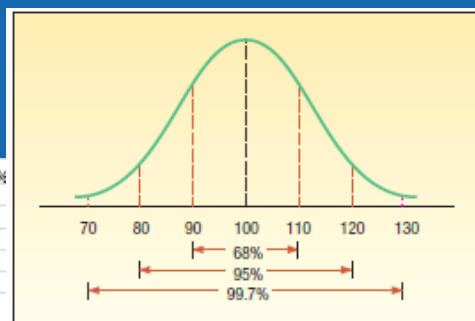
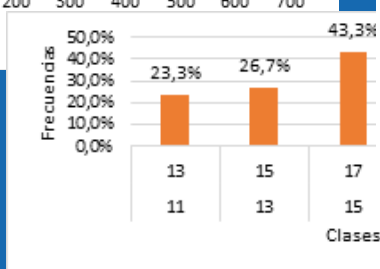
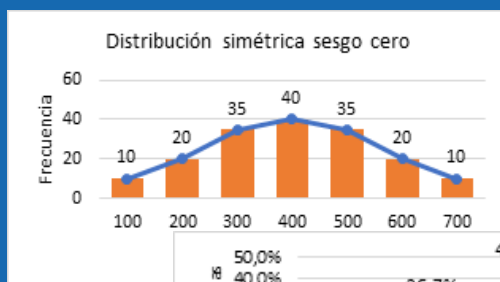


C.E.

ANÁLISIS ESTADÍSTICO BÁSICO PARA NEGOCIOS



Leonardo Salvador Pérez
Renato Revelo Oña
Ramiro Jurado Zurita

Jacqueline Ruiz Salas
Mónica Andrade Salcedo

SALVADOR, LEONARDO P.; REVELO, RENATO E.; JURADO, RICHARD R.; RUIZ, JACQUELINE Y.; ANDRADE, MÓNICA R.

Análisis estadístico básico para negocios

Primera edición

Quito, Ecuador, 2017-12-12

ISBN: 978-9942-30-471-1

Área: Administración

Formato: 14.85 x 21 cm

Páginas: 101

Colección Empresarial

De los autores:

Salvador Pérez Leonardo Patricio, ipsalvador@uce.edu.ec

Magister en Auditoría de Gestión de la Calidad, Magister en Gerencia de Proyectos Sociales y Educativos, Docente de la Universidad Central del Ecuador y Universidad Politécnica Salesiana.

Revelo Oña Renato Esteban, rrevelo@uce.edu.ec

Magister en Finanzas, Docente de la Universidad Central del Ecuador, Universidad Técnica Particular de Loja y Universidad Tecnológica Indoamérica.

Jurado Zurita Richard Ramiro, rrjurado@uce.edu.ec

Magister en Gestión de Proyectos Socio Productivos, Docente de la Universidad Central del Ecuador y Universidad Politécnica Salesiana.

Ruiz Salas Jacqueline Yovana, jyruiz@uce.edu.ec

Magister en Gestión de Proyectos Socio Productivos, Docente de la Universidad Central del Ecuador.

Andrade Salcedo Mónica del Rocío, mrandrade@uce.edu.ec

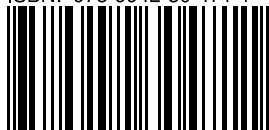
Magister en Administración y Marketing, Magister en Auditoría Integral, Docente de la Universidad Central del Ecuador y Universidad Técnica Particular de Loja.

Reservados todos los derechos. Ni la totalidad ni parte de esta publicación puede reproducirse, registrarse o transmitirse, por un sistema de recuperación de información, en ninguna forma ni por ningún medio, sea electrónico, mecánico, fotoquímico, magnético o electroóptico, por fotocopia, grabación o cualquier otro, sin permiso previo por escrito del editor.

El préstamo, alquiler o cualquier otra forma de cesión de uso de este ejemplar requerirá también la autorización del autor o de sus representantes.

Conforme lo establece el Art. 71 y 72 del Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior (Codificación), esta obra fue revisada por dos pares académicos externos y cuenta con el aval emitido por el Decano de la Facultad de Ciencias Administrativas, UCE, mediante oficio No. DFA-920 de fecha 12 de diciembre de 2017.

ISBN: 978-9942-30-471-1



9789942304711

ANÁLISIS ESTADÍSTICO BÁSICO PARA NEGOCIOS



Leonardo Patricio
Salvador Pérez



Renato Esteban
Revelo Oña



Richard Ramiro
Jurado Zurita



Jacqueline Yovana
Ruiz Salas



Mónica del Rocío
Andrade Salcedo

Primera Edición

Dedicatoria

A nuestros familiares

Contenidos

1. Introducción	8
1.1 ¿Por qué se debe estudiar la estadística?.....	8
1.2 Definición de estadística	8
1.3 Ética y estadística	9
1.4 Tipos de estadística	9
1.5 Variables	10
2. Niveles de medición.....	14
2.1 Datos de nivel nominal.....	14
2.2 Datos de nivel ordinal.....	14
2.3 Datos de nivel intervalo.....	14
2.4 Datos de nivel de razón.....	15
3. Tabla de frecuencias.....	18
3.1 Tabla de distribución de frecuencias.....	18
4. Medidas de tendencia central	30
4.1 Medidas de ubicación	30
5. Medidas de dispersión.....	46
5.1 Rango.....	46
5.2 Desviación media.....	46
5.3 Varianza y desviación estándar	48
6. Relación entre variables.....	70
6.1 Análisis de correlación	70
6.2 Coeficiente de correlación	71
6.3 Regresión lineal.....	74
7. Números Índice	78
7.1 Números índice simples	78
7.2 Índices no ponderados	80
7.3 Índices ponderados	87
7.4 Índice de valores.....	92
7.5 Índices para propósitos especiales	92

Capítulo

1

Generalidades

1. Introducción

Las técnicas estadísticas se aplican de manera amplia en mercadotecnia, contabilidad, control de calidad y en otras actividades; estudios de consumidores; análisis de resultados en deportes; administradores de instituciones; en la educación; organismos políticos; médicos; y por otras personas que intervienen en la toma de decisiones.

En los negocios cotidianos se utiliza la estadística: el promedio de compras del ama de casa en el mercado, el volumen de ventas del almacén TIA en las épocas de navidad, el volumen de compras de mercadería de la Corporación Favorita para el mes de febrero, etc., y así un sin número de actividades en donde se necesite trabajar con un registro de datos.

1.1 ¿Por qué se debe estudiar la estadística?

Por qué permite realizar un análisis comparativo e interpretativo de datos que servirán para la toma de decisiones y de esta manera determinar los cambios o variaciones que se dan en cualquier ámbito de estudio.

Por medio de la recopilación de datos obtenidos en la muestra, podremos interpretar de mejor manera la información para solucionar un problema.

1.2 Definición de estadística

Cuando se habla de estadística, las personas rápidamente se imaginan estar ante una colección de datos, tales como la estatura y peso de personas, el promedio de calificaciones de los estudiantes de Administración, número de partidos ganados por la LDU, cantidad de accidentes de autos, etc.

Spiegel & Stephens (2009) definen que “La estadística se ocupa de los métodos científicos que se utilizan para recolectar, organizar, resumir, presentar y analizar datos así como para obtener conclusiones válidas y tomar decisiones razonables con base en este análisis” (pág. 1).

Coincidiendo con los autores antes mencionados, Lind, Marchal, & Whaten (2012) definen que la estadística es la “Ciencia

que recoge, organiza, presenta, analiza e interpreta datos con el fin de propiciar una toma de decisiones más eficaz” (pág. 5).

La información generalmente se presenta en gráficos donde se visualiza y deduce rápidamente los resultados de los análisis.

1.3 Ética y estadística

Dado que la estadística trabaja con datos, es muy importante que las personas involucradas en todas las fases donde se recoge, organiza, presenta, analiza, interpreta resultados, y finalmente se informa a la sociedad, en cualquier ámbito, sean íntegros y honestos, de conducta y moral intachable para no sesgar los resultados, esto quiere decir que debe imperar la ética en el proceder de la aplicación de la estadística.

1.4 Tipos de estadística

La estadística se divide en dos categorías: descriptiva e inferencial.

1.4.1 Estadística descriptiva

Se ocupa de describir y analizar la información de una población.

Una definición de estadística descriptiva señala que son los “(...) Métodos para organizar, resumir y presentar datos de manera informativa” (Lind , Marchal, & Whaten, 2012, pág. 6).

Ejemplos: a) El volumen de ventas de la empresa Movistar durante el año 2017; b) El número de estudiantes de estadística en la Carrera de Administración de Empresas.

1.4.2 Estadística inferencial

Técnica que analiza determinadas características de una muestra para posteriormente deducir sobre la población investigada.

Lind, Marchal, & Whaten (2012) definen a la estadística inferencial como los “(...) Métodos que se emplean para determinar una propiedad de una población con base en la información de una muestra de ella” (pág. 7)

Ejemplo: a) Una persona tiene mareos y el médico en base a una muestra de sangre puede determinar si tiene el colesterol elevado; b) Para probar la resistencia de los vidrios de seguridad que una fábrica está produciendo, se somete a prueba hasta la rotura de una muestra de vidrios y de allí se deduce cual es la resistencia de toda la producción.

Población.

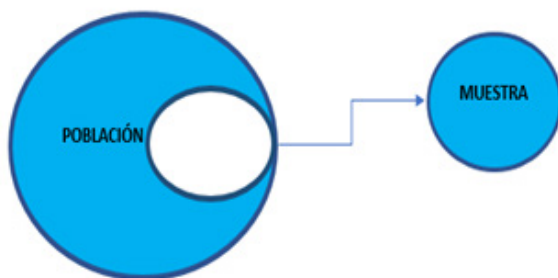
Es el conjunto universo de personas, cosas u objetos que se van a estudiar en su totalidad.

Ejemplo: La población de estudiantes de la Universidad Central del Ecuador.

Muestra.

Parte de la población que se quiere estudiar.

Ejemplo: Los estudiantes de la Facultad de Ciencias Administrativas.



1.5 Variables

La variable es una característica de una población o muestra que se está observando (Webster, 2000, pág. 9).

Existen dos tipos de variables: cualitativas y cuantitativas.

1.5.1 Variable cualitativa o atributo

Su característica es que no se expresan numéricamente. Detallan cualidades de un objeto, persona, problema.

1.5.2 Variable cuantitativa

Se expresan en forma numérica y se dividen en discretas y continuas.

Variable discreta.

Se expresan en números enteros como la edad, el número de unidades vendidas, etc.

Variable continua.

Se expresan en números decimales y son el resultado de mediciones. Se expresan con números decimales o con fracciones.

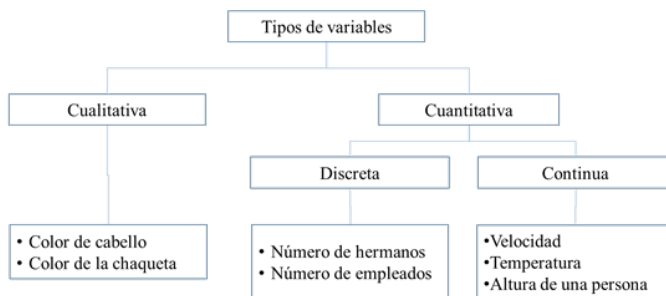


Figura . Resumen de los tipos de variables. Modelo tomado de (Lind , Marchal, & Whaten, 2012, pág. 9).

Capítulo

2

Niveles de medición

2. Niveles de medición

El resumen y presentación de los datos estadísticos se expresan a través de cuatro niveles de medición:

2.1 Datos de nivel nominal

Los resultados solo se clasifican y se cuentan, no necesitan un orden al momento de expresar los resultados. Al aplicar la variable de interés se divide en categorías o resultados. (Lind , Marchal, & Whaten, 2012, pág. 10).

2.2 Datos de nivel ordinal

Los resultados se clasifican y ordenan. Así, una calificación es más alta o menor: excelente, muy bueno, bueno, regular.

Propiedades:

1. Las clasificaciones de los datos se encuentran representadas por conjuntos de etiquetas o nombres (alto, medio, bajo), las cuales tienen valores relativos.
2. En consecuencia, los valores relativos de los datos se pueden clasificar u ordenar (Lind , Marchal, & Whaten, 2012, pág. 11).

2.3 Datos de nivel intervalo

Los resultados se ordenan de acuerdo con el grado de la característica analizada. Incluye todas características del nivel ordinal, pero la diferencia entre valores constituye una magnitud constante.

Las variables de intervalo dan un sentido de “cuanto” o de “que tamaño”, “que tan caliente”, “que tan obstinado”, “que tan deprimido”, etc. Con las variables de intervalo se piensan en términos de distancia entre puntuaciones sobre una línea recta.

Las propiedades son:

1. Las clasificaciones de datos se ordenan de acuerdo con el grado que posea de la característica en cuestión.
2. Diferencias iguales en la característica representan di-

ferencias iguales en las mediciones (Lind , Marchal, & Whaten, 2012, pág. 12).

2.4 Datos de nivel de razón

Los resultados se expresan comparando valores. Expresan cuanto más grande o más pequeño es un valor en relación con otro.

Posee todas las características del nivel de intervalo, pero considera que el punto 0 tiene sentido y la razón entre dos números es significativa. Las propiedades son:

1. La clasificación de datos se ordena de acuerdo con la cantidad de características que poseen.
2. Diferencias iguales en la característica representan diferencias iguales en los números asignados a las clasificaciones.
3. El punto cero representa la ausencia de características y la razón entre dos números es significativa (Lind , Marchal, & Whaten, 2012, pág. 13).



Figura : Resumen de niveles de medición. Modelo tomado de (Lind , Marchal, & Whaten, 2012, pág. 13).

Capítulo

3

Tablas de frecuencias

3. Tabla de frecuencias

Casi todos los trabajos que se realizan en la estadística inician con la recolección de datos.

La sola recolección de datos registrados en el papel poco o nada revelan por si solos, por lo que es indispensable que sean organizados de manera concisa para determinar su significancia y se pueda ver a simple vista que es lo que nos dicen.

Tabla de frecuencias es la “Agrupación de datos cualitativos en clases mutuamente excluyentes que muestra el número de observaciones en cada clase” (Lind , Marchal, & Whaten, 2012, pág. 23).

Entre las herramientas estadísticas que se utilizan para organizar los datos, se tienen:

- La tabla de distribución de frecuencias, donde los datos se encuentran en clases específicas.
- Los diferentes tipos de gráficos que expresan visualmente los resultados.
- Las tablas de contingencia y diagramas de “tallo y hojas”.

3.1 Tabla de distribución de frecuencias

En la tabla de frecuencias los datos de la variable considerada se agrupan en clases, de acuerdo con las necesidades de la investigación.

Las clases definidas tienen una distribución de frecuencias, lo que se denomina datos agrupados.

De acuerdo con los dos tipos de variables, se construye la tabla de distribución de frecuencias para datos cuantitativos y cualitativos.

3.1.1 Construcción de una tabla de frecuencias de datos cuantitativos

Lind, Marchal & Wathen (2012) señalan que la distribución de

frecuencias es la “Agrupación de datos en clases mutuamente excluyentes, que muestra el número de observaciones que hay en cada clase” (pág. 29)

Para desarrollar una tabla de frecuencias, se deben considerar los siguientes pasos:

Paso 1: Determinar el valor máximo y mínimo del conjunto de datos recolectados.

Paso 2: Definir el rango (valor máximo menos el valor mínimo).

Paso 3: Determinar el número de clases o categorías presentes en el conjunto de datos.

Se utiliza la fórmula:

El 2 es un valor constante que no será cambiado por ninguna circunstancia.

k = número de clases. Este valor será siempre número par iniciando con el 2

n = número de datos

Ejemplos:

Determine cuántas clases o categorías serán necesarias para la distribución de frecuencias, en los siguientes casos:

a) $n = 45$

$n = 45$		
Constante	k	2^k
2^1	2	4
2^2	3	8
2^3	4	16
2^4	5	32
2^5	6	64

Como 2 elevado a la 6 es 64 y es mayor a 45, entonces $k = 6$

2 ^

b)

7

n = 250

128

n = 250			
Constante	k	2^k	
2 ^	2	4	
2 ^	3	8	
2 ^	4	16	
2 ^	5	32	
2 ^	6	64	
2 ^	7	128	
2 ^	8	256	Como 2 elevado a la 8 es 256 y es mayor a 250, entonces k = 8

Paso 4: Se determina el intervalo que es el espacio que habrá entre dos valores del conjunto de datos.

Si el conjunto de datos contiene solo números enteros, el intervalo siempre será el número entero siguiente a la fracción resultante, sin importar si los decimales sean menores a 0,5.

Si el conjunto de datos contiene valores con decimales, el intervalo deberá ser aproximado a la décima o centésima más alta.

Para obtener el intervalo se divide el rango para el número de clases k.

Ejemplos:

En un conjunto de datos se tienen de 230 observaciones entre \$135 y \$321. Determinar el intervalo conveniente.

N° datos =	230
V Max =	321
V min =	135
Rango =	186

N° clases =	8
Intervalo =	23,25
Intervalo =	24,00

En un conjunto de datos se tienen de 115 observaciones entre \$ 16,35 y \$ 45,24. Determinar el intervalo conveniente.

N° datos =	115
V Max =	45,24
V min =	16,35
Rango =	28,89
N° clases =	7
Intervalo =	4,13
Intervalo =	4,15

Ejercicio:

Las siguientes son las calificaciones de 30 estudiantes de estadística en un curso de la Carrera de Administración de Empresas. Organice los datos en una distribución de frecuencias.

13	16	15	16	14
17	17	16	12	14
13	13	16	14	17
14	17	13	12	17
16	19	14	16	15
17	11	16	20	15

Paso 5: Construir la tabla de frecuencias

N° datos =	30	N° clases	Límite inferior	Límite superior	Punto medio	Frecuencia absoluta	Frecuencia absoluta acumulada	Frecuencia absoluta relativa %	Frecuencia acumulada relativa %
V Max	20	1	11	13	12,00	7	7	23,3%	23,3%
V min	11	2	13	15	14,00	8	15	26,7%	50,0%
Rango	9	3	15	17	16,00	13	28	43,3%	93,3%
N° clases	5	4	17	19	18,00	1	29	3,3%	96,7%
Intervalo	1,80	5	19	21	20,00	1	30	3,3%	100,0%
	2					30		100,0%	

Procedimiento:

- El primer valor del límite inferior es el valor mínimo o el que el profesional estadístico elija utilizar, pero siempre será menor al valor mínimo. En este caso se utilizó el valor mínimo.
- El límite superior es el límite inferior más el intervalo seleccionado. En este caso es 2.
- El siguiente límite inferior es el límite superior anterior más el intervalo, y así sucesivamente.
- El Punto medio es el promedio o media aritmética entre los límites inferior y superior en cada clase.
- La frecuencia absoluta es el conteo de datos que se encuentran entre los límites. Si un valor es igual al límite superior se coloca en la primera clase y se excluye de la siguiente clase (el valor será igual al límite inferior y no se puede repetir). Al trabajar de esta manera se puede comprobar usando el Excel.
- La frecuencia absoluta acumulada es la suma sucesiva de las frecuencias parciales de cada clase. (Clase 1: 7; clase 2: 7+8 = 15; etc.)
- La frecuencia absoluta relativa es el porcentaje que representa cada una de las frecuencias absolutas dividido para el total de datos. (Clase 1: $7/30 \cdot 100 = 23,3\%$; clase 2: $8/30 \cdot 100 = 26,7\%$; clase 3: $13/30 \cdot 100 = 43,3\%$; etc.)
- La frecuencia acumulada relativa es la suma sucesiva de las frecuencias absolutas relativas parciales de cada clase. (Clase 1: 23,3%; clase 2: 23,3% + 26,7% = 50,0%; etc.)

se 1: 23,3%; clase 2: $23,3\% + 26,7\% = 50\%$; etc.)

Interpretación:

- a) De los 30 estudiantes, 13 tienen calificaciones entre 15 y 17, que representan el 43,3% del total del curso.
- b) 15 estudiantes tienen calificaciones entre 11 y 15, que representa el 50%, mientras el otro 50% tienen calificaciones entre 16 y 20.

Con los resultados obtenidos, se puede tomar una decisión, en este caso se observa que apenas un 50% de estudiantes tiene buenas calificaciones y para el docente es importante que este porcentaje sea mayor, por lo tanto, deberá implementar una nueva estrategia de enseñanza para que los estudiantes mejoren las calificaciones.

Este es un caso sencillo trabajado con calificaciones de estudiantes. De la misma manera se puede trabajar con datos obtenidos en ventas de productos, en compras, en rentabilidades de las empresas, etc.

Representación gráfica de datos cuantitativos

Los datos obtenidos a través de la distribución de frecuencias se pueden mostrar de manera más expresiva a través de gráficos estadísticos, donde se pueden visualizar de mejor manera las características de la variable observada.

El gráfico más común de utilizar es el diagrama de barras, más conocido como “histograma de frecuencias” ya sea absolutas o relativas.

En el diagrama de barras, en el eje de las ordenadas (eje Y) se colocan las clases con sus límites y en el eje de las abscisas (eje X) se colocan los valores de las frecuencias.

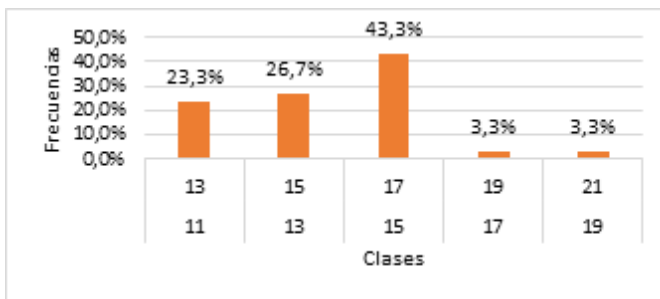
Cada barra representa una clase.

Histograma de frecuencias absolutas



Interpretación: Con calificaciones de 16 a 20 se encuentran 15 ($13 + 1 + 1$) estudiantes, así como 15 ($7 + 8$) estudiantes tiene entre 11 y 15 puntos.

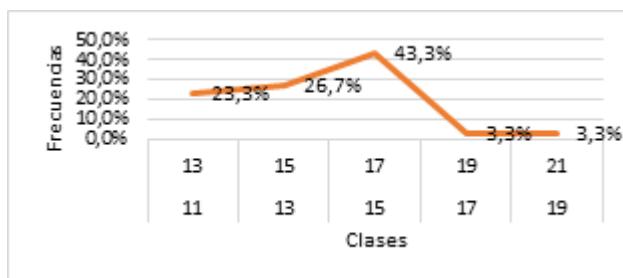
Histograma de frecuencias relativas



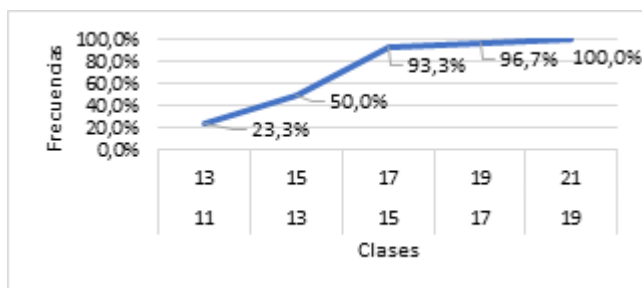
Interpretación: Con calificaciones de 15 a 20 se encuentran 50% ($43,3\% + 3,3\% + 3,3\%$) de los estudiantes, mientras que el otro 50% tienen entre 11 y 15 puntos.

Otro gráfico comúnmente utilizado es el polígono de frecuencias, donde se trabaja con los valores de los puntos medios en el eje de las abscisas.

Polígono de frecuencia absolutas relativas



Polígono de frecuencias acumuladas relativas



La interpretación es similar que en el histograma.

3.1.2 Construcción de una tabla de frecuencias de datos cualitativos

Para los datos cualitativos, Lind, Marchal & Wathen (2012) señalan que la distribución de frecuencias es la “Agrupación de datos cualitativos en clases mutuamente excluyentes que muestra el número de observaciones que hay en cada clase” (pág. 23).

En este caso, la información se clasifica en categorías, sin un orden particular. Es más sencillo trabajar con los datos cualitativos.

Ejemplo:

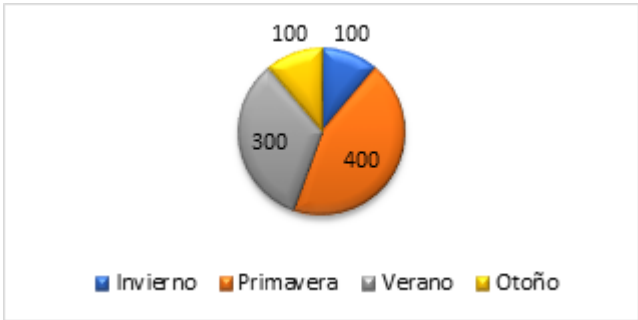
Una empresa de turismo realizó una encuesta a una muestra de 1.000 residentes de Sangolquí, para saber que estación del

año preferían. Los resultados obtenidos fueron los siguientes: a 100 les gustaba el invierno; a 400 la primavera; a 300 el verano y a 100, el otoño. Se desea resumieran los datos en una tabla de frecuencias.

Tabla de frecuencias cualitativas:

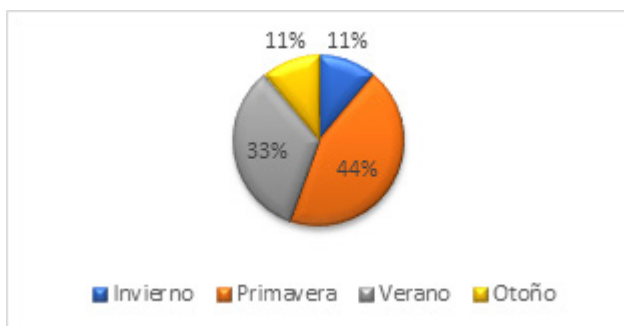
Clases	Frecuencia absoluta	Frecuencia relativa %
Invierno	100	11%
Primavera	400	44%
Verano	300	33%
Otoño	100	11%
	900	100%

Gráfico de frecuencias absolutas



Interpretación: De las 1.000 personas entrevistadas se observa que 400 prefieren la primavera, seguido de 300 que prefieren el verano, mientras que el invierno y el otoño son preferidos por 100 personas respectivamente.

Gráfico de frecuencias relativas



Interpretación:

De las 1.000 personas entrevistadas se observa que el 44% prefieren la primavera, seguido del 33% que prefieren el verano, mientras que el invierno y el otoño son preferidos por 11% respectivamente.

Capítulo

4

Medidas de tendencia

4. Medidas de tendencia central

Se tienen dos formas numéricas para describir los datos cuantitativos: medidas de ubicación y medidas de dispersión.

4.1 Medidas de ubicación

Generalmente se les denomina promedio o media aritmética.

De acuerdo con la clasificación de población y muestra se tiene:

4.1.1 Media poblacional

$$\begin{aligned} & \textit{Media poblacional} \\ &= \frac{\textit{Suma de valores presentes en la población}}{\textit{Nº de datos de la población}} \end{aligned}$$

En el lenguaje matemático también esta expresada esta fórmula con su respectiva nomenclatura:

$$\mu = \frac{\Sigma X}{N}$$

Símbolo	Significado
	Media poblacional
	Sigma (operación de suma) letra griega
X	Valores analizados
N	Número de valores (N mayúscula)
	Sumatoria de X valores

La media poblacional que se busca es una característica de la población y se denomina “Parámetro poblacional”.

Ejemplo:

Existen 42 tiendas comerciales en el centro comercial el Condamado ubicado al norte de Quito. A continuación, tenemos los valores de ingresos anuales en miles de dólares.

12	15	14	47	45	22	12	21	9	20	5	19	55	18
2	8	12	2	13	59	14	25	15	20	16	45	17	50
4	5	8	1	10	12	26	11	22	10	5	9	2	8

Se requiere encontrar la media aritmética e interpretar porque esta información es una población.

Al considerar la información de las ganancias anuales de todas las tiendas del centro comercial se convierte en una población. Para calcular la media aritmética se tiene que sumar todos los valores de la población obteniendo un ingreso total de 745 y dividir para e 42. Así la media aritmética es de 17.7 (miles de dólares).

$$\mu = \frac{\sum X}{N} = \frac{12+15+14+...+2+8}{42} = \frac{745}{42} = 17.7$$

La media poblacional de 17.7 es un parámetro poblacional y representa el número típico de ingresos anuales de las tiendas.

4.1.2 Media de una muestra

La media maestral es una característica del análisis de una porción de la población que se denomina “Estadístico”, y permite identificar una característica específica de la población.

Por ejemplo, una empresa dedicada a la fabricación de mayonesa no puede medir todos los envases de lote de producción para conocer cuál es el contenido, así que se selecciona 30 envases y se obtiene la media muestral; este valor se infiere (deduce o generaliza) para todo el lote.

$$\text{Media Muestral} = \frac{\text{Suma de todos los valores de la muestra}}{\text{Nº de datos de la muestra}}$$

En el lenguaje matemático también esta expresada esta fórmula con su respectiva nomenclatura:

$$S=\frac{\sum x}{n}$$

Símbolo	Significado
S	Media muestral
	Sigma (operación de suma) letra griega
X	Valores encontrados
n	Número de valores de la muestra (n minúscula)
	Sumatoria de X valores de la muestra

Ejemplo:

CNT estudia los minutos que consumen del plan de sus usuarios estrellas. Una muestra aleatoria de 16 personas muestra la cantidad de min consumidos ese mes.

5	6	10	45	15	18	20	29
23	19	9	5	24	36	14	17

Determine la media muestral

$$s = \frac{\sum X}{n} = \frac{5 + 6 + 10 + \cdots + 17}{16} = \frac{295}{16} = 18.44$$

El valor de la media aritmética de los minutos consumidos por los usuarios estrella de CNT de la muestra es de 18.44 minutos y constituye un “estadístico”.

4.1.3 Propiedades de la media aritmética

- 1. La media contiene todos los valores o cada uno de los datos.
- 2. La media no es igual a ningún valor de los datos, es decir la media es única.
- 3. Todo conjunto de datos de un intervalo contiene una media.
- 4. La suma de los valores que influenciaron a la media o

también llamados desviaciones con respecto al valor de la media es cero.

Expresado simbólicamente la propiedad $N^{\text{a}} 4$ se tiene que:

$$\Sigma(X - X_m) = 0$$

Ejemplo: La media de los valores 9, 5, 14 y 3 es 7.75.

$$\begin{aligned}\Sigma(X - X_m) &= (9 - 7.75) + (5 - 7.75) \\ &\quad + (14 - 7.75) + (3 - 7.75) = 0\end{aligned}$$

Esto significa que la media es el punto de equilibrio de los datos.

La media aritmética tiene un punto débil cuando se tiene un conjunto de datos con valores extremadamente pequeños o extremadamente grandes comparados con la mayoría de datos, lo que hace que este valor no tenga una representación significativa de los datos.

Ejemplo:

La rentabilidad en el año 2016 de 4 tiendas TOYOTA en el Ecuador fue de \$5, \$15, \$62, y \$90 millones de dólares. El ingreso medio de este caso es de \$ 43, valor que no es representativo dados los valores extremos.

Ejercicios de aplicación

1. Encontrar la media de los siguientes valores poblacionales: 12, 10, 8, 5, 7, 30 y 15

$$\begin{aligned}\mu &= \frac{\Sigma X}{N} = \frac{12 + 10 + 8 + 5 + 7 + 30 + 15}{7} = \frac{87}{7} \\ &= 12.43\end{aligned}$$

2. Encontrar la media de los siguientes valores muestrales: 5, 11, 9, 15 y 18

$$S = \frac{\Sigma X}{n} = \frac{5 + 11 + 9 + 15 + 18}{5} = \frac{58}{5} = 11.60$$

Con los datos del ejercicio anterior, demuestre que:

$$\Sigma(X - X_m) = 0$$

$$\Sigma(X - X_m) = (5 - 11,60) + (11 - 11,60) + (9 - 11,60) + (15 - 11,60) + (18 - 11,60) = 0$$

4.1.4 Media ponderada

La media ponderada es un caso especial de la media aritmética que se aplica cuando se tienen grupos de observaciones con el mismo valor.

Ejemplo:

Tropi Burger vende hamburguesas de carne, pollo y cerdo a \$2.50, \$3.00 y \$2.75 respectivamente. De las últimas 10 hamburguesas que se vendieron 5 fueron de carne, 3 de pollo y 2 de cerdo. Para determinar el precio promedio aplicamos la siguiente fórmula.

Calculando de la forma tradicional se tiene:

$$\begin{aligned} X_m &= \frac{2,50 + 2,50 + 2,50 + 2,50 + 2,50 + 3,00 + 3,00 + 3,00 + 2,75 + 2,75}{10} \\ &= 2,70 \end{aligned}$$

Una forma más fácil de calcular es aplicando la media ponderada:

$$X_w = \frac{\Sigma(WX)}{\Sigma W}$$

$$X_w = \frac{W_1 X_1 + W_2 X_2 + \dots + W_n X_n}{W_1 + W_2 + \dots + W_n}$$

Donde W es el peso o número de veces que se repite cada grupo de valores y X es el valor observado.

$$X_w = \frac{5(2,50) + 3(3,00) + 2(2,75)}{10} = 2,70$$

4.1.5 Mediana

Es una medida de ubicación de un conjunto de datos ordenados que muestra el valor central o la media de los valores centrales.

Esta medida de ubicación se utiliza cuando se tienen valores muy pequeños o muy grandes en relación con los demás del conjunto, tal que la media aritmética no resulta representativa.

Lind, Marchal & Wathen (2012) señalan las siguientes principales propiedades de la mediana:

1. No influyen en ella valores extremadamente grandes o pequeños. Por consiguiente, la Mediana es una valiosa medida de ubicación cuando dichos valores se presentan.
2. Es calculable en el caso de datos de nivel ordinal o más altos (...). (pág. 65).

Importante: Para facilitar la búsqueda de la Mediana, los datos tienen que estar ordenados de mayor a menor o viceversa

Ejemplo:

- a) Determine la Mediana de la cantidad de llamadas telefónicas realizadas por un grupo de estudiantes en el lunes:

3	5	6	4	2
---	---	---	---	---

2	3	4	5	6
---	---	---	---	---

El punto medio o posición es la casilla 3. Por tanto, la mediana es el valor 4 que se encuentra en la posición 3.

- a) En la siguiente muestra de calificaciones, determinar la Mediana:

8	15	11	19	14	10
---	----	----	----	----	----

8	10	11	14	15	19
---	----	----	----	----	----

El punto medio (posición) está entre las casillas 3 y 4, por tanto

la mediana es el promedio de los valores que se encuentra de estas casillas: $(11+14)/2 = 12,5$

En grupos de datos pequeños como los de los ejemplos son fáciles de ubicar la Mediana solo por observación. En casos donde se tienen decenas o más cantidad de datos, es conveniente aplicar la siguiente fórmula para ubicar la posición de Mediana:

$$\text{Posición de la mediana}(PM) = \frac{n + 1}{2}$$

4.1.6 Moda

La Moda también es una medida de ubicación. Es el valor que más se repite en el conjunto de datos. Para facilitar la búsqueda de la moda los datos tienen que estar ordenados de mayor a menor o viceversa, permitiendo una mayor visualización del valor repetido.

Ejemplo:

La empresa AMBACAR vendió durante el año 2016 la siguiente cantidad de autos. Determine la Moda respecto al número de autos vendidos.

Año 2016	Ventas	Año 2016	Datos ordenados
Enero	3	Enero	3
Febrero	15	Marzo	6
Marzo	6	Abril	11
Abril	11	Julio	11
Mayo	19	Diciembre	12
Junio	14	Junio	14
Julio	11	Octubre	14
Agosto	15	Febrero	15
Septiembre	15	Agosto	15

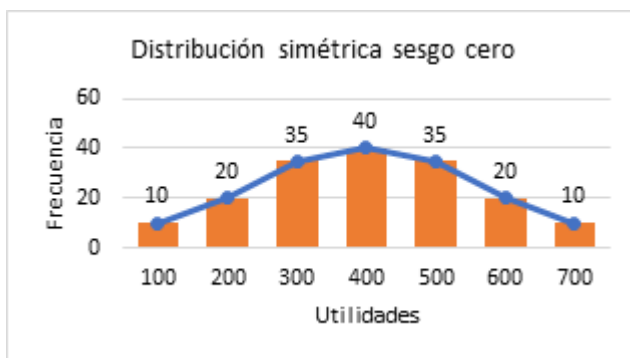
Octubre	14	Septiembre	15
Noviembre	20	Mayo	19
Diciembre	12	Noviembre	20

Conclusión: En los meses de febrero, agosto y septiembre se vendieron 15 autos en cada mes.

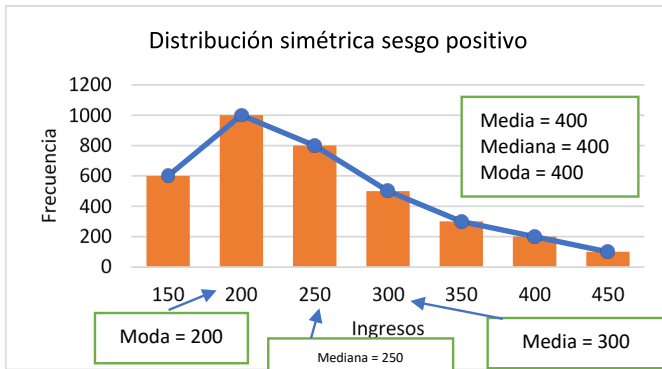
4.1.7 Posiciones relativas de la media, mediana y moda

En un histograma se puede observar la distribución de la media, mediana y moda para lo cual dependiendo del gráfico podremos saber si tiene una distribución simétrica o asimétrica.

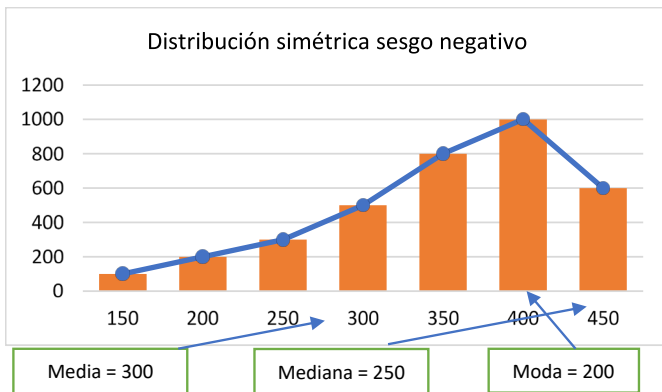
En el caso de una distribución de forma simétrica o forma campana, la media, mediana y moda siempre son iguales, como podemos observar en la imagen la moda, mediana y media son iguales a 6.



En el caso de que una distribución no sea simétrica, podemos tener dos tipos de sesgos, el sesgo positivo el cual tiene como característica que la media es mayor que la mediana y la moda y su gráfica se encuentra ligeramente hacia la izquierda:



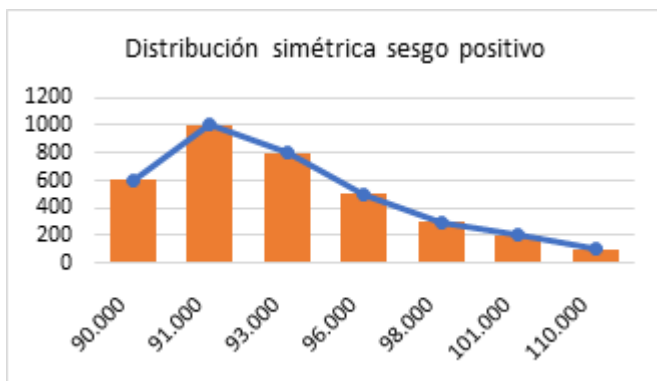
En el caso de que un sesgo negativo, la moda es mayor que la mediana y la media y su gráfica se encuentra inclinada hacia la derecha:



En un conjunto de datos, no siempre se tendrá la Moda, entonces, la evaluación se realizará en base a la Mediana y a la Media.

Ejemplo:

Los ingresos mensuales de uno de los almacenes TIA se organizaron en una distribución de frecuencias y se determinó una media de \$ 96.000, la mediana de \$ 93.000 y la moda de \$ 91.000. Grafique las ventas y determine si la distribución es simétrica, tiene sesgo positivo o sesgo negativo. Explique la respuesta.



Solución: La distribución tiene sesgo positivo por cuanto la Media de las ventas es mayor que la Mediana y la Moda

4.1.8 Media geométrica

Es muy útil debido a que permite calcular los cambios porcentuales, índices y cifras relativas, siendo de gran importancia tanto para la matemática como para la estadística y la economía ya que ayuda a conocer cuál es el cambio verdadero que se produce ya sea en cuanto a salarios, ventas realizadas u otro tipo de datos numéricos.

La media geométrica siempre será menor o igual que la media aritmética y trabaja únicamente con números positivos.

La fórmula de la media geométrica es la siguiente:

$$MG = \sqrt[n]{(X_1)(X_2)(X_3) \dots (X_n)}$$

Ejercicios:

- 1) Un inversionista recupero en 3 años el 2%, 4% y 5% respectivamente de la inversión que realizó en un proyecto inmobiliario, ¿Cuál es la media geométrica de la recuperación de la inversión?

El 2% de la recuperación de la inversión corresponde a 102%

(2% más que el 100%) y se expresa como 1,02. El mismo procedimiento para todos los casos.

Por lo tanto tenemos:

$$MG = \sqrt[3]{(1,02)(1,04)(1,05)}$$

$$MG = 1,0366$$

Esto significa que la tasa promedio de recuperación es de 3,66%.

2) El incremento salarial de un obrero de los dos últimos años fue de 8% y 10% determine la media geométrica del incremento salarial.

$$MG = \sqrt[n]{(X1)(X2)}$$

$$MG = \sqrt[2]{(1,08)(1,1)}$$

$$MG = 1,0899$$

El incremento salarial del obrero fue de 8,99%.

3) La utilidad de una constructora en cuatro proyectos fue de 2%, 4%, 7% y 9% ¿Cuál es la media geométrica de las ganancias?

$$MG = \sqrt[n]{(X1)(X2)(X3)(X4)}$$

$$MG = \sqrt[4]{(1,02)(1,04)(1,07)(1,09)}$$

$$MG = 1,0546$$

La media geométrica de las utilidades es de 5,46%.

4) Calcular la media geométrica de los siguientes porcentajes: 6%, -8%, 10% y 15% que presentan las utilidades reportadas en los últimos 4 años por el departamento de ventas en la empresa SFDK S.A.

$$MG = \sqrt[n]{(X1)(X2)(X3)(X4)}$$

$$MG = \sqrt[4]{(1,06)(0,92)(1,10)(1,15)}$$

$$MG = 1,0539$$

Las ganancias del departamento de ventas fueron el 5,39%.

5) En la empresa AUTOV S.A. dedicada a la compra y venta de autos se presentaron incrementos porcentuales en las ganancias en los 5 últimos años siendo estos 5%, 7%, 9%, 10%, 11% y 11,5%. Calcular la media geométrica del incremento de las ganancias.

$$MG = \sqrt[n]{(X1)(X2)(X3)(X4)(X5)}$$

$$MG = \sqrt[5]{(1,05)(1,07)(1,1)(1,11)(1,115)}$$

$$MG = 1,0887$$

Por lo tanto el la media geométrica del incremento porcentual en ventas es de 8,87%.

La Media Geométrica también puede aplicarse para determinar el cambio porcentual promedio en ciertos períodos de tiempo y para ello se usa la siguiente fórmula:

$$MG = \sqrt[n]{\frac{\text{Valor final del período}}{\text{Valor inicial del período}}} - 1$$

Ejercicios:

1) En la ciudad de Latacunga se muestra un mayor aumento en el número de empleos entre los años 1990 y 2015, se estimó que los empleos subieron de 5.800 hasta 9.500. ¿Cuál es la media geométrica del incremento anual de dicho período?

Datos:

$$n = 25 \text{ (diferencia entre 2015 - 1990)}$$

$$\text{Valor inicial} = 5.800$$

$$\text{Valor final} = 9.500$$

$$MG^n = \sqrt[n]{\frac{\text{Valor final del periodo}}{\text{Valor inicial del periodo}}} - 1$$

$$MG^{25} = \sqrt[25]{\frac{9500}{5800}} - 1 \quad MG=0,0199$$

Por lo tanto, el incremento anual de empleos en el período es de 1,99%.

2) En el proceso de reclutamiento para aspirantes a bomberos en la ciudad de Guayaquil entre los períodos de 1996 y 2016 aumentó de 150 a 650 aspirantes. Calcule la media geométrica del incremento anual de aspirantes

$$MG^n = \sqrt[n]{\frac{\text{Valor final del periodo}}{\text{Valor inicial del periodo}}} - 1$$

$$MG^{20} = \sqrt[20]{\frac{650}{150}} - 1$$

$$MG=0,076$$

Hubo un incremento anual de aspirantes del 7,60%.

3) En el año 2000 existían 20.200 jóvenes amantes de la música electrónica en la ciudad de Quito, para el 2016 el número de jóvenes aumentó en 39.800. ¿Cuál es la media geométrica del incremento anual en el período establecido?

$$MG^n = \sqrt[n]{\frac{\text{Valor final del periodo}}{\text{Valor inicial del periodo}}} - 1$$

$$MG^{16} = \sqrt[16]{\frac{39800}{20200}} - 1$$

$$MG=0,0433$$

Por lo tanto, el incremento anual de jóvenes amantes de la música electrónica fue de 4,33%.

4) En la ciudad de Guayaquil se reportaron 5.800 infracciones de tránsito en el 2002, para el año 2016 se reportaron 11.320 infracciones de tránsito. Calcular la media geométrica del incremento anual de infracciones de tránsito.

$$MG^n = \sqrt[n]{\frac{\text{Valor final del periodo}}{\text{Valor inicial del periodo}}} - 1$$

$$MG^{14} = \sqrt[14]{\frac{11320}{5800}} - 1$$

$$MG=0,0489$$

El incremento anual de infracciones de tránsito es de 4,89%.

5) La empresa KISKO S.A. debido a sus buenos productos obtenía ganancias de \$600500 en el año 2001, para el año 2016 la empresa arrojo resultados de ganancias de \$1300000. Calcule la media geométrica del aumento anual de las ganancias.

$$MG = \sqrt[n]{\frac{\text{Valor final del período}}{\text{Valor inicial del período}}} - 1$$

$$MG = \sqrt[15]{\frac{1300000}{600500}} - 1$$

$$MG = 0,0528$$

El incremento anual de ganancias en la empresa es de 5,28%.

Capítulo

5

Medidas de dispersión

5. Medidas de dispersión

La dispersión es el grado de concentración de valores respecto a la media aritmética.

Una medida de dispersión pequeña indica que los datos se encuentran acumulados cerca de la media aritmética, mientras que una dispersión grande indica que los datos se encuentran alejados.

Entre las medidas de dispersión se encuentran: el Rango; la Desviación media; la Desviación estándar.

5.1 Rango

La media más simple que existe es el rango porque representa la diferencia entre los valores máximo y mínimo de un conjunto de datos.

Rango = Valor máximo – valor mínimo

Nota: El rango se emplea mucho en aplicaciones de control de procesos estadísticos, debido a que resulta fácil calcular e interpretar.

5.2 Desviación media

Es la “Media aritmética de los valores absolutos de las desviaciones con respecto a la media aritmética” (Lind , Marchal, & Whaten, 2012, pág. 76) y su fórmula de cálculo es:

$$DM = \frac{\sum |X - \bar{X}|}{n}$$

Ejemplo:

En una muestra de 10 clientes, la empresa Che Farina contabilizó el tiempo que se demora en servir la pizza a sus clientes. Determine la Desviación media y comente al respecto.

Nº	X
1	18
2	12
3	13
4	15
5	16
6	11
7	17
8	16
9	13
10	10

Desarrollo:

Nº	X	X-Xm	X-Xm	Li sup	Li inf	Xm
1	18	3,90	3,90	16,40	11,80	14,10
2	12	- 2,10	2,10	16,40	11,80	14,10
3	13	- 1,10	1,10	16,40	11,80	14,10
4	15	0,90	0,90	16,40	11,80	14,10
5	16	1,90	1,90	16,40	11,80	14,10
6	11	- 3,10	3,10	16,40	11,80	14,10
7	17	2,90	2,90	16,40	11,80	14,10
8	16	1,90	1,90	16,40	11,80	14,10

9	13	-	1,10	1,10	16,40	11,80	14,10
10	10	-	4,10	4,10	16,40	11,80	14,10
141			23,00				

$$n = 10$$

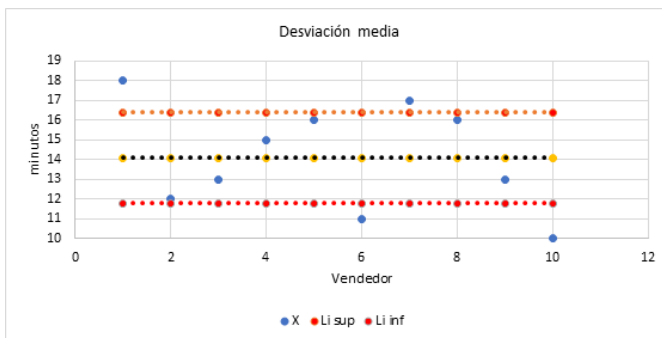
$$DM = 16/10 = 2,30$$

$$X_m = 14,10$$

$$Li \text{ sup} = 16,40$$

$$Li \text{ inf} = 11,80$$

Representación gráfica:



Conclusión: Los clientes son atendidos en un tiempo medio de 14,10 minutos con una desviación media de 2,30 minutos. (14,10 $\pm$ 2,30).

5.3 Varianza y desviación estándar

La Varianza y la desviación estándar se fundamentan en la desviación media, pero en lugar de trabajar con el valor absoluto de las desviaciones lo hacen con el cuadrado de las desviaciones.

La Varianza y la Desviación estándar se calcula tanto para la población como para la muestra. En esta última varía porque se utiliza (n-1) en lugar de N.

Las fórmulas son las siguientes:

Fórmulas	Población	Muestra
Varianza	$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N}$	$S^2 = \frac{\Sigma(X-Xm)^2}{n-1}$
Desviación estándar	$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$	$S = \sqrt{\frac{\Sigma(X-Xm)^2}{n-1}}$

5.3.1 Varianza

Es la media aritmética de las desviaciones de la media elevadas al cuadrado.

La Varianza nunca será negativa y será cero **sólo si todas las observaciones son las mismas**.

5.3.2 Desviación Estándar

En los dos casos, la Desviación estándar es la raíz cuadrada de la Varianza

Nomenclatura:

$\sigma^2 =$ varianza de la población

$\sigma =$ desviación estándar de la población

$S^2 =$ varianza de la muestra

$S =$ desviación estándar de la muestra

$\mu =$ media de la población

$X =$ valor observado de una población o muestra

$Xm =$ media de la muestra

$N =$ número de datos de la población

$n =$ número de datos de la muestra

Ejemplos:

- 1) La empresa Leo Salvador, instala el sistema para puertas automáticas para garajes. Los 5 empleados que trabajan en la empresa realizaron instalaciones el sábado anterior y registraron el siguiente tiempo en minutos para instalar los 5 sistemas automáticos:

28	32	24	46	44
----	----	----	----	----

Determine la media, varianza, la desviación estándar y comente el resultado.

N°	X	$X - U$	$(X - U)^2$	U	Lim sup	Lim inf
1	28	(6,80)	46,24	34,80	43,53	26,07
2	32	(2,80)	7,84	34,80	43,53	26,07
3	24	(10,80)	116,64	34,80	43,53	26,07
4	46	11,20	125,44	34,80	43,53	26,07
5	44	9,20	84,64	34,80	43,53	26,07
	174		380,80			

$$N = 5$$

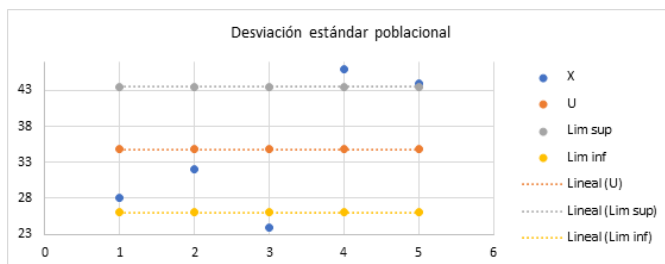
$$U = 174/5 = 34,8$$

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

$$\sigma^2 = \frac{380.80}{5}$$

$$\sigma^2 = 76,16$$

$$\sigma = 8,73$$



Ejemplos de conclusiones:

- a) La empresa instala los sistemas de puertas automáticas en un tiempo medio de 34,8 minutos, con una desviación estándar de 8,73 minutos ($34,8 \pm 8,73$).
 - b) La empresa instala las puertas automáticas en un tiempo medio de 34,8 minutos, con un tiempo mínimo de 29,07 minutos y un máximo de 43,58 minutos.
 - c) Tan solo dos empleados instalan dentro de la banda de tiempo establecida.
- 2) La Asociación de Propietarios de Hoteles de Atacames, realizó una encuesta para conocer las tarifas de los hoteles entre semana en el área. En seguida aparece la tarifa por cuarto para huéspedes de negocios en una muestra de 5 hoteles.

101	97	103	120	78
-----	----	-----	-----	----

Determine la media, varianza, la desviación estándar y comente el resultado.

Nº	X	X - Xm	(X - Xm) ²	Xm	Lim sup	Lim inf
1	101	1,20	1,44	99,8	114,82	84,78
2	97	(2,80)	7,84	99,8	114,82	84,78
3	103	3,20	10,24	99,8	114,82	84,78

4	120	20,20	408,04	99,8	114,82	84,78
5	78	(21,80)	475,24	99,8	114,82	84,78
	499		902,80			

$$N = 5$$

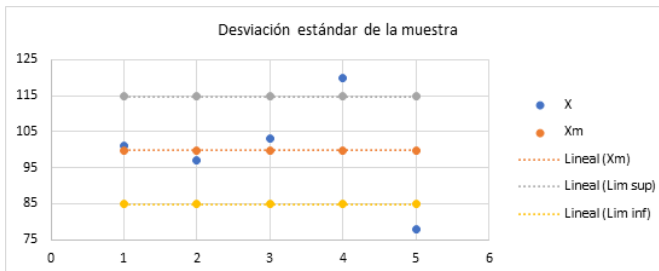
$$U = 499/5 = 99,80$$

$$S^2 = \frac{\sum (X - X_m)^2}{n-1}$$

$$S^2 = \frac{902,80}{4}$$

$$S^2 = 225,70$$

$$S = 15,02$$



Ejemplos de Conclusiones:

- Las tarifas de los hoteles entre semana tuvieron una media de \$ 99,8 con una desviación estándar de \$ 15,02. $(99,8 \pm 15,02)$
- Las tarifas de los hoteles entre semana presentaron una media de \$ 99,8 fluctuando entre \$ 84,72 y \$ 114,82.
- Tres de los hoteles cobraron tarifas dentro de la banda calculada.

5.3.3 Interpretación y usos de la desviación estándar

La desviación estándar generalmente se utiliza como medida para comparar la dispersión de dos o más conjuntos de observaciones.

Teorema de Chebyshev

Una desviación estándar pequeña indica que los datos se encuentran cercanos a la media (poco dispersos), mientras que la desviación estándar grande significa que los datos se encuentran alejados de la media (muy dispersos).

El matemático ruso P. L. Chebyshev establece un teorema que permite determinar la mínima porción de los valores que se encuentran a cierta cantidad de desviaciones estándares de la media.

Lind, Marchal & Wathen (2012) resaltan que “En cualquier conjunto de observaciones (muestra o población), la proporción de valores que se encuentran a k desviaciones estándares de la media es de por lo menos $1 - 1/k^2$, siendo k cualquier constante mayor que 1” (pág. 85)

Ejemplo:

La media aritmética de la suma mensual que aportan los empleados de la Empresa Leo Salvador S.A. al plan de reparto de utilidades de la compañía es de \$ 25,40 y la desviación estándar, de \$ 4,20. ¿Por lo menos qué porcentaje de las aportaciones se encuentra en más de 2 desviaciones estándares y menos 2 desviaciones de la media?

$$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

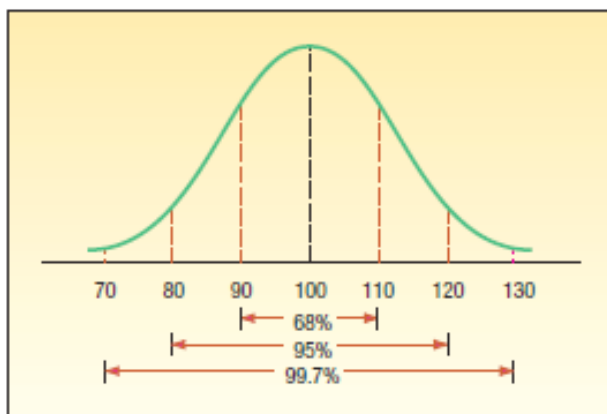
Respuesta: alrededor de 75%

La regla empírica o regla normal

La regla empírica menciona que:

En cualquier distribución de frecuencias simétrica con forma de campana, aproximadamente 68% de las observaciones se encontrarán entre más y menos una desviación estándar de la media; cerca de 95% de las observaciones se encontrarán entre más y menos dos desviaciones estándares de la media y, de hecho todas (99.7%), estarán entre más y menos tres desviaciones estándares de la media (Lind , Marchal, & Whaten, 2012, pág. 86).

Estas características se demuestran en la figura que tiene forma de campana.



Curva simétrica con forma de campana que muestra las relaciones entre la desviación estándar y las observaciones. Figura tomada de (Lind , Marchal, & Whaten, 2012, pág. 86)

Todas las observaciones (se asumen la totalidad pero en realidad es el 99,7%) se encuentran bajo la curva acampanada que representa 3 desviaciones estándar.

$$X_m = 100$$

$$S = 10$$

$$\text{Rango} = X_m \pm Y (S)$$

$$Y = \text{desviación } 1, 2, 3$$

$$\text{Rango} = 100 \pm 3 (10)$$

$$\text{Rango} = \text{Valor máximo} - \text{valor mínimo}$$

$$V_{\text{Max}} = 100 + 30 = 130$$

$$V_{\text{min}} = 100 - 30 = 70$$

Conclusión:

Para el ejemplo de un conjunto de datos observados con una media de 100 y $S = 10$, el total de observaciones se encuentra entre 70 y 130, y el rango es 60.

Ejemplo:

Las calificaciones de una muestra de estudiantes de un curso de Estadística son semejantes a una distribución simétrica acampanada. La media muestral es 14; la desviación estándar es \$2. Aplicando la regla empírica:

- a) ¿El 68% de las calificaciones entre que calificaciones se encuentra?

$$\text{Rango} = X_m \pm Y (S)$$

$$\text{Rango} = 14 \pm 1 (2)$$

$$V_{\text{máx}} = 14 + 2 = 16$$

$$V_{\text{min}} = 14 - 2 = 12$$

Respuesta: El 68% de las calificaciones se encuentra en-

tre 12 y 16 puntos

b) ¿El 95% en que rango se sitúa?

$$\text{Rango} = 14 \pm 2 \text{ (2)}$$

$$V_{\text{máx}} = 14 + 4 = 18$$

$$V_{\text{mín}} = 14 - 4 = 10$$

Respuesta: El 95% de las calificaciones se encuentra entre 10 y 18 puntos

c) ¿El 99,7% está entre que valores?

$$\text{Rango} = 14 \pm 3 \text{ (2)}$$

$$V_{\text{máx}} = 14 + 6 = 20$$

$$V_{\text{mín}} = 14 - 6 = 8$$

Respuesta: El 99,7 de las calificaciones se encuentra entre 8 y 20 puntos

5.3.4 Media aritmética y desviación estándar de datos agrupados

Media aritmética

Cuando los datos se encuentran en una tabla de frecuencias entre límites de intervalos, se dice que son datos agrupados.

Se representa de la siguiente manera:

$$\bar{x}_m = \frac{\sum fM}{n}$$

Siendo:

f = la frecuencia en cada clase

n = la suma total de las frecuencias

$M = P_m$ = el punto medio de cada clase (promedio de los valores de cada clase).

Desviación estándar de datos agrupados

Se representa con la siguiente formula:

$$s = \sqrt{\frac{\sum f(M - \bar{X})^2}{n - 1}}$$

Como en el caso de la desviación estándar de datos no agrupados, nos indica el grado de dispersión de un conjunto de datos analizados, respecto a la media.

Ejemplos:

- 1) Los gastos en publicidad constituyen un elemento significativo del costo de los artículos vendidos. En seguida aparece una distribución de frecuencias que muestra los gastos de publicidad de una muestra de 60 compañías manufactureras del Ecuador.

Determinar la media y la desviación estándar.

Gastos en pu- b l i c i d a d (\$ millones)		Nº com- pañías
Límite inferior	Límite s u p e - rior	frecuen- cia
25	35	5
35	45	10
45	55	21

55	65	16
65	75	8

Desarrollo:

Li	Ls	f	Punto medio M	f*M	(M - Xm) ²	f*(M-X) ²
25	35	5	30	150	484,00	2.420,00
35	45	10	40	400	144,00	1.440,00
45	55	21	50	1050	4,00	84,00
55	65	16	60	960	64,00	1.024,00
65	75	8	70	560	324,00	2.592,00
		60		3120		7.560,00

$$n = 60$$

$$X_m = 52$$

$$X_m = \frac{\sum fM}{n}$$

$$X_m = \frac{3.120}{60}$$

$$S = \sqrt{\frac{\sum f(M - \bar{X})^2}{n-1}}$$

$$S = \sqrt{\frac{7.560}{59}}$$

$$S = 11,32$$

Conclusiones:

- a) Los gastos en publicidad presentan una media de \$ 52 millones con una desviación estándar de \$ 11,32 millo-

nes.

- b) La media se encuentra en la tercera clase (45 - 55).
 - c) Se tiene una media de $52 \pm 11,32$ millones.
 - d) Se tiene una media de 52 millones en un rango de 40,68 y 63,32.
- 2) Las ventas en miles de dólares del supermercado SANTA MARÍA en los 30 días del mes de noviembre son las siguientes:

0,8	6,5	7,1	3,6	3,3	7,9	5,8	6,2	1,2	3,7
6,1	7,4	8,7	4,0	6,2	8,4	2,8	4,5	2,2	7,8
9,2	3,4	2,6	5,3	3,9	7,0	7,3	9,3	6,6	7,6

Determinar la media y la desviación estándar.

Desarrollo:

A conveniencia del interesado, se forman 5 clases con límites determinados a conveniencia:

Límites
0-2
2-4
4-6
6-8
8-10

Determinar la frecuencia, la media aritmética y la desviación estándar.

Límites	Frecuencia	M	f * M	$(M - \bar{X})^2$	f * $(M - \bar{X})^2$
0-2	2	1	2	19,951	39,902

2-4	9	3	27	6,084	54,760
4-6	3	5	15	0,218	0,653
6-8	12	7	84	2,351	28,213
8-10	4	9	36	12,484	49,938
	30		164	41,089	173,467

$$n = 30$$

$$X_m = 5,47$$

$$X_m = \frac{\sum fM}{n}$$

$$X_m = \frac{164}{30}$$

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$

$$\sigma = \sqrt{\frac{176,467}{30}}$$

$$\sigma = 2,43$$

Conclusiones:

- Las ventas presentan una media de \$ 5,47 mil dólares con una desviación estándar de \$ 2,43.
- La media se encuentra en la tercera clase (4 - 6).
- Se tiene una media de $5,57 \pm 2,43$ miles de dólares.
- Se tiene una media de 5,57 mil en un rango de 3,04 y 7,90

5.3.5 Otras medidas de posición

La medida de dispersión más utilizada es la desviación estándar.

dar, sin embargo, existen otras formas de describir la variación o dispersión de un conjunto de datos, siendo estas, los cuartiles, deciles y percentiles. (Lind , Marchal, & Whaten, 2012, pág. 111).

La mediana también es parte de esta familia de medidas de dispersión, que divide el conjunto de datos en dos partes iguales.

Para trabajar con estas medidas de dispersión se requiere que el conjunto de datos u observaciones se encuentren ordenados de menor a mayor o viceversa.

Las cuatro medidas de posición trabajan en base a percentiles, que se indican a continuación:

Los cuartiles (C): dividen un conjunto de datos ordenados en 4 partes iguales, el 25%, 50%, 75% de las observaciones.

25%	50%	75%	100%
cuar- til 1	cuar- til 2	cuar- til 3	cuartil 4

Los deciles (D): dividen un conjunto de datos ordenados en 10 partes iguales.

10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
de - cil 1	2	3	4	5	6	7	8	9	10

Los percentiles (P): dividen un conjunto de datos ordenados en 100 partes iguales.

1%	2%	3%		41%	42%		98%	99%	100%
----	----	----	-------	-----	-----	-------	-----	-----	------

La mediana (Med): divide al conjunto de datos en dos partes iguales:

50%	100%
-----	------

Otra medida de posición muy utilizada a nivel mundial es el Quintil. Es una medida socio económica oficial que sirve para

medir los ingresos de la población. El quintil uno representa la población más pobre; el quintil cinco representa la población más rica.

Los **quintiles (Q)**: dividen un conjunto de datos ordenados en 5 partes iguales:

20%	40%	60%	80%	100%
quin- til 1	quin- til 2	quin- til 3	quin- til 4	quin- til 5

Otra medida de dispersión estadística es el “Rango intercuartil”, que es la diferencia entre el cuartil 1 y el cuartil 3:

$$RC = C_3 - C_1$$

Con esta medida, se eliminan los valores extremadamente alejados, y permite desarrollar el diagrama de caja, que es un instrumento para evaluar la dispersión.

En todos los casos de cálculos de percentiles, se debe ubicar inicialmente la posición del percentil, mediante la fórmula:

$$Lp = (n + 1) \frac{p}{100}$$

p: representa el percentil de la medida que se busca.

n: número de datos a analizar.

Ejemplo:

Pollo GUS realiza entregas a domicilio en la zona de San Rafael. En una muestra de diez entregas el Gerente recopiló la siguiente información en minutos y quiere conocer: el cuartil 1, el cuartil 3, el rango intercuartil; la mediana; el decil 3, el quintil 2.

20	26	31	27	28	30	17	22	34	18
----	----	----	----	----	----	----	----	----	----

Desarrollo:

Se ordenan los datos:

17	18	20	22	26	27	28	30	31	34
----	----	----	----	----	----	----	----	----	----

Cálculo del cuartil 1 C1:

$$L_p = (n + 1) \frac{p}{100}$$

$$L_{25} = (10 + 1) \frac{25}{100} \quad L_{25} = 2,75$$

El C1 se encuentra entre los valores de la segunda y tercera posición.

2da posición, valor = 18

3era posición, valor = 20 diferencia entre los dos valores = 2

Decimal de la ubicación = 0,75 se multiplica por la diferencia entre valores 2 = 1,50

Posición real del C1 = el valor menor = 18 + 1,50 = 19,50

$$\frac{2}{3} \frac{18}{20} (2 * 0,75) = 1,50 \rightarrow \llbracket 18 + 1,50 = 19,50 \rrbracket$$

C1 = 19,50 minutos

Cálculo del cuartil 3 C3:

$$L_{75} = (10 + 1) \frac{75}{100} \quad L_{75} = 8,25$$

El C3 se encuentra entre los valores de la octava y novena posición.

$$\frac{8}{9} \frac{30}{31} (1 * 0,25) = 0,25$$

$$\rightarrow \llbracket 30 + 0,25 = 30,25 \rrbracket$$

C3 = 30,25 minutos

El rango intercuartil RC = 30,25 - 19,50 = 10,75 minutos

Cálculo de la mediana:

Posición de la mediana:

$$L_{50} = (10 + 1) \frac{50}{100} \quad L_{50} = 5,50$$

Se encuentra en el medio de los valores de la quinta y sexta posición

$$Med = \frac{26+27}{2} = 26,5 \quad Med = 26,5 \text{ minutos}$$

Cálculo del decil 3, D3

$$L_{30} = (10 + 1) \frac{30}{100} \quad L_{75} = 3,30$$

El D3 se encuentra entre los valores de la tercera y cuarta posición.

$$\frac{3}{4} \frac{20}{22} (1 * 0,30) = 0,30 \rightarrow \llbracket 20 + 0,20 = 20,20 \rrbracket$$

D3 = 20,20 minutos

Cálculo del quintil 2 Q1:

$$L_{40} = (10 + 1) \frac{40}{100} \quad L_{75} = 4,40$$

El Q2 se encuentra entre los valores de la cuarta y quinta posición.

$$\frac{4}{5} \frac{20}{26} (6 * 0,40) = 2,40 \\ \rightarrow \llbracket 20 + 2,40 = 22,40 \rrbracket$$

Q2 = 22,40 minutos

5.3.6 Diagrama de caja

También se conoce como diagrama de caja y bigotes. Es un gráfico que se construye en base a los cuartiles y permite visualizar

la distribución de un conjunto de datos.

En su representación se utilizan: los tres cuartiles, los valores mínimos, máximo y mediana de los datos.

Ejemplo:

En base al ejemplo anterior, el Gerente desea conocer ¿Cuánto tiempo tarda una entrega normal? ¿En qué margen de tiempo se realizan la mayoría de entregas?

Se presenta la información ordenada

17	18	20	22	26	27	28	30	31	34
----	----	----	----	----	----	----	----	----	----

Valor mínimo = 19

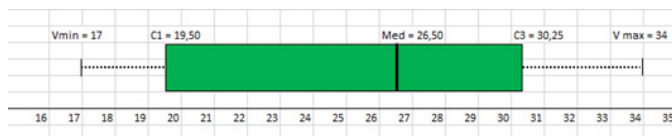
Valor máximo = 34

Cuartil 1 C_1 = 19,50

Cuartil 3 C_3 = 30,25

Mediana = 26,50

Diagrama de caja



El diagrama de caja muestra que el 50% de las entregas se encuentra entre 19,50 y 30,25 minutos.

El rango intercuartil representa la dispersión de la mayoría de las entregas (extremos de la caja) y es 10,75 minutos.

5.3.7 Sesgo

Es un error que generalmente aparece en los resultados de

un estudio debido a factores que dependen de la recolección, análisis, interpretación, publicación o revisión de los datos que pueden conducir a conclusiones que son sistemáticamente diferentes de la verdad o incorrectas acerca de los objetivos de una investigación

En la literatura estadística se utilizan diversas fórmulas para calcular el sesgo. La más sencilla, ideada por el profesor Karl Pearson (1857-1936), se basa en la diferencia entre la media y la mediana.

El sesgo puede variar entre -3 a +3.

eleva- do	modera- do	bajo	bajo	modera- do	eleva- do
-3	-2	-1	1	2	3

Un valor de 1,60 indica un sesgo positivo moderado, mientras que un valor de 2,8 indica un sesgo positivo alto.

Un valor de -0,30 indica un sesgo negativo bajo y un valor de -1,7 indica un sesgo negativo moderado.

Coefficiente de sesgo de Pearson

$$sk = \frac{3(Xm - mediana)}{S}$$

Coefficiente de sesgo calculado con software

$$sk = \frac{n}{(n-1)(n-2)} \left[\sum \left(\frac{X - Xm}{S} \right)^3 \right]$$

El más utilizado es el método de Pearson

Ejemplo:

Los siguientes valores son los sueldos iniciales, en miles de dólares, de una muestra de cinco graduados de contabilidad, quie-

nes aceptaron puestos de contaduría pública el año pasado.

35	22	23	18	26
----	----	----	----	----

Calcule el coeficiente de sesgo por el método de Pearson.

$$\bar{X}_m = 24,80$$

$$\text{Med} = 23$$

$$S = 6,38$$

$$S_k = \frac{(\bar{X}_m - \text{Med})}{s}$$

$$S_k = \frac{(24,80 - 23)}{6,38}$$

$$S_k = 0,85$$

Que representa un sesgo positivo bajo

Capítulo

6

Relación entre variables

6. Relación entre variables

Las organizaciones en general, públicas o privadas, grandes o pequeñas, de interés social o de inversión, se mueven a través de planes, programas y proyectos.

Muchas veces la toma de decisiones en una empresa depende de dos o más variables que están relacionadas entre sí, y es muy importante conocer qué tipo de relación hay entre ellas y la intensidad de la misma, esto se puede hacer mediante diversas técnicas para medir la asociación entre dos variables, entre ellos tenemos: el coeficiente de correlación creado por Karl Pearson que indica el grado de relación que existe entre una variable y otra dentro de un mismo plan o situación, además de usar un diagrama de dispersión que muestra gráficamente la relación entre ambas variables a través de la ecuación de regresión, entre otros.

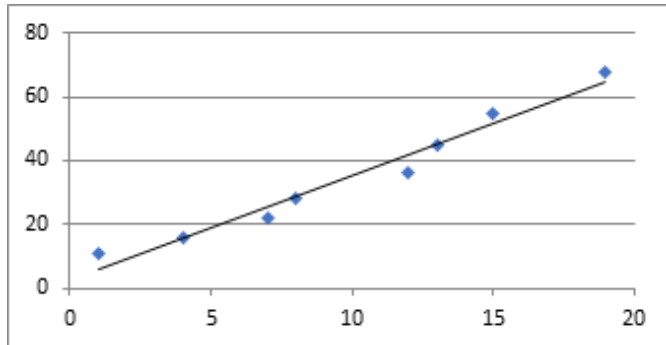
6.1 Análisis de correlación

Conjunto de métodos para medir la relación que existe entre dos variables y la magnitud de dicha relación, estas variables pueden ser cuantitativas o cualitativas, para una representación gráfica de la relación entre variables cuantitativas podemos usar el diagrama de dispersión.

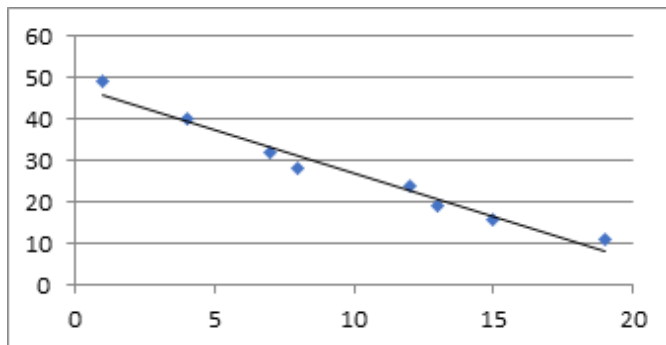
Existen dos tipos de variables que intervienen en este análisis, tenemos a la variable independiente o “causa” representada en el eje “X” y la variable dependiente o “efecto” representada en el eje “Y”.

Pueden existir varios tipos de relaciones, entre ellas:

- La relación lineal positiva que indica que entre las variables “X” y “Y” existe una relación directa es decir si los valores de “X” aumentan, los valores de “Y” también tienden a aumentar.



-La relación lineal negativa quiere decir que existe una relación inversa entre las variables “X” y “Y”, es decir que, si los valores de “X” tienden a aumentar, los valores de “Y” tienden a disminuir y viceversa.



6.2 Coeficiente de correlación

Es una medida de la magnitud de la relación lineal que existe entre dos variables.

Varía entre -1 y 1, inclusive.

Un valor cercano a 0 indica que hay poca asociación entre las variables.

Un valor cercano a 1 indica una asociación directa o positiva entre las variables.

Un valor cercano a -1 indica una asociación inversa o negativa entre las variables.

Aproximadamente se puede representar de la siguiente manera:

de elevada a fuerte	moderado	débil	débil	moderado	de elevada a fuerte
-1	-0,66	-0,33	0,33	0,66	1

Fórmula para cálculo:

$$r = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{(n - 1)(S_x)(S_y)}$$

Se considerará la desviación estándar dependiendo el caso: población o muestra:

Ejemplo:

Una empresa se plantea cambiar la composición de uno de sus productos utilizando un nuevo material. Antes de tomar una decisión, la empresa decide realizar un ensayo para estudiar la posible relación entre la utilización de dicho material y el número de defectos. Para ello analiza lotes con diferentes porcentajes del nuevo material y toma los siguientes datos:

Caso	% nuevo material X	Nº defectos Y
1	1	3
2	3	6
3	10	27
4	5	8
5	7	32
6	22	40
7	5	13

8	13	21
9	2	4
10	18	28
11	4	8
12	9	15

Desarrollo: Al tratarse de un ensayo, es un análisis muestral.

X	Y	X - Xm	Y - Ym	(X - Xm)²	(Y - Ym)²	(X - Xm) (Y - Ym)
1	3	- 7,25	- 14,08	52,56	198,34	102,10
3	6	- 5,25	- 11,08	27,56	122,84	58,19
10	27	1,75	9,92	3,06	98,34	17,35
5	8	- 3,25	- 9,08	10,56	82,51	29,52
7	32	- 1,25	14,92	1,56	222,51	- 18,65
22	40	13,75	22,92	189,06	525,17	315,10
5	13	- 3,25	- 4,08	10,56	16,67	13,27
13	21	4,75	3,92	22,56	15,34	18,60
2	4	- 6,25	- 13,08	39,06	171,17	81,77
18	28	9,75	10,92	95,06	119,17	106,44
4	8	- 4,25	- 9,08	18,06	82,51	38,60
9	15	0,75	- 2,08	0,56	4,34	- 1,56
8,25	17,08	-	0,00	470,25	1.658,92	760,75

$$Sx = \sqrt{\frac{\sum(x-Xm)^2}{n-1}} \quad Sx = \sqrt{\frac{470,25}{11}}$$

$$Sx = 6,54$$

$$Sy = \sqrt{\frac{\sum(y-ym)^2}{n-1}} \quad Sy = \sqrt{\frac{1.658,92}{11}}$$

$$Sy = 12,28$$

$$r = \frac{\sum(x-xm)(y-ym)}{(n-1)(Sx)(Sy)} \quad r = \frac{760,75}{11 \cdot 6,54 \cdot 12,28} \quad r = 0,86$$

Conclusión: relación entre la utilización de nuevo material y el número de defectos es elevada.

6.3 Regresión lineal

Ecuación de regresión, forma general

$$y = a + bx$$

Donde “y” va a depender de los valores que se le da a “x”

a = Intersección de la recta con el eje “y”, siendo “x” igual a cero.

b = pendiente de la recta

X = variable independiente

Y = variable dependiente

$$b = r \frac{Sy}{Sx} ; \quad a = Ym - bXm$$

Continuando con el ejercicio anterior:

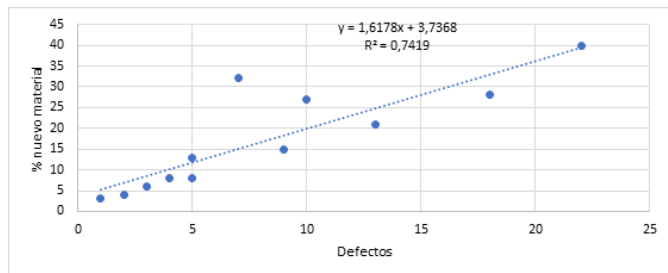
$$b = r \frac{Sy}{Sx} \quad b = 0,86 \frac{12,28}{6,54} \quad b = 1,618$$

$$a = Ym - bXm \quad a = 17,08 - 1,618(8,25) \quad a = 3,737$$

Aplicando estos resultados en la ecuación de la línea recta se tiene:

$$y = a + bx \qquad y = 3,737 + 1,618x$$

Gráfico de regresión lineal



Capítulo

7

Números Índice

7. Números Índice

El índice es una herramienta estadística útil que expresa el cambio relativo de un valor de un período a otro (Lind , Marchal, & Whaten, 2012, pág. 574). Compara una magnitud en dos situaciones distintas, una de las cuales se considera como referencia. Se trata de comparar períodos de tiempos distintos, situación geográfica, ingresos, o cualquier otra característica, expresado como porcentaje, que mide un periodo dado contra un periodo base determinado.

Un caso de estos es el **Índice de precios al consumidor IPC**, que es un indicador mensual o nacional que mide los cambios en el tiempo del nivel general de los precios, correspondientes al consumo final de bienes y servicios de los hogares de estratos de ingreso: alto, medio y bajo, residentes en el área urbana del país. La variable principal que se investiga es el precio.

Clasificación

1. Simples
2. Compuestos
 - 2.1 No ponderados
 - 2.2 Ponderados

7.1 Números índice simples

Expresa el cambio relativo de precio, cantidad o el valor en función de un **período base** (Lind , Marchal, & Whaten, 2012, pág. 574)

$$P = \frac{P_t}{P_o} \times 100$$

Ejemplos

1. El precio de un paquete de vacaciones de fin de semana

durante el verano (con alojamiento y todos los alimentos) en Montañita en el año 2013, fue de \$160. El precio aumentó a \$275 en 2016. ¿Cuál es el índice de precios de 2016 con el año 2013 como periodo base y 100 como valor base?:

$$P = \frac{P_t}{P_o} \times 100 \quad P = \frac{275}{160} \times 100 = 171,88$$

Interpretación:

El precio del paquete de fin de semana durante el verano aumentó 71.88% de 2013 a 2016.

2. Los precios de faros delanteros Manson, modelo G7, convertidos en índices con dos periodos bases distintos.

Año	Precio del faro	Índice de precios (2000=100)	Índice de precios (2005-2010=100)
2000	90	$\frac{90}{90} \times 100$ = 100	$\frac{90}{127,5} \times 100$ = 70,58
2005	120	$\frac{120}{90} \times 100$ = 133,33	$\frac{120}{127,5} \times 100$ = 94,12
2010	135	$\frac{135}{90} \times 100$ = 150	$\frac{135}{127,5} \times 100$ = 105,88
2015	147	$\frac{147}{90} \times 100$ = 163,33	$\frac{147}{127,5} \times 100$ = 115,29

Los números índice de 2015 con las dos bases distintas no son iguales. En los años 2000 y 2005 se reportan una disminución

de precios al contrario para los años 2010 y 2015 se reportan un incremento de los precios.

7.2 Índices no ponderados

Se combinan varios artículos u otras variables para la elaboración de un índice para comparar el costo de este agregado de artículos en dos periodos distintos para analizar su estudio económico. Este índice puede abarcar el costo de libros, colegiatura, alojamiento, alimentos y entretenimiento. Hay distintas formas de combinar los artículos para determinar un índice.

7.2.1 Promedio simple de los índices de precios

Ejemplo:

El 1 de enero se reportan los precios de varios artículos de alimentos de 2005 a 2015. Usted desea elaborar con ellos el índice de 2015, usando 2005 como base. Esto se expresa con el código abreviado 2005= 100. Debemos recalcar que 2005=100 es tan solo una notación convencional para especificar que 2005 es el año base y no una igualdad entre valores numéricos como tal. Inicie con el cálculo de un promedio simple de los índices de precios de cada artículo, emplee 2005 como año base y 2015 como año dado. El índice simple del pan es 220, que se determina con la fórmula. Calcule el índice simple de los demás artículos de manera similar.

$$P = \frac{P_t}{P_o} \times 100 \qquad P = \frac{110}{0.50} \times 100 = 220$$

Cálculo del índice de precios de alimentos

Artículo	Precio en 2005	Precio en 2015	Índice simple
Huevos (docena)	0,70	1,20	$\frac{1,20}{0,70} \times 100$ = 171,43
Pan (docena)	0,50	1,10	$\frac{1,10}{0,50} \times 100$ = 220
Leche (litro)	0,55	0,85	$\frac{0,85}{0,55} \times 100$ = 154,54
Queso (libra)	1,00	2,30	$\frac{2,30}{1,00} \times 100$ = 230
Arroz (libra)	0,30	0,50	$\frac{0,50}{0,30} \times 100$ = 166,67

El aumento de precio mayor afectó al queso en 130%, y el pan quedó en segundo lugar, con 120%.

Promedio simple de los precios relativos

$$P = \frac{\sum P_i}{n}$$

Donde P_i representa el índice simple de cada uno de los artículos, y n , el número de artículos.

Artículo	Precio en 2005	Precio en 2015	Índice simple
Huevos (docena)	0,70	1,20	$\frac{1,20}{0,70} \times 100$ = 171,43
Pan (docena)	0,50	1,10	$\frac{1,10}{0,50} \times 100$ = 220
Leche (litro)	0,55	0,85	$\frac{0,85}{0,55} \times 100$ = 154,54
Queso (libra)	1,00	2,30	$\frac{2,30}{1,00} \times 100$ = 230
Arroz (libra)	0,30	0,50	$\frac{0,50}{0,30} \times 100$ = 166,67
Total	3,05	5,95	942,64

$$\begin{aligned}
 P &= \frac{\sum P_i}{n} \\
 &= \frac{171,43 + 220 + 154,54 + 230 + 166,67}{5} \\
 &= 188,53
 \end{aligned}$$

Esto indica que la media del grupo de índices aumentó 88.53% de 2005 a 2015

7.2.2 Índice agregado simple

Consiste en considerar un índice simple agregados de magnitudes o precios con el que se calcula la razón de la media aritmética de los precios de n artículos en el periodo t como periodo

base.

$$P = \frac{\sum Pt}{\sum Po} \times 100$$

Se denomina índice agregado simple. El índice de los artículos de alimentos anteriores se determina al sumar los precios de 2005 y 2015. La suma de los precios del periodo base es \$3,05, y del periodo dado, \$5,95. El índice agregado simple es 195,08, lo que significa que el grupo de precios agregado aumentó 95,08% en el periodo de 10 años.

$$P = \frac{\sum Pt}{\sum Po} \times 100 = \frac{5,95}{3,05} \times 100 = 195,08$$

Ejercicios

1. En la siguiente tabla se reportan las ganancias de cada una de la empresa Supermaxi en años recientes. Desarrolle un índice, con 2003 como base, que muestre el cambio de las ganancias por acción durante el periodo.

Año	Ganancia en millones	Índice de precios (2003=100)
2002	234,56	$\frac{234,56}{345,93} \times 100 = 67,81$
2003	345,93	$\frac{345,93}{345,93} \times 100 = 100$
2004	567,43	$\frac{567,43}{345,93} \times 100 = 164,03$
2005	475,98	$\frac{475,98}{345,93} \times 100 = 137,59$
2006	378,78	$\frac{378,78}{345,93} \times 100 = 109,50$
2007	679,34	$\frac{679,34}{345,93} \times 100 = 196,38$
2008	789,45	$\frac{789,45}{345,93} \times 100 = 228,21$

Interpretación:

El precio de las ganancias de SUPERMAXI en el año 2008 aumento en 128,21% al contrario en el año 2002 disminuyo en 32,19%. Se concluye que cada año va aumentando sus ganancias considerablemente.

2. En enero de 2014, el precio de un par de zapatos Adidas fue \$95,50. En abril de 2016, el precio del mismo par de zapatos fue de \$130,69. Utilice el precio de enero de 2014 como periodo base y 100 como valor base para desarrollar un índice simple. ¿Qué porcentaje aumentó el precio del par de zapatos Adidas?

$$P = \frac{Pt}{Po} \times 100 = \frac{130,69}{95,50} \times 100 = 136,85$$

$$136,85 - 100 = 36,85 \%$$

El precio de los zapatos aumento en un 36,85% para el año de 2016

3. A continuación, se enumeran las ventas netas de Corporación Beltrán, minorista de ventas ubicadas en Ecuador durante los años de 2002 a 2007. En 2007, Beltrán se convirtió en una subsidiaria de Tejar. Su sitio en la red es www.beltran.com. Determine el índice de 2003 y (2006-2007). ¿En cuánto aumentaron las ventas netas desde el periodo base?

Años	Ventas (Millones)	Índice de precios (2003=100)	Índice de precios (2006-2007) =100
2002	437,56	$\frac{437,56}{568,76} \times 100$ = 76,93	$\frac{437,56}{482,83} \times 100$ = 90,62
2003	568,76	$\frac{568,76}{568,76} \times 100$ = 100	$\frac{568,76}{482,83} \times 100$ = 117,79
2004	457,65	$\frac{457,65}{568,76} \times 100$ = 80,46	$\frac{457,65}{482,83} \times 100$ = 94,78
2005	597,56	$\frac{597,56}{568,76} \times 100$ = 105,06	$\frac{597,56}{482,83} \times 100$ = 123,76
2006	532,12	$\frac{532,12}{568,76} \times 100$ = 93,55	$\frac{532,12}{482,83} \times 100$ = 110,21
2007	433,54	$\frac{433,54}{568,76} \times 100$ = 76,22	$\frac{433,54}{482,83} \times 100$ = 89,79

Interpretación

- Las ventas netas aumentaron en el año 2005 en un 5,06 %, con respecto al periodo base 2003.
 - Las ventas aumentaron 23,76% en el año de 2005, con respecto al periodo base (2006-2007).
4. En el año 2012, el precio de un paquete turístico a Galápagos fue \$650,50 con hospedaje y alimentación incluida para el año de 2016, el precio del mismo paquete turístico es de \$895,30. Utilice el precio de enero de 2012 como periodo base y 100 como valor base para desarrollar un índice simple. ¿Qué porcentaje aumentó el precio del paquete turístico?

$$P = \frac{P_t}{P_o} \times 100 = \frac{895,30}{650,50} \times 100 = 137,63$$

$$137,63 - 100 = 37,63 \%$$

El precio de los zapatos aumento en un 37,63% para el año de 2016

5. Cálculo del índice de precios de repuestos eléctricos para vehículos.

Repuestos (Unidades)	2012	2015	Índice simple
Halógenos	\$2,80	\$2,50	$\frac{2,50}{2,80} \times 100$ = 89,29
Pitos	5,20	5,60	$\frac{5,60}{5,20} \times 100$ = 107,69
Relays	3,00	2,70	$\frac{2,70}{3,00} \times 100$ = 90
Sokes	0,45	0,55	$\frac{0,55}{0,45} \times 100$ = 122,22
Fusibles	0,06	0,10	$\frac{0,10}{0,06} \times 100$ = 166,67
Bendix	8,00	8,50	$\frac{8,50}{8,00} \times 100$ = 106,25
Automáticos	12,00	12,80	$\frac{12,80}{12,00} \times 100$ = 106,67
Terminales	0,07	0,06	$\frac{0,06}{0,07} \times 100$ = 85,71
Total	31,58	32,81	874,5

El aumento de precio mayor afecto a los fusibles con, 66,67%, y los pitos quedó en segundo lugar, con 7,69%.

Promedio simple de los precios relativos

$$P = \frac{\sum Pi}{n} = \frac{89,29 + 107,69 + 90 + 122,22 + 166,67 + 106,25 + 106,75 + 85,71}{8} = 109,31$$

Esto indica que la media del grupo de índices aumentó 9,31% de 2012 a 2015

7.2.3 Índice agregado simple

El índice de los artículos de repuestos eléctricos se determina al sumar los precios de 2012 y 2015. La suma de los precios del periodo base es \$31,58, y del periodo dado, \$32,81. El índice agregado simple es 103,89, lo que significa que el grupo de precios agregado aumentó 3,89% en el periodo de 3 años.

$$P = \frac{\sum Pt}{\sum Po} \times 100 = \frac{32,81}{31,58} \times 100 = 103,89$$

7.3 Índices ponderados

Estos indicadores nos permiten observar cómo es la evolución del costo de vida de un país, a través de la recopilación de información para así procesarla y que, como resultado, los números índices ponderados obtenidos, nos muestren las variaciones de los diferentes productos o servicios, en un determinado periodo de tiempo.

Los métodos que más se utilizan para calcular el índice de precios ponderado son el de Laspeyres y el de Paasche. Sin embargo, existen otros métodos expresados por otros investigadores como Fisher y Edgeworth.

Las variables que intervienen son: P es el índice de precios; pt es el precio actual; po es el precio anterior o denominado precio en el periodo base; qo cantidad anterior o cantidad en el periodo base y finalmente qt cantidad actual.

7.3.1 Índice de Precios de Laspeyres:

En este método el período base considera precios (p_0) y cantidades (q_0) constantes; en los períodos analizados se mantiene la cantidad de productos del período base (q_0) y se considera los precios actualizados (p_t).

Este método es el más acogido para calcular los **índices de precios**, porque implica gastar menos recursos. Únicamente se tiene que actualizar los precios de los productos para el período analizado.

El Instituto Nacional de Estadísticas y Censos INEC hace uso de este método puesto no requiere de la actualización de muchos datos.

$$Lp = \frac{\sum_{i=1}^n p_t * q_0}{\sum_{i=1}^n p_0 * q_0} * 100$$

Ejemplo 1:

La madre de Paola decide salir a comprar algunos artículos de aseo con la misma cantidad de dinero que salía hace dos años atrás, es decir en el año 2015. El valor es de \$30 dólares, y lo que necesita saber es si puede comprar la misma cantidad de artículos con la misma cantidad de dinero.

Artículos	2015		2017		$p_0 * q_0$	$p_{2017} * q_0$	$p_{2017} * q_{2017}$	$p_{2017} * q_{2017}$
	Precio (p_0)	Cantidad (q_0)	Precio (p_t)	Cantidad (q_t)				
Desinfectante	0,80	5	1,20	3	4,00	2,40	6,00	3,60
Desinfectante	2,15	2	3,54	1	4,30	2,15	7,08	3,54
Jabón antibacterial	1,00	3	1,38	2	3,00	2,00	6,14	2,76
Toallita	4,75	1	5,80	1	4,75	4,75	5,80	5,80
Lava vajilla	1,25	2	2,30	1	2,50	1,25	6,00	2,30
Suma n					18,55	12,55	27,62	18,00

$$Lp = \frac{\sum_{i=1}^n p_t * q_0}{\sum_{i=1}^n p_0 * q_0} * 100$$

$$Lp = \frac{27,62}{18,55} * 100 = 148,89$$

Interpretación:

Según el índice de precios de Laspeyres los precios en el mercado han aumentado en 48,49% y la madre de Paola ya no puede comprar los mismos artículos con los US\$ 30,00

Ejemplo 2:

La tienda de barrio decide hacer una comparación entre la cantidad de productos vendidos y sus respectivos precios entre los años 2010 y 2017.

Artículos	2010		2017		pa ²⁰¹⁰ q ²⁰¹⁰	pa ²⁰¹⁷ q ²⁰¹⁰	pa ²⁰¹⁰ q ²⁰¹⁷	pa ²⁰¹⁷ q ²⁰¹⁷
	Precio (p ₀)	Cantidad (q ₀)	Precio (p ₁)	Cantidad (q ₁)				
Arroz (libra)	0,40	50	0,65	40	20,00	17,20	32,50	27,95
Cuente (pa)	0,60	65	1,00	48	39,00	28,80	65,00	48,00
Pan (mitigate)	1,00	32	2,70	29	32,00	29,00	86,40	78,30
Leche	0,60	28	0,85	22	16,80	13,20	23,80	18,70
Pollo (car)	0,30	80	0,50	50	24,00	15,00	40,00	25,00
Sumas					131,80	103,20	247,70	197,95

$$Lp = \frac{247,7}{131,8} * 100 = 187,94$$

Interpretación:

Los precios del año 2017 se han incrementado en 87,94% respecto al año 2010

7.3.2 Índice de precios de Paasche

Este método surge como alternativa a la fórmula planteada por Laspeyres, puesto que al parecer su teoría no es muy realista, pues se dice, que no puede ser que las cantidades de productos utilizadas en un año base sean las mismas para años posteriores, y que siempre habrá una variación por más mínima que sea.

Ante esta situación se origina la propuesta de Paasche, misma que dedujo que se utilizaría las cantidades del periodo actual; no obstante, se han presentado algunos inconvenientes, uno de los más relevantes, es el ponderar para cada año corriente, ya que es algo laborioso y lleva de mucho tiempo para tener los

resultados y poder ser analizado.

$$P_p = \frac{\sum_{i=1}^n p_t * q_t}{\sum_{i=1}^n p_0 * q_t} * 100$$

Mediante este índice ponderado de Paasche, se puede observar los cambios de precios dentro de un determinado periodo; es decir que nos permitirá conocer la situación actual, sea el caso de que un país quiere conocer el valor o el costo de la canasta básica, por lo que actualizado sus valores se obtendrá un porcentaje de incremento o disminución en comparación al año base.

Ejemplo 3

En base al ejemplo 1, determinar la variación de precios según Paasche.

Artículos	2015		2017		po*qp	po*qt	pt*qp	pt*qt
	Precio (po)	Cantidad (qp)	Precio (pt)	Cantidad (qt)				
Detergente	0,80	5	1,20	3	4,00	2,40	6,00	3,60
Desinfectante	2,15	2	3,54	1	4,30	2,15	7,08	3,54
jabon antibacterial	3,00	3	1,38	2	3,00	2,00	4,14	2,76
Toalla	4,75	1	5,80	1	4,75	4,75	5,80	5,80
Lava vajilla	1,25	2	2,30	1	2,50	1,25	4,60	2,30
Sumin					18,55	12,55	27,62	18,00

$$P_p = \frac{\sum_{i=1}^n p_t * q_t}{\sum_{i=1}^n p_0 * q_t} * 100$$

$$P_p = \frac{18,00}{12,55} * 100 = 143,43$$

Según Paasche, los precios del 2017 al 2015 variaron en 43,43%.

Comparando los índices de Laspeyres y Paasche en el ejemplo 1 y 3 se tiene que son 48,89% y 43,43% respectivamente.

Se observa una variación de 5,06%, que puede ser significativa según el caso en análisis, pero cuando se trata de calcular mensualmente y anualmente la variación de los **índices de precios**

de un país, se pone de manifiesto el gasto económico que representa.

Ejemplo 4:

En base al ejemplo 2, determinar la variación de precios según Paashe.

Artículos	2010		2017		po*qo	po*qt	pt*qo	pt*qt
	Precio (p0)	Cantidad (q0)	Precio (p1)	Cantidad (q1)				
Arroz (lbs)	0,40	50	0,65	43	20,00	17,20	32,50	27,95
Lenteja (lbs)	0,60	65	1,00	48	39,00	28,80	65,00	48,00
Pan integral	1,00	32	2,20	29	32,00	29,00	86,40	78,30
Leche	0,60	28	0,85	22	16,80	13,20	23,80	18,70
Anícar	0,30	80	0,50	50	24,00	15,00	40,00	25,00
Suma					131,80	103,20	247,70	197,95

$$P_p = \frac{\sum_{i=1}^n p_t * q_t}{\sum_{i=1}^n p_0 * q_t} * 100$$

$$P_p = \frac{197,95}{103,20} * 100 = 191,81$$

Según Paashe, los precios del 2017 al 2010 variaron en 91,81%.

7.3.3 Índice ideal de Fisher

Este índice trata de suavizar los valores de Laspeyres y de Paashe, calculado mediante la raíz cuadrada del producto de los mismos.

$$F_p = \sqrt{L_p * P_p}$$

Ejemplo 5:

Tomando los mismos valores de los ejemplos 1 y 3, determinar el índice según Fisher.

$$F_p = \sqrt{L_p * P_p}$$

$$F_p = \sqrt{(48,89 * 43,43)}$$

$$F_p = 46,07$$

En realidad, para calcular este índice, obligatoriamente se deben calcular los otros, lo que implica utilizar más recursos.

7.4 Índice de valores

Para el cálculo de este índice se toma en cuenta los precios y cantidades del año base, asimismo, los precios y cantidades del año actual, solo así se podrá observar los cambios o variaciones entre precios y cantidades de los dos años.

$$IV_0^t = \frac{V_t}{V_0} = \frac{\sum_{i=1}^n p_t * q_t}{\sum_{i=1}^n p_0 * q_0}$$

Verificación:

$$IV_0^t = L_{p_0}^t * P_{q_0}^t = L_{q_0}^t * P_{p_0}^t = F_{p_0}^t * F_{q_0}^t$$

Ejemplo 6:

Aplicando el ejemplo 1, determinar el Índice de Valores.

Artículos	2015		2017		po*qo	pt*qt	pt*qo	pt*qt
	Precio (ps)	Cantidad (qs)	Precio (pt)	Cantidad (qt)				
Detergente	0,80	5	1,20	3	4,00	2,40	6,00	3,60
Casinfecantes	2,55	2	3,54	1	4,10	2,15	7,08	3,54
Talco sin alcohol	1,00	3	1,38	2	3,00	2,00	4,14	2,76
Toallita	4,75	1	5,80	1	4,75	4,75	5,80	5,80
Sabon jabonilla	1,25	2	2,30	1	2,50	1,25	4,60	2,30
Sumas					18,55	12,55	27,62	18,00

$$Vp = \frac{\sum p_t * q_t}{\sum p_0 * q_0} * 100$$

$$Vp = \frac{18,00}{18,55} * 100$$

$$Vp = 97,03$$

Los precios y cantidades del año 2017 han decrecido en 3% respecto al año 2015.

7.5 Índices para propósitos especiales

El índice especial, también es denominado índice de satisfacción del Consumidor, ya que sirve para medir el grado de bienestar en relación del consumidor y factores como el producto, salario, entre otros. En la actualidad uno de los índices para propósitos especiales es la elaboración de índice de precios, ya que para una nación aporta con información muy importante y valledera, puesto genera un análisis por las personas a cargo, para

mediante esto calcular el costo de la canasta familiar y por ende hacer o no un ajuste a las remuneración, para que así puedan tener cobertura o por lo menos cubrir la canasta vital de los habitantes; durante el desarrollo de este índice se podrá evidencia un porcentaje de inflación, deduciendo la situación de un país, mediante la observación de los cambios dentro de un lapso de tiempo.

Dentro de los índices para propósitos especiales también se encuentra inmerso el índice de precios al productor que, como menciona se trata de recopilar información sobre precios en base a los productores.

Capítulo

R

Referencias Bibliográficas

Ditutor. (2015). [www.ditutor.com](http://www.ditutor.com/estadistica/medidas_posicion.html). Obtenido de http://www.ditutor.com/estadistica/medidas_posicion.html

Equipo de redacción, R. e. (2016). Tipos de gráficos estadísticos. Obtenido de http://www.mastiposde.com/graficos_estadisticos.html

Lind, D., Marchal, W., & Whaten, S. (2012). Estadística aplicada a los negocios y la economía. México, DF: Mc Graw Hill.

Lind, D., Marchal, W., & Wathen, S. (2012). Estadística aplicada a los negocios y la economía. México: McGRAW-HILL.

Spiegel, M., & Stephens, L. (2009). Estadística. México: Mc Graw Hill.

Webster, A. (2000). Estadística aplicada a los negocios y la economía. Santa Fe de Bogotá: Mc Graw Hill.

Capítulo

E

Evaluación de pares



UNIVERSIDAD CENTRAL DEL ECUADOR
FACULTAD DE CIENCIAS ADMINISTRATIVAS
DECANATO

Quito D.M., 12 de diciembre del 2017
Oficio No. DFA-920

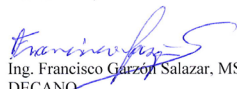
Señor Ingeniero
Renato Revelo Oña
DOCENTE DE LA FACULTAD DE CIENCIAS ADMINISTRATIVAS
Universidad Central del Ecuador
Presente

De mi consideración:

En atención a su pedido, adjunto al presente remito los informes originales de revisión de la obra "*Análisis estadístico básico para negocios*", de su autoría conjuntamente con otros profesionales, presentados por el Magister Celio Fabián Guayasamin Segovia, Docente de la Universidad de las Fuerzas Armadas-ESPE y Magister Jorge Eduardo Zapata Lara, Docente de la Universidad Politécnica Salesiana, en calidad de pares evaluadores.

Gracias por su atención.

Atentamente,


Ing. Francisco Garzon Salazar, MSc.
DECANO



Adjunto: 2 fojas útiles originales

Copia: Magister Leonardo Salvador Pérez; Magister Richard Jurado Zurita; Magister Jacqueline Ruiz Salas; Magister Mónica Andrade Salcedo, Docentes - FCA

Sandra S.



Quito, 04 de diciembre de 2017

Señor Magister
Francisco Garzón Salazar
DECANO
Facultad de Ciencias Administrativas de la Universidad Central del Ecuador
Presente,

De mi consideración:

Atendiendo el Oficio No. DFA-888, del 01 de diciembre de 2017, agradezco la invitación para ser **par evaluador** de la obra **Análisis estadístico básico para negocios**, de los autores: Mgt. Leonardo Patricio Salvador Pérez, Mgt. Renato Esteban Revelo Oña, Mgt. Richard Ramiro Jurado Zurita, Mgt. Jaqueline Yovana Ruiz Salas, Mgt. Mónica del Rocio Andrade Salcedo; comunico a usted que se ha revisado dicha obra con los criterios de relevancia y pertinencia que especifica el respectivo reglamento, por lo tanto, la misma es recomendable como una **OBRA RELEVANTE**.

Cabe indicar que los contenidos cumplen con estándares de calidad para los procesos de enseñanza y aprendizaje, es inédita y contribuyen al conocimiento y formación de los estudiantes universitarios, de tal manera que resultan fundamentales y sustanciales en la Educación Superior.

Informe que pongo a vuestra consideración para los fines pertinentes.

Atentamente,

Mgt. Celio Fabián Guayasamín Segovia

CC: 1705695870

DOCENTE DE LA UNIVERSIDAD DE LAS FUERZA ARMADAS ESPE

*Recibido 22/12-22/12
Señado
11-22*



Quito, 04 de diciembre de 2017

Señor Magister

Francisco Garzón Salazar

DECANO

Facultad de Ciencias Administrativas de la Universidad Central del Ecuador

Presente,

De mi consideración:

Atendiendo el Oficio No. DFA-889, del 01 de diciembre de 2017, agradezco la invitación para ser **par evaluador** de la obra **Análisis estadístico básico para negocios**, de los autores: Mgt. Leonardo Patricio Salvador Pérez, Mgt. Renato Esteban Revelo Oña, Mgt. Richard Ramiro Jurado Zurita, Mgt. Jaqueline Yovana Ruiz Salas, Mgt. Mónica del Rocio Andrade Salcedo; comunico a usted que se ha revisado dicha obra con los criterios de relevancia y pertinencia que especifica el respectivo reglamento, por lo tanto, la misma es recomendable como una **OBRA RELEVANTE**.

Cabe indicar que los contenidos cumplen con estándares de calidad para los procesos de enseñanza y aprendizaje, es inédita y contribuyen al conocimiento y formación de los estudiantes universitarios, de tal manera que resultan fundamentales y sustanciales en la Educación Superior.

Informe que pongo a vuestra consideración para los fines pertinentes.

Atentamente,

Jorge Eduardo Zapata Lara Msc. Mai.

CC: 1708066533

DOCENTE DE LA UNIVERSIDAD POLITÉCNICA SALESIANA

*Recibido 2017-12-12
Salazar
17:20*

1 **ra. Edición**

2017

www.coleccionempresarial.com

ISBN: 978-9942-30-471-1



9789942304711

Leonardo Salvador Pérez
Renato Revelo Oña
Ramiro Jurado Zurita
Jacqueline Ruiz Salas
Mónica Andrade Salcedo