito, probabilidad de 78
Expansión binomial 21, 94
Experimentación 19
clases de experimentos 167
disposición de los bloques 179
error experimental 177
control del 178
fases de la 166
inferencia estadística 183
introducción 163
objetivos 165
precisión relativa del experimento 180
randomización 182
repeticiones 173
factores que determinan el número de 177
selección del diseño experimental 182
selección del sitio y del material experimental 179
tratamientos 169
selección de los 170
unidad experimental 171
Experimentos, en confundido 242
factoriales 225
comparación de grados de libertad individuales 233
respuesta superficial: tratamientos igualmente espaciados 236
Factorial N! 84
Factoriales (ver Experimentos factoriales) 225
Fracaso, probabilidad de 78
Figuras, presentación de 34, 35
Fisher, R. A. 23, 24
Frecuencia, acumulada 70
de clase 70
relativa. 70
Fuentes de variación 186
en regresión 146
Funciones 51
Gates, C. E. 22
Gauss 96
Genética 21
Gosset, W. S. 24
Grados de libertad 56, 187
comparaciones de 233
Griffing, B. 22
Hechos 29
Hipótesis 105
H_0, hipótesis nula 107
pruebas de 105
Histogramas 71
Inferencia estadística 20, 183
Inducción, inductivo 25

UNIVERSIDAD CENTRAL DEL ECUADOR
FUNDADA EN 1620
QUITO

Interacción 227
Intervalo, de clase 73
de confianza 64
Investigación 19
Ji cuadrado, X^2 . . . . . 21, 115
Laplace 96
Línea recta, ecuación de la 141
Límite, de clase 73
de confianza 64
Marca, de clase 73
Material experimental 177
Media aritmética 53
Medidas 52
de dispersión 55
el rango 55
la variancia 56
la desviación típica 59
de tendencia central 52
la media 52
propiedades de la 54
la mediana 55
el modo 55
Mendel G. 21
Método, científico 24
estadístico 20
Modelo lineal aditivo 64
Muestras 30
al azar 30
representativa 30
Muestreo 39
Nivel de significación 108
Normal, curva 96
distribución 96
Observaciones 29
Parcelas divididas 251
parcelas perdidas 255
modelo lineal 256
Parámetros 51
Pascal, triángulo de 95
Pearson, K. 24
Permutación 85
Población 30
finita 30
infinita 30
Polígonos de frecuencia 71
Potencia de la prueba 110
Presentación de datos 33
Probabilidad 21
sucesos dependientes 81

sucesos independientes 79
sucesos mutuamente excluyentes 82
Prueba, de bondad de ajuste 116
de dos colas 110
de hipótesis 105
entre medias de tratamientos 193
diferencia mínima significativa 193
de rango múltiple de Duncan 195
de Scheffé 198
de Tukey 197
de una cola 110
Rango 55
Randomización 182
Redondeo de cifras 51
Referencias 280
Regresión 139
cálculo del coeficiente y de la ecuación de 143
ecuación de la línea recta 141
el método de los cuadrados mínimos 143
fuentes de variación en 146
propiedades y suposiciones en 148
valores ajustados 145
Repeticiones 173
factores que determinan el número de 177
Respuesta superficial 143
"diseño" de diamante doble 241
grados de libertad individuales 233
tratamientos igualmente espaciados 236
Scheffé, prueba de 198
Significación, prueba de 193
Simbología matemática 47
Spiegel, M. R. 20
Sprague, G. F. 22
Srb, A. M. 21
Steel, R. C. 20
Student, t da 24
Sucesos 77
dependientes 81
independientes 79
mutuamente excluyentes 82
Suma, de cuadrados 57
corregidos 187
no corregidos 57
de productos 156
t, distribución de 125
prueba de observaciones no pareadas 126
prueba de observaciones pareadas 128
valor de 125

Transformaciones 221
a raíz cuadrada 222
arco-seno 223
logarítmica 222
Tratamientos 169
selección de 170
Tukey, prueba de 197
Unidad experimental 171
Universo 25
Variables 29
cualitativa 29
cuantitativa 29
continua 29
discreta 29
Variación 27
ambiental 28
hereditaria 28
Variancia 56
Yates, corrección de 123
Zeta, prueba de 22, 98

UNIVERSIDAD CENTRAL DEL ECUADOR
POTENTIOR EST OMNIUM SAPIENTIA
FUNDADA EN 1620
QUITO

SIB - UCE

N.A.5118606

519.5
G643 González Bahamonde, Germán
Métodos estadísticos y principios de diseño experimental

FECHA | NOMBRE

N.A.5118606

519.5
G643 González Bahamonde, Germán
Métodos estadísticos y principios de diseño experimental

UNIVERSIDAD CENTRAL DEL ECUADOR
OMNIUM POTENTIOR EST SAPIENTIA
FUNDADA EN 1620
QUITO

ISBN 978-9978-92-883-7

DITORIAL
UNIVERSITARIA
editorial@ac.uce.edu.ec