

Apuntes de Estadística y Probabilidad

Leonardo H. Añez Vladimirovna¹

*Universidad Autónoma Gabriel René Moreno,
Facultad de Ingeniería en Ciencias de la Computación y Telecomunicaciones,
Santa Cruz de la Sierra, Bolivia*

30 de junio de 2018

¹Correo Electrónico: toborochi98@outlook.com

Notas del Autor

Estos apuntes fueron realizados durante mis clases en la materia MAT202 (Probabilidad y Estadística), acompañados de referencias de libros, fuentes y código que use a lo largo del curso, en el período I-2018 en la Facultad de Ingeniería en Ciencias de la Computación y Telecomunicaciones.

Para cualquier cambio, observación y/o sugerencia pueden enviarme un mensaje al siguiente correo:

`toborochi98@outlook.com`

Índice general

1. Aspectos Básicos de las Estadísticas	5
1.1. Conceptos Generales de la Estadística	5
1.1.1. Generalidades	5
1.1.2. Variables Estadísticas	5
1.2. Representación de la Información en Cuadros o Tablas	6
1.2.1. Conceptos Iniciales	6
1.2.2. Frecuencia	7
1.2.3. Tablas de Distribución de Frecuencia	7
1.3. Representaciones Gráficas	8
1.3.1. Diagrama de Barras	8
1.3.2. Diagrama de Sectores	8
1.3.3. Diagrama de Bastones	9
1.3.4. Histograma y Polígono de Frecuencia	9
2. Estadígrafos de Tendencia Central	11
2.1. Media Aritmetica Simple (\bar{X})	11
2.1.1. Cálculo	11
2.1.2. Propiedades	11
2.1.3. Métodos Abreviados para el Cálculo de la Media	12
2.1.4. Ventajas y Desventajas	12
2.1.5. Ejemplos	12
2.2. Media Aritmetica Ponderada (\bar{X}_p)	12
2.2.1. Cálculo	13
2.2.2. Ejemplos	13
2.3. Mediana (M_e)	13
2.3.1. Cálculo	13
2.3.2. Ejemplos	14
2.4. Cuantiles	14
2.4.1. Cuartiles (Q_i)	14
2.4.2. Deciles (D_i)	14
2.4.3. Percentiles (P_i)	14
2.4.4. Cálculo de Cuantiles	14
2.4.5. Ejemplos	15
2.5. Moda o Modo (M_o)	15
2.5.1. Cálculo	15
2.5.2. Ventajas y Desventajas	15
2.6. Media Geométrica (M_g)	16
2.6.1. Cálculo	16
2.6.2. Propiedades y Usos	16
2.6.3. Ejemplos	17
2.7. Media Armónica (M_h)	17
2.7.1. Cálculo	17
2.7.2. Usos de la Media Armónica	17
2.7.3. Ventajas y Desventajas	17
2.7.4. Ejemplos	17
2.8. Media Cuadrática (\bar{X}_c)	17
2.8.1. Cálculo	17
2.8.2. Ejemplos	18

3. Estadígrafos de Dispersión	19
3.1. Recorridos	19
3.1.1. Recorrido (R)	19
3.1.2. Recorrido Intercuartílico (R_I)	19
3.2. Desviación Media (D_M)	19
3.3. Varianza (S^2, \hat{S})	19
3.3.1. Cálculo	19
3.3.2. Propiedades	20
3.3.3. Componentes de la Varianza	20
3.3.4. Métodos Abreviados para el Cálculo de la Varianza	20
3.3.5. Ejemplos	21
3.4. Desviación Estándar	21
3.4.1. Propiedades	21
3.4.2. Ejemplos	21
3.5. Medidas de Dispersión Relativa	21
3.5.1. Coeficiente de Variación (C_v)	22
3.5.2. Momentos	22
3.6. Medidas de Forma	22
3.6.1. Coeficientes de Asimetría	22
3.6.2. Curtosis o Apuntalamiento	23
3.7. Medidas de Concentración	23
3.7.1. Índice de Gini (G)	23
3.7.2. Curva de Lorenz	24
3.8. Distribución Bivariante	24
3.8.1. Distribución de Frecuencias Absolutas Conjuntas	24
3.8.2. Distribución de Frecuencias Relativas Conjuntas	25
3.8.3. Distribución de Frecuencias Absolutas Acumuladas Conjuntas	25
3.8.4. Distribución de Frecuencias Relativa Acumuladas Conjuntas	25
3.8.5. Distribuciones Marginales	25
3.9. Covarianza (S_{xy})	26
3.9.1. Cálculo	26
3.9.2. Propiedades	26
3.9.3. Interpretación	26
4. Análisis de Regresión y Correlación	27
4.1. Modelo de Regresión Lineal Simple	27
4.2. Regresión Lineal	27
4.3. Regresión Potencial	28
4.4. Regresión Exponencial	28
4.5. Regresión Parabólica	28
4.6. Coeficiente de Correlación Rectilíneo	28
4.6.1. Propiedades	28
4.6.2. Cálculo	29
5. Probabilidad	31
5.1. Eventos	31
5.1.1. Operaciones con Eventos	32
5.2. σ -álgebra de Eventos	32
5.3. Definición Axiomática de Probabilidad	33
5.3.1. Propiedades de la Probabilidad de Eventos	33
5.4. Definición Clásica de Probabilidad	33
5.4.1. Probabilidad Condicionada	34
5.4.2. Eventos Dependientes	34
5.4.3. Eventos Independientes	34
5.4.4. Teorema de la Probabilidad Total	36
5.4.5. Teorema de Bayes	36
5.4.6. Ejemplos	37

Capítulo 1

Aspectos Básicos de las Estadísticas

Concepto de la Estadística

Es una ciencia que proporciona un conjunto de métodos y técnicas que permiten recolectar, clasificar, organizar, analizar y presentar datos de forma adecuada para tomar decisiones frente a una incertidumbre o describir o afirmar algo acerca de la población. Esta definición permite clasificar la Estadística en: Estadística **Descriptiva** y Estadística **Inferencial**.

Estadística Descriptiva

Es la parte de la Estadística que proporciona un conjunto de métodos, con el fin de describir o analizar datos.

Estadística Inferencial

1.1. Conceptos Generales de la Estadística

1.1.1. Generalidades

- (I) **Población:** Es el conjunto o totalidad de elementos, sean personas, objetos o cosas que presentan uno o más características que nos interesan acerca de las cuales intentamos obtener conclusiones.
- (II) **Tamaño de la Población:** Número de elementos de la población, denotado usualmente por: N .
- (III) **Muestra:** Es parte o subconjunto de la población.
- (IV) **Tamaño de la Muestra:** Cantidad de elementos de la Muestra. Denotado por: n .
- (V) **Parámetro:** Es una medida numérica que describe una característica de la población. Los parámetros mas usados de la población son:
 - Media Poblacional (μ)*
 - Varianza Poblacional (σ^2)*
 - Desviación Estándar Poblacional (σ)*
 - Proporción Poblacional (π ó p)*
- (VI) **Estadígrafos:** Es una medida numérica que describe una característica de la muestra. Los estadígrafos mas usados son:
 - Media Muestral (\bar{x}, \bar{y})*
 - Varianza Muestral (S^2, \hat{S}^2)*
 - Desviación Estándar (S, \hat{S})*
 - Proporción Muestral ($\hat{\pi}$ ó \hat{p})*

1.1.2. Variables Estadísticas

Se denomina variable estadística a la característica o combinación de características de una población:

Todos los puntos marcados con un * serán explicados mas adelante, o puedes ver el índice

Tipos de Variables Estadísticas

Están clasificadas de la siguiente manera:

■ Cualitativas:

- **Variables Nominales:** Es aquella cuyo valor se expresa en categorías sin ningún tipo de orden o clasificación.
- **Variables Ordinales:** Es aquella cuyo valor se expresa en categorías pero se busca una clasificación de orden.

■ Cuantitativas:

- **Variables Discretas:** Se dice cuando es resultado de la operación de contar.
- **Variables Continuas:** Una variable es continua cuando se encuentra por medición o comparación, y es expresada por cualquier número real.

Nominales	Ordinales	Discretas	Continuas
Sexo	Tamaño	Número de Hijos	Edad
Nacionalidad	Grado de Instrucción	Número de Alumnos	Estatura
Color	Categoría	Número de Accidentes	Peso

Figura 1.1: Tabla con algunos ejemplos de los tipos de Variables

1.2. Representación de la Información en Cuadros o Tablas

Para entender como hacer tablas Estadísticas y Representaciones Gráficas, hay ciertos conceptos que es necesario conocer antes:

1.2.1. Conceptos Iniciales

- (I) **Clase:** Es el conjunto de observaciones cualitativas o cuantitativamente iguales o muy próximas.
- (II) **Límite de Clase:** Son los valores de las observaciones, mínimo y máximo que pueden tener los datos, incluidos en la *Clase*.
- (III) **Intervalo de Clase:** Es la distancia entre los límites de la *Clase* expresados de la siguiente manera:

$$y'_{i-1}y'_i$$

- (IV) **Marca de Clase:** Es el punto medio de cada uno de los intervalos. Si es constante se puede expresar como c_j , caso contrario y_i . Se tiene la siguiente fórmula para calcularlo:

$$c_j = \frac{y'_{i-1} + y'_i}{2}$$

- (V) **Recorrido de la Variable:** Es el valor de la diferencia entre los límites de los datos de la muestra (el máximo y el mínimo), está representado por: R .

- **Datos No Agrupados¹:** $R = y_k - y_i$
- **Datos Agrupados:** $R = y'_k - y'_i$

- (VI) **Amplitud de Intervalo:** Expresado por C y C_i , para cuando es constante y variable respectivamente.

$$C_i = y'_i - y'_{i-1}$$

- (VII) **Número de Intervalos:**

$$k = \frac{R}{C_i}$$

¹También llamados **Datos Originales**

1.2.2. Frecuencia

Es la cantidad de Observaciones incluidas en una clase.

Frecuencias

(I) **Frecuencia Absoluta Simple** (f_i): es el numero de veces que se repite cada observación.

(II) **Frecuencia Absoluta Acumulada** (F_i): Expresado como:

$$F_i = \sum_{j=1}^i f_j$$

(III) **Frecuencia Relativa Simple** (h_i):

$$h_j = \frac{f_j}{n}$$

(IV) **Frecuencia Relativa Acumulada** (H_i):

$$H_i = \frac{F_i}{n} = \sum_{j=1}^i h_j$$

Propiedades de la Frecuencias

- $f_1 + f_2 + f_3 + \dots + f_k = \sum_{i=1}^k f_i = n$
- $h_1 + h_2 + h_3 + \dots + h_k = \sum_{i=1}^k h_i = 1$
- $\forall i \in I, \begin{cases} 0 \leq h_i \leq 1 \\ 0 \leq f_i \leq n \end{cases}$
- $f_1 = F_1 \leq F_2 \leq F_3 \leq \dots \leq F_k = n$
- $h_1 = H_1 \leq H_2 \leq H_3 \leq \dots \leq H_k = 1$
- $H_j = H_{j-1} + h_j$
- $F_j = F_{j-1} + f_j$

1.2.3. Tablas de Distribución de Frecuencia

Distribución de Frecuencia para Datos Agrupados no en Intervalos de Clase

x_i	f_i	F_i	h_i	H_i
x_1	f_1	F_1	h_1	H_1
x_2	f_2	F_2	h_2	H_2
x_3	f_3	F_3	h_3	H_3
\vdots	\vdots	\vdots	\vdots	\vdots
x_k	f_k	$F_k = n$	h_k	$H_k = 1$
Totales	n		1	

Distribución de Frecuencias para Datos Agrupados en Intervalo de Clase

$y'_{i-1} - y'_i$	f_i	F_i	h_i	H_i
$y'_0 - y'_1$	f_1	F_1	h_1	H_1
$y'_1 - y'_2$	f_2	F_2	h_2	H_2
$y'_2 - y'_3$	f_3	F_3	h_3	H_3
\vdots	\vdots	\vdots	\vdots	\vdots
$y'_{k-1} - y'_k$	f_k	$F_k = n$	h_k	$H_k = 1$
Totales	n		1	

Estructura de una Tabla Estadística**1. Título****2. Cuerpo Propiamente**

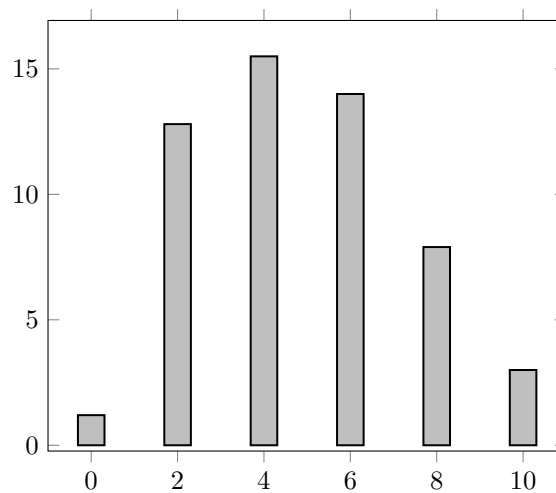
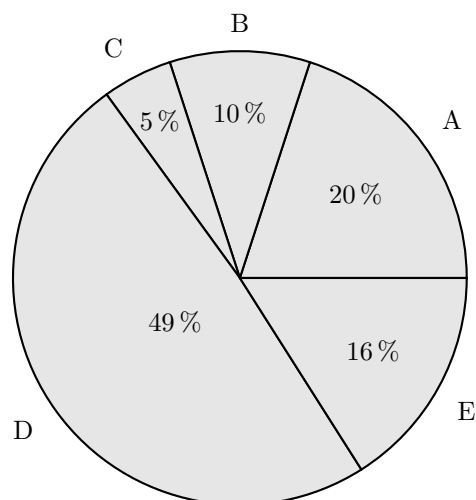
- Encabezamiento
- Cuerpo
- Columna Matriz

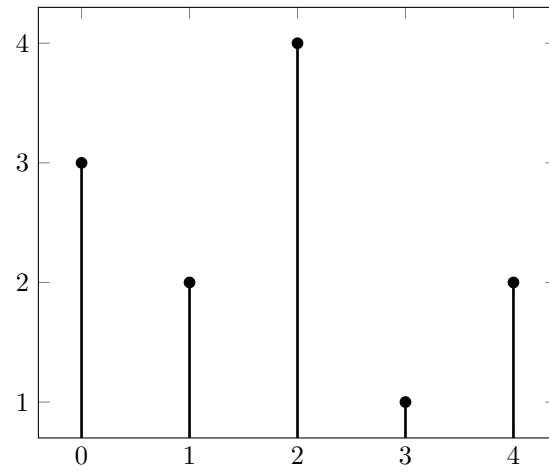
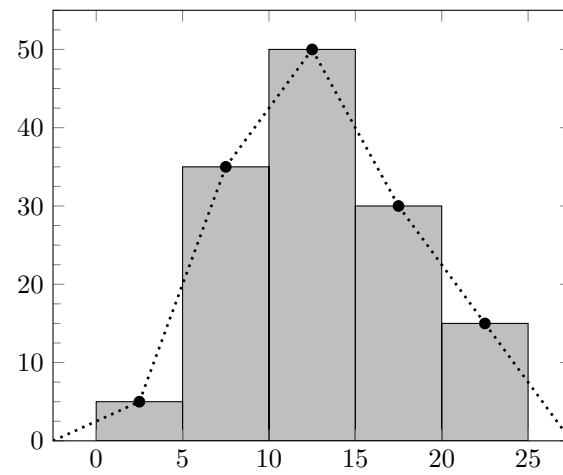
3. Indicadores Complementarios

- Fuente(s)
- Nota(s)
- Comentarios

1.3. Representaciones Gráficas

Las gráficas son instrumentos estadísticos muy útiles para dar idea global acerca de una situación estadística:

1.3.1. Diagrama de Barras**1.3.2. Diagrama de Sectores**

1.3.3. Diagrama de Bastones**1.3.4. Histograma y Polígono de Frecuencia**

Capítulo 2

Estadígrafos de Tendencia Central

Una *Medida de Tendencia Central* es un índice de localización central, empleado en la descripción de la distribución de frecuencias. Sirve también como base para medir y evaluar datos anormalmente altos o anormalmente bajos, llamados *Valores Extremos*.

Características del Valor Central

1. Debe estar definida en forma objetiva.
2. Debe depender de todas la observación obtenida en lo posible.
3. Debe ser fácil de comprender e interpretar.
4. Debe ser fácil de calcular.
5. Debe ser estable.
6. Debe ser adecuado a cálculos algebraicos posteriores.

2.1. Media Aritmetica Simple (\bar{X})

2.1.1. Cálculo

Para Datos no Agrupados (Originales)

Sean $x_1, x_2, x_3, \dots, x_n$ valores de la variable X . La media Arítmetica simple se denota por $M(x), \theta$ o \bar{X} está dada por:

$$M(x) = \bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Para Agrupados

$$M(y) = \bar{Y} = \frac{\sum_{i=1}^n y_i \cdot f_i}{n}$$

2.1.2. Propiedades

1. En todas las distribuciones de frecuencias, la sumatoria simple de las desviaciones de los valores de la variable respecto a la media aritmética es 0.

Datos Originales

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

Datos Agrupados

$$\sum_{i=1}^n (y_i - \bar{y}) f_i = 0$$

2. En toda distribución, la sumatoria de los cuadrados de las desviaciones de los valores de las variables, respecto de la media es mínima

Datos Originales

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - B)^2$$

Para cualquier constante B .

3. $M(ax_i) = a\bar{x}_i = a \cdot M(x_i) = a\bar{x}$, a constante
4. $M(a \pm x_i) = \overline{a \pm x_i} = a \pm M(x_i) = a \pm \bar{x}$, a constante
5. $M(a) = \bar{a} = a$, a constante

Datos Agrupados

$$\sum_{i=1}^n (y_i - \bar{y})^2 \leq \sum_{i=1}^n (y_i - B)^2 \cdot f_i$$

2.1.3. Métodos Abreviados para el Cálculo de la Media**Primer Método Abreviado**

$$\bar{y} = O_t + \frac{\sum_{i=1}^n z'_i \cdot f_i}{n}$$

Donde: $z'_i = y_i - O_t$

Segundo Método Abreviado

$$\bar{y} = O_t + C \cdot \frac{\sum_{i=1}^n z''_i \cdot f_i}{n}$$

Donde: $z''_i = \frac{y_i - O_t}{C} = \frac{z'_i}{C}$

Y para ambos métodos: O_t es el origen del trabajo (cualquier valor de la marca de clase).

Media Aritmética en Distribuciones Simétricas

Si k es impar

$$\bar{y} = y_{\frac{k+1}{2}}$$

Si k es par

$$\bar{y} = y'_{\frac{k}{2}}$$

2.1.4. Ventajas y Desventajas**Ventajas**

- Es la medida de tendencia central mas usada.
- El promedio es estable en el muestreo.
- Puede ser usado como un detector de variación en los datos.
- Se emplea en cálculos estadísticos posteriores.

Desventajas

- Es sensible a los valores extremos.
- Si el conjunto de datos es muy grande puede ser tedioso su cálculo.
- No se puede calcular para datos cualitativos.
- No se puede calcular para datos que tengan clases de extremos abiertos, tanto inferior o superior.

2.1.5. Ejemplos**2.2. Media Aritmetica Ponderada (\bar{X}_p)**

Permite calcular un promedio que toma en cuenta la importancia o peso relativo de cada valor sobre el total.

2.2.1. Cálculo

Si $w_1, w_2, w_3, \dots, w_n$ son los pesos asociados a los valores de X .

$$X : x_1, x_2, x_3, \dots, x_n$$

La Media Aritmética Ponderada está dada por:

$$\bar{X}_p = \frac{\sum_{i=1}^n x_i \cdot w_i}{\sum_{i=1}^n w_i}$$

Donde:

- x_i : Valor de las Observaciones Individuales.
- w_i : Peso asignado en cada observación.

2.2.2. Ejemplos

2.3. Mediana (M_e)

Es el valor de la variable que supera a no mas de la mitad de las observaciones y es superado por no mas de la mitad de las observaciones.

2.3.1. Cálculo

Datos no Agrupados

En el cálculo de la mediana para datos no agrupados hay que distinguir 2 observaciones:

- **Número de Observaciones Impares:**

Si

$$x_1, x_2, x_3, \dots, x_n$$

Son los valores de las variables ordenados creciente o decreciente entonces la mediana es el valor de la variable situado en el centro del conjunto de variables, esto es:

$$Me = x_{\frac{n+1}{2}}$$

- **Número de Observaciones Par:** En este caso esta definido por:

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Datos Agrupados

- **Agrupados no en Intervalo de Clase:** En este caso bastará con identificar la frecuencia absoluta acumulada inmediatamente mayor a la mitad de las observaciones. La *Mediana* es aquel valor de la variable que corresponde a dicha frecuencia acumulada.
- **Agrupados en Intervalo de Clase:**

$$Me = y'_{j-1} + C_j \cdot \frac{\frac{n}{2} - F_{j-1}}{f_j}$$

Mediana en Distribuciones Simétricas

Si k es impar

$$M_e = y_{\frac{k+1}{2}}$$

Si k es par

$$M_e = y'_{\frac{k}{2}}$$

2.3.2. Ejemplos

2.4. Cuantiles

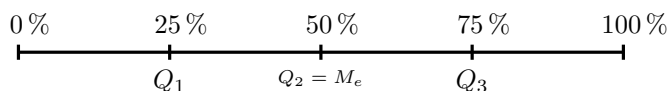
Como consecuencia del estudio de la mediana, surgen otros estadígrafos que dividen a las observaciones en otras proporciones y no solo en mitades como lo hace la media, los estadígrafos mas frecuentes de este tipo en el análisis estadístico son:

- Cuantiles (Q_i)
- Deciles (D_i)
- Percentiles (P_i)

Los cuantiles se emplean para describir el comportamiento de una población y sus resultados se expresan en tanto por ciento.

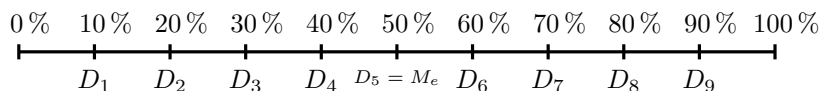
2.4.1. Cuantiles (Q_i)

Son valores que dividen a un conjunto de datos ordenados de forma ascendente o descendente en cuatro partes iguales:



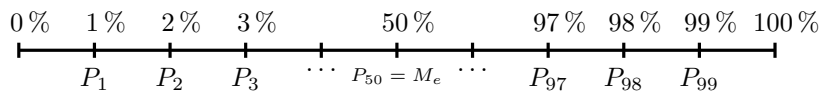
2.4.2. Deciles (D_i)

Son valores que dividen a un conjunto de datos ordenados de forma ascendente o descendente en diez partes iguales:



2.4.3. Percentiles (P_i)

Son valores que dividen a un conjunto de datos ordenados de forma ascendente o descendente en cien partes iguales:



2.4.4. Cálculo de Cuantiles

Para datos Originales

- Cuantiles:

$$Q_i = x_{\frac{i(n+1)}{4}}$$

- Deciles:

$$D_i = x_{\frac{i(n+1)}{10}}$$

- Percentiles:

$$P_i = x_{\frac{i(n+1)}{100}}$$

Para datos Agrupados

■ Cuartiles:

$$Q_i = y'_{j-1} + C_j \frac{i \frac{n}{4} - F_{j-1}}{f_j}$$

■ Deciles:

$$D_i = y'_{j-1} + C_j \frac{i \frac{n}{10} - F_{j-1}}{f_j}$$

■ Percentiles:

$$P_i = y'_{j-1} + C_j \frac{i \frac{n}{100} - F_{j-1}}{f_j}$$

2.4.5. Ejemplos**2.5. Moda o Modo (M_o)**

La Moda de una Muestra:

$$x_1, x_2, x_3, \dots, x_n$$

Es aquel valor de la variable que representa la mayor frecuencia, es decir, el valor de la variable que mas veces se repite. En términos sociológicos la moda es aquella de mayor aceptación popular, la moda es usada generalmente para representar y describir una tendencia en la economía.

Observación

La moda no siempre existe y no siempre es única. Como ejemplo podemos tener los siguientes conjuntos:

1. $\{2, 4, 5, 6, 3, 7, 3, 8, 3\}$
2. $\{5, 7, 4, 5, 8, 4, 5, 9, 4\}$
3. $\{1, 2, 3, 4, 5, 6, 7\}$

Para el primer conjunto, la moda es $M_o = 3$ ya que es el dato que mas se repite de las observaciones. Del segundo ejemplo, la moda es $M_o = 4$ y $M_o = 5$, ya que ambos datos son los que mas se repiten. Finalmente, en el tercer ejemplo, no hay moda.

2.5.1. Cálculo**En Intervalo de Clase**

■ De Amplitud Constante:

$$M_o = y'_{j-1} + C_j \cdot \frac{f_j - f_{j-1}}{(f_j - f_{j-1}) + (f_j - f_{j+1})}$$

■ De Amplitud Variable:

$$M_o = y'_{j-1} + C_j \cdot \frac{\frac{f_{j+1}}{a_{j+1}}}{\frac{f_{j+1}}{a_{j+1}} + \frac{f_{j-1}}{a_{j-1}}}$$

esto es:

$$M_o = y'_{j-1} + C_j \cdot \frac{h_{j+1}}{h_{j+1} + h_{j-1}}$$

2.5.2. Ventajas y Desventajas**Ventajas**

- El valor de la moda es totalmente independiente de los valores extremos.
- La moda se puede utilizar como una localización para datos cualitativos como cuantitativos.

- La moda es posible calcular aun en distribuciones que tienen intervalo de clase abierto.

Desventajas

- Es una medida inestable porque varía cuando se combina el intervalo de clase.
- Su significación es limitada, cuando no se dispone de una gran número de valores.
- La moda no se presta a manipulaciones algebraicas posteriores.
- Cuando el conjunto de observaciones contiene mas de una moda es difícil interpretar y de comprender.

2.6. Media Geométrica (M_g)**2.6.1. Cálculo****Para datos no Agrupados**

$$M_g = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$$

En notación:

$$M_g = \sqrt[n]{\prod_{i=1}^n x_i}$$

Usando logaritmos:

$$\log(M_g) = \log \left(\sqrt[n]{\prod_{i=1}^n x_i} \right)$$

$$\log(M_g) = \frac{1}{n} \log \left(\prod_{i=1}^n x_i \right)$$

Aplicando propiedades de logaritmos:

$$M_g = \text{antilog} \left[\frac{1}{n} \sum_{i=1}^n \log(x_i) \right]$$

Para datos Agrupados

$$M_g = \sqrt[n]{y_1^{f_1} \cdot y_2^{f_2} \cdot y_3^{f_3} \cdot \dots \cdot y_n^{f_n}}$$

En notación:

$$M_g = \sqrt[n]{\prod_{i=1}^k y_i^{f_i}}$$

Usando logaritmos¹

$$M_g = \text{antilog} \left[\frac{1}{n} \sum_{i=1}^k f_i \log(y_i) \right]$$

2.6.2. Propiedades y Usos

1. Los promedios geométricos dan menos importancia a las desviaciones extremas que los promedios aritméticos.
2. No se puede deducir, cuando la serie presenta un valor 0 o valores negativos.
3. Es apropiado cuando hay que promediar razones y proporciones.
4. La aplicación mas útil es para promediar tasas de cambio.

Cuando el cálculo de la tasa media de cambio abarca un número considerable de años se lo hace usando la siguiente formula:

$$P_n = P_o(1 + i)^n$$

Donde:

¹El procedimiento es el mismo que para datos Originales

- P_o : Cantidad al Principio del periodo.
- P_n : Cantidad al final del periodo.
- i : Tasa de cambio expresada en tanto por uno.
- n : Número de periodos.

2.6.3. Ejemplos

2.7. Media Armónica (M_h)

Es la media aritmética recíproca de los recíprocos de las variables:

2.7.1. Cálculo

Para datos Originales

$$M_h = \frac{1}{\frac{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}}{n}} = \frac{1}{\frac{\sum_{i=1}^n \frac{1}{x_i}}{n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Para datos Agrupados

$$M_h = \frac{1}{\frac{\frac{f_1}{y_1} + \frac{f_2}{y_2} + \frac{f_3}{y_3} + \dots + \frac{f_k}{y_k}}{n}} = \frac{1}{\frac{\sum_{i=1}^k \frac{f_i}{y_i}}{n}} = \frac{n}{\sum_{i=1}^k \frac{f_i}{y_i}}$$

2.7.2. Usos de la Media Armónica

1. Es apropiado para promediar velocidades.
2. En economía se usa en los índices de precio.
3. Se usa para promediar precio cuando la cantidad es variable y el monto gastado igual.

2.7.3. Ventajas y Desventajas

Ventajas

- Esta basada en todas las observaciones.
- Es capaz de un tratamiento algebraico.
- No se ve significativamente afectado por la fluctuación del muestreo.
- Es un promedio apropiado para promediar tasas.
- No le da mucho peso a los artículos grandes.

Desventajas

- No se puede calcular si alguno de los elementos es cero.
- En ciertas situaciones es difícil de calcular.
- Por lo general, es un valor que no existe en los datos proporcionados.

2.7.4. Ejemplos

2.8. Media Cuadrática ($\overline{X_c}$)

2.8.1. Cálculo

Datos Originales

$$\overline{X_c} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$$

Datos Agrupados

$$\overline{Y}_c = \sqrt{\frac{\sum_{i=1}^k y_i^2 \cdot f_i}{n}}$$

Relación entre los Promedios Computados

$$M_h \leq M_g \leq \bar{x} \leq \bar{x}_c$$

2.8.2. Ejemplos

Capítulo 3

Estadígrafos de Dispersión

Son los que cuantifican el grado de dispersión de los valores de la variable, en torno a un promedio o valor central de la distribución. Los estadígrafos de dispersión se necesitan para dos propósitos básicos:

1. Para verificar la confiabilidad de los promedios.
2. Para que sirva como base para el control de la variación.

3.1. Recorridos

3.1.1. Recorrido (R)

Datos Originales

$$R = x_{\max} - x_{\min}$$

Datos Agrupados

$$R = y_k - y_0$$

3.1.2. Recorrido Intercuartílico (R_I)

$$R_I = Q_3 - Q_1$$

3.2. Desviación Media (D_M)

Datos Originales

$$D_M = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Datos Agrupados

$$D_M = \frac{\sum_{i=1}^k |y_i - \bar{y}| \cdot f_i}{n}$$

3.3. Varianza (S^2, \hat{S})

3.3.1. Cálculo

Datos Originales

- Muestras Grandes:

$$\hat{S} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad n \geq 30$$

- Muestras Pequeñas:

$$\hat{S} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad n < 30$$

3.3.2. Propiedades

1. $\hat{S} \geq 0$
2. $V(a) = 0$, a constante.
3. $V(x \pm a) = V(x)$
4. $V(ax) = a^2 V(x)$
5. $V(ax + b) = a^2 V(x)$, b constante.

Observaciones

- La diferencia entre S^2 y \hat{S} es grande para muestras pequeñas y es mínima para muestras grandes, casi son iguales.
- La Varianza siempre viene expresada en unidades al cuadrado.

3.3.3. Componentes de la Varianza

Si un conjunto de datos se divide en subconjuntos por categoría o estratos, es posible descomponer la Varianza en dos componentes.

Intervarianza

Es un estadígrafo que representa la variabilidad entre los estratos y se define como la varianza de la media de los estratos:

$$S_b^2 = V(\bar{y}_h) = \frac{\sum_{h=1}^L (\bar{y}_h - \bar{y})^2 \cdot f_h}{n}$$

Donde:

- \bar{y}_h : Media de los estratos
- \bar{y} : Media General
- f_h : Frecuencia
- L : Número de Estratos

Intravarianza

Presenta la variabilidad dentro de los estratos definida como la media de la varianza de los estratos:

$$S_w^2 = M(S_h^2) = \frac{\sum_{h=1}^L S_h^2 \cdot f_h}{n}$$

- S_h^2 : Varianza de los estratos.

Como ambos componen la Varianza, se cumple la siguiente igualdad:

$$S^2 = S_b^2 + S_w^2$$

3.3.4. Métodos Abreviados para el Cálculo de la Varianza

Primer Método Abreviado

Se sabe que:

$$\hat{S} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Si desarrollamos el Binomio y distribuimos la sumatoria:

$$\hat{S} = \frac{\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n} = \frac{\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2}{n}$$

Separamos el denominador:

$$\hat{S} = \frac{\sum_{i=1}^n x_i^2}{n} - 2\bar{x} \cdot \frac{\sum_{i=1}^n x_i}{n} + \frac{n\bar{x}^2}{n} \quad \text{Recordemos que: } \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

Reemplazamos y simplificamos:

$$\hat{S} = \frac{\sum_{i=1}^n x_i^2}{n} - 2\bar{x} \cdot \bar{x} + \bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n} - 2\bar{x}^2 + \bar{x}^2$$

Finalmente tendremos:

$$\hat{S} = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$$

Segundo Método Abreviado

$$\hat{S} = \frac{\sum_{i=1}^k z_i'^2 f_i}{n} - \frac{\left(\sum_{i=1}^k z_i' f_i \right)^2}{n^2}$$

Tercer Método Abreviado

$$\hat{S} = C^2 \left[\frac{\sum_{i=1}^k z_i''^2 f_i}{n} - \frac{\left(\sum_{i=1}^k z_i'' f_i \right)^2}{n^2} \right]$$

3.3.5. Ejemplos

3.4. Desviación Estándar

Es la raíz cuadrada positiva de la varianza:

$$D(x) = \sigma = \sqrt{V(x)}$$

3.4.1. Propiedades

1. $D(x) \geq 0$
2. $D(a) = 0$, a constante.
3. $D(x \pm a) = D(x)$

3.4.2. Ejemplos

3.5. Medidas de Dispersión Relativa

Se utilizan para comparar dos o mas distribuciones en torno a la media.

3.5.1. Coeficiente de Variación (C_v)

Es una medida de dispersión relativa muy útil cuando:

- Los datos están en unidades diferentes.
- Los datos están en las mismas unidades, pero las medias muy distantes.

Cálculo del Coeficiente de Variación

$$C_v = \frac{S}{\bar{x}} \quad \vee \quad C_v = \frac{S}{\bar{x}} 100 \quad \text{Tanto Porciento}$$

3.5.2. Momentos

Son de uso frecuente los promedios de las series de potencias de las variables estos promedios reciben el nombre de *Momentos*. Pueden definirse respecto a cualquier punto:

- Momentos Respecto a un punto
- Momentos Respecto del Origen
- Momentos Respecto a la Media o Momentos Centrales

Momentos Respecto a un Punto ($M_{r,A}$)

Datos Originales

$$M_{r,A} = \frac{\sum_{i=1}^n (x_i - A)^r}{n}$$

Datos Agrupados

$$M_{r,A} = \frac{\sum_{i=1}^n (y_i - A)^r \cdot f_i}{n}$$

Momentos Respecto del Origen (M'_r)

Datos Originales

$$M'_r = \frac{\sum_{i=1}^n (x_i)^r}{n}$$

Datos Agrupados

$$M'_r = \frac{\sum_{i=1}^n (y_i)^r \cdot f_i}{n}$$

Momentos Respecto a la Media (M_r)

Datos Originales

$$M_r = \frac{\sum_{i=1}^n (x_i - \bar{x})^r}{n}$$

Datos Agrupados

$$M_r = \frac{\sum_{i=1}^n (y_i - \bar{y})^r \cdot f_i}{n}$$

3.6. Medidas de Forma

3.6.1. Coeficientes de Asimetría

- Primer Coeficiente de Pearson (S_p)

$$S_p = \frac{\bar{x} - M_o}{S} \quad \text{en distribuciones Unimodales}$$

- Segundo Coeficiente de Pearson (S_p)

$$S_p = \frac{3(\bar{x} - M_e)}{S}$$

- Coeficiente de Asimetría de Bowley (S_q)

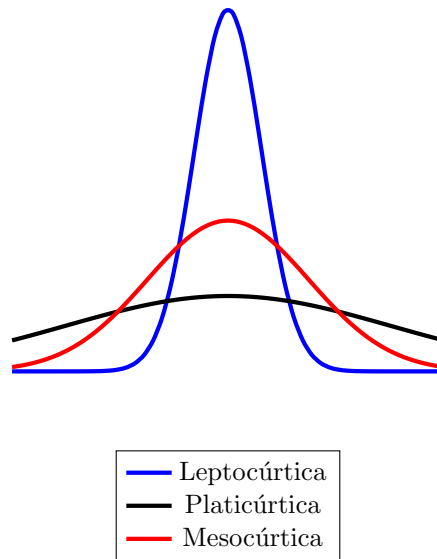
$$S_q = \frac{(Q_3 - M_e) - (M_e - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

- Coeficiente de Asimetría de Fisher (S_m)

$$S_m = \frac{M_3}{\sqrt{M_2^3}}$$

3.6.2. Curtosis o Apuntalamiento

Se entiende como el grado de deformación vertical a la curva de frecuencia, con relación al grado de apuntalamiento tenemos:



Curtosis en función del Momento

$$a_4 = \frac{M_4}{S^4} = \frac{M_4}{M_2^2}$$

Interpretación

- $a_4 = 3 \Rightarrow$ la Curva es Mesocúrtica.
- $a_4 > 3 \Rightarrow$ la Curva es Leptocúrtica.
- $a_4 < 3 \Rightarrow$ la Curva es Platicúrtica.

3.7. Medidas de Concentración

Las medidas de concentración sirven para estimar el grado de igualdad en el reparto o grado de desigualdad de la distribución de cantidades económicas, riquezas, sueldos, etc. También sirve para analizar la concentración poblacional de un territorio.

3.7.1. Índice de Gini (G)

Mide el grado de concentración a través de la diferencia entre p_i y q_i correspondientes, donde p_i y q_i son porcentajes acumulados de la distribución.

La sumatoria es desde $i = 1$ a $k - 1$ porque resultaría en cero.

Cálculo

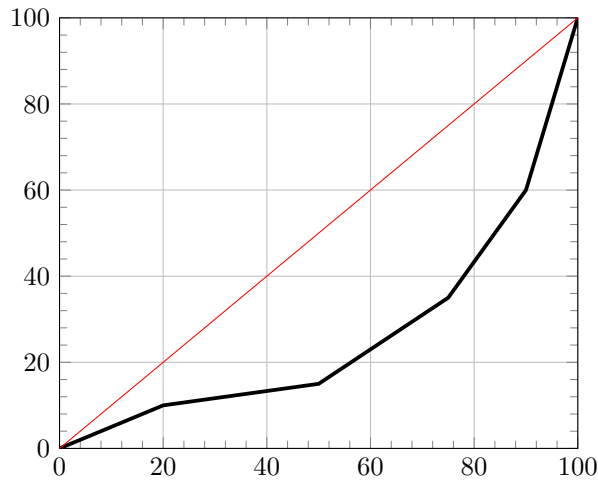
$$G = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i}$$

Observar que $0 \leq G \leq 1$ y $0 \leq \frac{G}{2} \leq \frac{1}{2}$.

Interpretación de $\frac{G}{2}$

Si $\frac{G}{2} > 10\%$ se considera que la distribución es alta.

3.7.2. Curva de Lorenz



3.8. Distribución Bivariante

Cuando se observan 2 características conjuntamente surgen las estadísticas de 2 variables, entonces los resultados numéricos de las observaciones vienen dado por parejas, es decir de forma conjunta e inseparable. Las veces que se repite cada pareja de valores es la frecuencia absoluta conjunta.

Paralelamente a los conceptos de frecuencia acumulada y relativa para distribuciones de una sola variable podrían también definirse los correspondientes o frecuencias acumuladas conjuntas y frecuencias relativas conjuntas.

3.8.1. Distribución de Frecuencias Absolutas Conjuntas

	$y'_{j-1} - y'_j$	$y'_0 - y'_1$	$y'_1 - y'_2$	$y'_2 - y'_3$	\dots	$y'_{q-1} - y'_q$	Totales
$x'_{i-1} - x'_i$	y_i x_i	y_1	y_2	y_3	\dots	y_q	$\sum_{j=1}^q f_{ij} = f_{i.}$
$x'_0 - x'_1$	x_1	f_{11}	f_{12}	f_{13}	\dots	f_{1q}	$f_{1.}$
$x'_1 - x'_2$	x_2	f_{21}	f_{22}	f_{23}	\dots	f_{2q}	$f_{2.}$
$x'_2 - x'_3$	x_3	f_{31}	f_{32}	f_{33}	\dots	f_{3q}	$f_{3.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
$x'_{p-1} - x'_p$	x_p	f_{p1}	f_{p2}	f_{p3}	\dots	f_{pq}	$f_{p.}$
Totales	$\sum_{j=1}^p f_{ij} = f_{.j}$	$f_{.1}$	$f_{.2}$	$f_{.3}$	\dots	$f_{.q}$	$n = \sum_{j=1}^p \sum_{i=1}^q f_{ij}$

3.8.2. Distribución de Frecuencias Relativas Conjuntas

	$y'_{j-1} - y'_j$	$y'_0 - y'_1$	$y'_1 - y'_2$	$y'_2 - y'_3$	\dots	$y'_{q-1} - y'_q$	Totales
$x'_{i-1} - x'_i$	y_i x_i	y_1	y_2	y_3	\dots	y_q	$\sum_{j=1}^q h_{ij} = h_{i.}$
$x'_0 - x'_1$	x_1	h_{11}	h_{12}	h_{13}	\dots	h_{1q}	$h_{1.}$
$x'_1 - x'_2$	x_2	h_{21}	h_{22}	h_{23}	\dots	h_{2q}	$h_{2.}$
$x'_2 - x'_3$	x_3	h_{31}	h_{32}	h_{33}	\dots	h_{3q}	$h_{3.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
$x'_{p-1} - x'_p$	x_p	h_{p1}	h_{p2}	h_{p3}	\dots	h_{pq}	$h_{p.}$
Totales	$\sum_{j=1}^p h_{ij} = h_{.j}$	$h_{.1}$	$h_{.2}$	$h_{.3}$	\dots	$h_{.q}$	$\sum_{j=1}^p \sum_{i=1}^q h_{ij} = 1$

3.8.3. Distribución de Frecuencias Absolutas Acumuladas Conjuntas

	$y'_{j-1} - y'_j$	$y'_0 - y'_1$	$y'_1 - y'_2$	$y'_2 - y'_3$	\dots	$y'_{q-1} - y'_q$
$x'_{i-1} - x'_i$	y_i x_i	y_1	y_2	y_3	\dots	y_q
$x'_0 - x'_1$	x_1	F_{11}	F_{12}	F_{13}	\dots	F_{1q}
$x'_1 - x'_2$	x_2	F_{21}	F_{22}	F_{23}	\dots	F_{2q}
$x'_2 - x'_3$	x_3	F_{31}	F_{32}	F_{33}	\dots	F_{3q}
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
$x'_{p-1} - x'_p$	x_p	F_{p1}	F_{p2}	F_{p3}	\dots	$F_{pq} = n$

3.8.4. Distribución de Frecuencias Relativa Acumuladas Conjuntas

	$y'_{j-1} - y'_j$	$y'_0 - y'_1$	$y'_1 - y'_2$	$y'_2 - y'_3$	\dots	$y'_{q-1} - y'_q$
$x'_{i-1} - x'_i$	y_i x_i	y_1	y_2	y_3	\dots	y_q
$x'_0 - x'_1$	x_1	H_{11}	H_{12}	H_{13}	\dots	H_{1q}
$x'_1 - x'_2$	x_2	H_{21}	H_{22}	H_{23}	\dots	H_{2q}
$x'_2 - x'_3$	x_3	H_{31}	H_{32}	H_{33}	\dots	H_{3q}
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
$x'_{p-1} - x'_p$	x_p	H_{p1}	H_{p2}	H_{p3}	\dots	$H_{pq} = 1$

3.8.5. Distribuciones Marginales

Distribución Marginal de X

X	$f_{i.}$	$h_{i.}$
x_1	$f_{1.}$	$h_{1.}$
x_2	$f_{2.}$	$h_{2.}$
x_3	$f_{3.}$	$h_{3.}$
\vdots	\vdots	\vdots
x_p	$f_{p.}$	$h_{p.}$

Distribución Marginal de Y

Y	$f_{.j}$	$h_{.j}$
y_1	$f_{.1}$	$h_{.1}$
y_2	$f_{.2}$	$h_{.2}$
y_3	$f_{.3}$	$h_{.3}$
\vdots	\vdots	\vdots
y_p	$f_{.p}$	$h_{.p}$

Son las distribuciones correspondientes a cada una de las filas o columnas de la tabla de doble entrada, si la variable x toma p valores y y toma q valores, entonces existen: $p + q$ distribuciones condicionadas.

3.9. Covarianza (S_{xy})

La Covarianza de la estadística de 2 variables se define como: La media aritmética de los productos de las desviaciones de cada una de las variables respecto a su media aritmética.

3.9.1. Cálculo

Para datos Originales

$$S_{xy} = \frac{\sum_{j=1}^p \sum_{i=1}^q (x_i - \bar{x})(y_j - \bar{y})}{n}$$

Para datos Agrupados

$$S_{xy} = \frac{\sum_{j=1}^p \sum_{i=1}^q (x_i - \bar{x})(y_j - \bar{y})f_{ij}}{n}$$

3.9.2. Propiedades

1. $Cov(ax + b, cy + d) = acCov(x, y)$
2. $V(x \pm y) = V(x) + [V(y) \pm 2Cov(x, y)]$
3. $V(ax \pm by) = a^2V(x) + b^2V(y) \pm 2abCov(x, y)$

3.9.3. Interpretación

- Si $S_{xy} > 0$, entonces existe relación directa entre x y y .
- Si $S_{xy} < 0$, entonces existe relación inversa entre x y y .
- Si $S_{xy} = 0$, **no** existe relación entre x y y (no necesariamente independientes).
- x y y son independientes entonces $S_{xy} = 0$

Capítulo 4

Análisis de Regresión y Correlación

Frecuentemente se presentan situaciones en las que es necesario estudiar las relaciones entre variables, esto se puede estudiar por análisis de regresión y/o correlación.

El análisis de regresión se refiere a la naturaleza de las relaciones entre las variables. El análisis de correlación tiene que ver con la fuerza o intensidad de las variables. Cuando el estudio de las relaciones está limitado a dos variables se denomina *Análisis de Regresión Simple* y *Análisis de Correlación Simple*, respectivamente.

4.1. Modelo de Regresión Lineal Simple

Vamos a rotular las dos variables implícitas x, y : x denominado variable independiente, predictora, o explicativa ya que puede ser controlada por el investigador. La variable y se la denomina, variable dependiente, predictoria o expectando.

4.2. Regresión Lineal

Si al representar los puntos de un gráfica, estos muestran un comportamiento rectilíneo, como en la siguiente figura:

Es necesario calcular los parámetros o coeficientes de la recta:

$$\hat{y} = a + bx$$

Para determinar a y b , que es la tarea principal de la estimación de la ecuación de regresión, recurrimos al *Método de Mínimos Cuadrados (MMC)*:

$$z = \sum_{i=1}^n (y_i - \hat{y})^2 \Leftrightarrow z = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Derivando parcialmente respecto a cada uno de los parámetros se obtienen dos ecuaciones llamadas, *Ecuaciones Normales*:

Derivamos respecto de a :

$$\frac{\partial z}{\partial a} = 2 \sum_{i=1}^n (y_i - a - bx_i)(-1) = 0$$

Derivamos respecto de b :

$$\frac{\partial z}{\partial b} = 2 \sum_{i=1}^n (y_i - a - bx_i)(-x_i) = 0$$

Multiplicamos todo por -1 y dejamos igualado a cero: Multiplicamos todo por $-x_i$ y dejamos igualado a cero:

$$\sum_{i=1}^n (-y_i + a + bx_i) = 0$$

$$\sum_{i=1}^n (-y_i x_i + ax_i + bx_i^2) = 0$$

Distribuimos la sumatoria:

Distribuimos la sumatoria:

$$-\sum_{i=1}^n y_i + \sum_{i=1}^n a + b \sum_{i=1}^n x_i = 0$$

$$-\sum_{i=1}^n y_i x_i + a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = 0$$

Dejamos de la siguiente manera:

$$\sum_{i=1}^n y_i = a \cdot n + b \sum_{i=1}^n x_i$$

Dejamos de la siguiente manera:

$$\sum_{i=1}^n y_i x_i = a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2$$

Hacemos el siguiente sistema de ecuaciones:

$$\begin{cases} \sum_{i=1}^n y_i = a \cdot n + b \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i x_i = a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 \end{cases}$$

Hallamos b mediante la determinante:

$$b = \frac{\begin{vmatrix} n & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \cdot x_i \end{vmatrix}}{\begin{vmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{vmatrix}} = \frac{n \sum_{i=1}^n y_i \cdot x_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$$

Para hallar a tomamos la primera ecuación del sistema y despejamos a :

$$a = \frac{\sum_{i=1}^n y_i}{n} - \frac{b \sum_{i=1}^n x_i}{n} \Leftrightarrow a = \bar{y} - b\bar{x}$$

Con esto, reemplazamos luego de hallar sus valores utilizando los datos, en:

$$\hat{y} = a + bx$$

4.3. Regresión Potencial

4.4. Regresión Exponencial

Tiene la siguiente forma:

4.5. Regresión Parabólica

4.6. Coeficiente de Correlación Rectilíneo

4.6.1. Propiedades

1. $-1 \leq r \leq 1$
2. Si $r > 0$, entonces existe correlación directa positiva.
3. Si $r < 0$, entonces existe correlación inversa negativa.
4. $r^2 = 0$, los datos forman una línea recta, en caso de correlación rectilínea.
5. Si $r = 1$, entonces existe correlación perfecta positiva.
6. Si $r = -1$, entonces existe correlación perfecta negativa.
7. Si $r = 0$, los datos son incorrelacionados.

Observaciones

- El signo de r es el mismo signo de b de la ecuación de regresión: $\hat{y} = a + bx$
- En la interpretación clásica se sostiene:
 1. Si $0,00 \leq r < 0,20$, existe correlación no significativa.
 2. Si $0,20 \leq r < 0,40$, existe baja correlación.
 3. Si $0,40 \leq r < 0,70$, existe correlación significativa.
 4. Si $0,70 \leq r < 1,00$, existe alto grado de asociación.

4.6.2. Cálculo

Mediante Fórmula propuesta por Pearson

$$r = \frac{n \cdot \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \cdot \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

En términos de Covarianza y Desviación Estándar

$$r = \frac{Cov(x, y)}{S_x \cdot S_y}$$

Error Estándar de Estimación

$$S_{yx} = \sqrt{\frac{\sum_{i=1}^n y_i^2 - a \cdot \sum_{i=1}^n y_i - b \cdot \sum_{i=1}^n y_i x_i}{n}}$$

Capítulo 5

Probabilidad

Es una disciplina abstracta que se usa como modelo para hacer deducciones acerca de eventos que posiblemente puedan suceder, por tanto la probabilidad no proporciona la base para construir medidas exactas de incertidumbre, para su estudio es necesario conocer los siguientes conceptos:

Experimentos

Experimentos Determinísticos

Un experimento es determinístico cuando el resultado de la observación es determinado en forma precisa, bajo las condiciones por las cuales se realiza dicho experimento. Algunos ejemplos son:

- Observar el color de una ficha extraída de una urna que solo contiene fichas rojas.
- Observar el resultado de la suma de dos números naturales impares.

Experimentos no Determinísticos

5.1. Eventos

Evento es un subconjunto del espacio muestral. Existen los siguientes tipos de eventos:

- **Evento Imposible:** es el evento que nunca ocurre y se denota por: ϕ .
- **Evento Seguro:** es el evento que siempre ocurre, denotado por: S o Ω .
- **Evento Elemental:** es el evento que esta compuesto por un elemento, se lo denota como conjunto *Unitario*.
- **Evento Compuesto:** formado por varios Eventos.
- **Eventos Incompatibles:** también llamados mutuamente excluyentes.

$$\left(\begin{array}{c} A \text{ y } B \\ \text{Son Eventos Excluyentes} \end{array} \right) \Leftrightarrow A \cap B = \phi$$

- **Eventos Compatibles:**

$$\left(\begin{array}{c} A \text{ y } B \\ \text{Son Eventos Compatibles} \end{array} \right) \Leftrightarrow A \cap B \neq \phi$$

- **Eventos Colectivamente Exhaustivos:**

$$\left(\begin{array}{c} A \text{ y } B \\ \text{Son E.C.E.} \end{array} \right) \Leftrightarrow \begin{cases} A \cup B = \Omega \\ A \cap B = \phi \end{cases}$$

5.1.1. Operaciones con Eventos

Unión

Sea A y B eventos cualquiera; la unión de estos eventos es el evento que ocurre, si A ocurre y B ocurre, o ambos ocurren.

- **Formalmente**

$$A \cup B = \{w \in S / w \in A \vee w \in B\}$$

- **Generalización**

$$\bigcup_{i=1}^n A_i = \{w \in S / \exists i \in I : w \in A_i\}$$

Intersección

- **Formalmente**

$$A \cap B = \{w \in S / w \in A \wedge w \in B\}$$

- **Generalización**

$$\bigcap_{i=1}^n A_i = \{w \in S / \exists i \in I : w \in A_i\}$$

Diferencia Relativa

- **Formalmente**

$$A - B = \{w \in S / w \in A \wedge w \notin B\}$$

Diferencia Simétrica

- **Formalmente**

$$A \triangle B = \{w \in S / w \in A \vee w \in B\}$$

Complemento

- **Formalmente**

$$A^c = \{w \in S / w \notin A\}$$

Inclusión

- **Formalmente**

$$A \subseteq B \Leftrightarrow \forall w \in S, (w \in A \rightarrow w \in B)$$

Igualdad de Eventos

- **Formalmente**

$$A = B \Leftrightarrow (A \subseteq B \wedge B \subseteq A)$$

5.2. σ -álgebra de Eventos

Se dice que la clase de subconjuntos del espacio muestral S , denotado por:

$$\Gamma = \{A \subset S / A \text{ es un Evento}\}$$

es un σ -álgebra de eventos si:

(I) $S \in \Gamma$

(II) $A \in \Gamma \Rightarrow A^c \in \Gamma$

(III) $\forall i \in I, \left(A_i \in \Gamma \Rightarrow \bigcup_{i=1} A_i \in \Gamma \right)$

5.3. Definición Axiomática de Probabilidad

Sea S un espacio muestral asociado a un cierto experimento aleatorio y sea Γ σ -álgebra de eventos subconjunto de S .

$$\begin{aligned} P: \Gamma &\rightarrow [0, 1] \\ A &\mapsto P(A) \end{aligned}$$

Tal que satisface los siguientes axiomas:

$$(I) \quad 0 \leq P(A) \leq 1$$

$$(II) \quad P(S) = 1$$

$$(III) \quad P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad A_i \text{ son eventos mutuamente excluyentes}$$

Una consecuencia inmediata del axioma (III) es:

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

Esto es si A y B son eventos mutuamente excluyentes.

5.3.1. Propiedades de la Probabilidad de Eventos

1. Si ϕ es un evento imposible entonces: $P(\phi) = 0$

2. Si A^c es el evento complementario de A , entonces:

$$P(A^c) = 1 - P(A) \quad \text{ó} \quad P(A) = 1 - P(A^c)$$

3. Si A y B son eventos tales que: $A \subset B$ entonces:

$$P(A) \leq P(B)$$

4. Si A y B son eventos cualesquiera, entonces:

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

5. Si A, B y C son eventos mutuamente excluyentes, entonces:

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

6. Si A, B y C son eventos cualesquiera, entonces:

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

7. Si $A_1, A_2, A_3, \dots, A_n$ son eventos cualesquiera:

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$$

5.4. Definición Clásica de Probabilidad

$$P(A) = \frac{n(A)}{n(S)} = \frac{\# \text{ de casos favorables del Evento } A}{\# \text{ total de casos posibles}}$$

5.4.1. Probabilidad Condicionada

Sean A y B eventos tal que $P(B) > 0$ la probabilidad condicional de que ocurra A dado que ha ocurrido el evento B se denota por: $P(A/B)$. Se define como:

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

Si B es un evento tal que $P(B) > 0$, entonces $P(\bullet/B)$ satisface los siguientes axiomas:

- (i) $0 \leq P(A/B) \leq 1$
- (ii) $P(S/B) = 1$
- (iii) $P\left(\bigcup_{i=1}^n A_i/B\right) = \sum_{i=1}^n P(A_i/B)$ A_i son eventos mutuamente excluyentes

Propiedades

Si B es un evento tal que $P(B) > 0$, entonces $P(\bullet/B)$ tiene las siguientes propiedades:

1. $P(\phi/B) = 0$
2. $P(A^c/B) = 1 - P(A/B) \quad \vee \quad P(A/B) = 1 - P(A^c/B)$
3. $P((A \cup C)/B) = P(A/B) + P(C/B) - P((A \cap C)/B)$
4. $A \in C \Rightarrow P(A/B) \leq P(C/B)$

5.4.2. Eventos Dependientes

Se dice que los eventos A y B son estadísticamente dependientes, si:

$$\left. \begin{aligned} P(A \cap B) &= P(B) \cdot P(A/B) \\ P(A \cap B) &= P(A) \cdot P(B/A) \end{aligned} \right\} \text{Regla General de la Multiplicación}$$

5.4.3. Eventos Independientes

Se dice que los eventos A y B son estadísticamente independientes, si:

$$P(A \cap B) = P(A) \cdot P(B)$$

Observaciones

1. $\left. \begin{aligned} P(A) \\ P(B) \end{aligned} \right\}$ Probabilidades Marginales
2. $\left. \begin{aligned} P(A/B) \\ P(B/A) \end{aligned} \right\}$ Probabilidades Condicionales
3. $P(A/B) \neq P(B/A)$
4. $P(A \cap B)$ Probabilidades Condicionales
5. Decir '*Eventos Independientes*' no es lo mismo que decir '*Evento Mutuamente Excluyente*'.

Teorema

Si A y B son eventos mutuamente independientes, entonces:

1. $P(A/B) = P(A)$
2. $P(B/A) = P(B)$
3. A^c y B son independientes.
4. A y B^c son independientes.
5. A^c y B^c son independientes.

En $P(\bullet/B)$ el símbolo \bullet significa *cualquiera*

Demostración de los Teoremas

- Demostración de 1. :

$$\begin{aligned} P(A/B) &\stackrel{1}{=} \frac{P(A \cap B)}{P(B)} \\ &\stackrel{2}{=} \frac{P(A) \cdot P(B)}{P(B)} \\ &\stackrel{3}{=} P(A) \end{aligned}$$

- Demostración de 2. :

$$\begin{aligned} P(B/A) &\stackrel{1}{=} \frac{P(B \cap A)}{P(A)} \\ &\stackrel{2}{=} \frac{P(A) \cdot P(B)}{P(A)} \\ &\stackrel{3}{=} P(B) \end{aligned}$$

- Demostración de 3. :

$$P(A^c \cap B) = P(A^c) \cdot P(B)$$

Formamos el siguiente sistema:

$$\begin{cases} B = (A \cap B) \cup (A^c \cap B) \\ \phi = (A \cap B) \cap (A^c \cap B) \end{cases}$$

Ya que son excluyentes:

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

Despejamos $P(A^c \cap B)$ ya que es lo que nos interesa demostrar:

$$P(A^c \cap B) = P(B) - P(A \cap B)$$

Como A y B son independientes, según la hipótesis, tenemos que $P(A \cap B) = P(A)P(B)$

Reemplazando en el anterior paso:

$$\begin{aligned} P(A^c \cap B) &\stackrel{1}{=} P(B) - P(A)P(B) \\ &\stackrel{2}{=} P(B) \cdot [1 - P(A)] \\ &\stackrel{3}{=} P(B) \cdot P(A^c) \end{aligned}$$

Con esto que demostrado 3.

- Demostración de 4. :

$$P(A \cap B^c) = P(A) \cdot P(B^c)$$

Formamos el siguiente sistema:

$$\begin{cases} A = (A \cap B) \cup (A \cap B^c) \\ \phi = (A \cap B) \cap (A \cap B^c) \end{cases}$$

Ya que son excluyentes:

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

Despejamos $P(A \cap B^c)$ ya que es lo que nos interesa demostrar:

$$P(A \cap B^c) = P(A) - P(A \cap B)$$

Como A y B son independientes, según la hipótesis, tenemos que $P(A \cap B) = P(A)P(B)$

Reemplazando en el anterior paso:

$$\begin{aligned} P(A \cap B^c) &\stackrel{1}{=} P(A) - P(A)P(B) \\ &\stackrel{2}{=} P(A) \cdot [1 - P(B)] \\ &\stackrel{3}{=} P(A) \cdot P(B^c) \end{aligned}$$

Con esto queda demostrado 4.

- Demostración de 5. : En este punto, lo que tratamos de demostrar es:

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

Comenzamos aplicando D'Morgan:

$$\begin{aligned}
 P(A^c \cap B^c) &\stackrel{1}{=} P[(A \cup B)^c] \\
 &\stackrel{2}{=} 1 - P(A \cup B) \\
 &\stackrel{3}{=} 1 - [P(A) + P(B) - P(A \cap B)] \\
 &\stackrel{4}{=} 1 - P(A) - P(B) + P(A \cap B) \\
 &\stackrel{5}{=} [1 - P(A)] - [P(B) - P(A \cap B)] \\
 &\stackrel{6}{=} [1 - P(A)] - [P(B) - P(A)P(B)] \\
 &\stackrel{7}{=} [1 - P(A)] - P(B)[1 - P(A)] \\
 &\stackrel{8}{=} [1 - P(A)][1 - P(B)] \\
 &\stackrel{9}{=} P(A^c)P(B^c)
 \end{aligned}$$

Con esto queda demostrado 5.

5.4.4. Teorema de la Probabilidad Total

También llamada *Regla de Eliminación*. Si los eventos mutuamente excluyentes:

$$B_1, B_2, B_3, \dots, B_n$$

Constituyen una partición del espacio muestral S , de tal manera que $\forall i \in I, P(B_i) > 0$. Entonces para cualquier evento A de S se cumple:

$$P(A) = \sum_{i=1}^n P(B_i)P(A/B_i)$$

Demostración

$$\begin{aligned}
 P(A) &\stackrel{1}{=} [(B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_n \cap A)] \\
 &\stackrel{2}{=} P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_n \cap A) \\
 &\stackrel{3}{=} \sum_{i=1}^n P(B_i \cap A) \\
 &\stackrel{4}{=} \sum_{i=1}^n P(B_i) \cdot P(A/B_i)
 \end{aligned}$$

5.4.5. Teorema de Bayes

Si los eventos: $B_1, B_2, B_3, \dots, B_n$, constituyen una partición del espacio muestral S , de tal modo que; $\forall i \in I, P(B_i) > 0$, entonces para cualquier evento A de S , tal que $P(A) > 0$ se cumple:

$$P(A) = \frac{P(B_r)P(A/B_r)}{\sum_{i=1}^n P(B_i)P(A/B_i)}$$

Demostración

$$\begin{aligned}
P(B_r/A) &\stackrel{1}{=} \frac{P(B_r \cap A)}{P(A)} \\
&\stackrel{2}{=} \frac{P(B_r)P(A/B_r)}{P(A)} \\
&\stackrel{3}{=} \frac{P(B_r)P(A/B_r)}{\sum_{i=1}^n P(B_i)P(A/B_i)}
\end{aligned}$$

5.4.6. Ejemplos

1. Demostrar que: $P(A/B) + P(A^c/B) = 1$

Solución

$$\begin{aligned}
P(A/B) + P(A^c/B) &\stackrel{1}{=} \frac{P(A \cap B)}{P(B)} + \frac{P(A^c \cap B)}{P(B)} \\
&\stackrel{2}{=} \frac{P(A \cap B) + P(A^c \cap B)}{P(B)} \\
&\stackrel{3}{=} \frac{P[(A \cap B) \cup (A^c \cap B)]}{P(B)} \\
&\stackrel{4}{=} \frac{P[(A \cup A^c) \cap B]}{P(B)} \\
&\stackrel{5}{=} \frac{P(\Omega \cap B)}{P(B)} \\
&\stackrel{6}{=} \frac{P(B)}{P(B)} \\
&\stackrel{7}{=} 1
\end{aligned}$$

2. Demostrar que: $P(A/B^c) + P(A^c/B) = 1$

Solución

3. La probabilidad de que llueva en Santa Cruz el 10 de Junio es de 40 %, de que truene es 5 % y de que, llueva y truene es de 3 %. ¿Cual es la probabilidad de que llueva o truene ese día?

Solución

4. Sea A y B dos eventos, tal que: $P(A) = 0,20 \%$; $P(B) = 0,30 \%$; $P(A \cup B) = 0,10 \%$
Hallar:

- (a) $P(A^c \cap B^c)$
- (b) $P(A^c \cap B)$
- (c) $P(A \cap B^c)$
- (d) $P(A^c \cup B)$

5. Con 7 ingenieros y 4 médicos se formaron comités de 6 miembros. ¿Cuál es la probabilidad que el comité incluya: ?
 - (a) Exactamente 2 médicos.
 - (b) Al menos 2 ingenieros.

6. En la UAGRM el 30 % de los estudiantes son cruceños, el 10 % estudia Ingeniería Informática, y el 1 % son cruceños y estudian Informática. Si se selecciona al azar un estudiante de la UAGRM. Hallar las siguientes probabilidades:

Bibliografía

- [1] Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, *Probabilidad y estadística para ingeniería y ciencias, Novena Edición*, Pearson, 2012.
- [2] Braulio Cáceres Chacón, Cindy Elisa Cáceres Antelo, *Problemas Resueltos y Propuestos de Estadística Descriptiva, Primera Edición*, Editorial e Imprenta Universitaria (U.A.G.R.M.), 2016.
- [3] Carlos Véliz Cupañay, *Estadística para la Administración y Negocios, Primera Edición*, Pearson, 2011.
- [4] Juan Martínez de Lejarza: Series Temporales
<https://www.uv.es/ceaces/pdf/series.pdf>
- [5] Series Temporales, Departamento de Estadística, Universidad Carlos III de Madrid.
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/EDescrip/tema7.pdf>
- [6] Manuel Córdova Zamora, *Estadística Descriptiva e Inferencial, Quinta Edición*, Editorial Moshera, 2003.