

Estadística Decriptiva Univariada

Dr. Cristian Carvajal Muquillaza

Universidad de Playa Ancha



Universidad de
Playa Ancha
Facultad de Ciencias
Naturales y Exactas

**DEPARTAMENTO
DISCIPLINARIO
DE MATEMÁTICA
Y ESTADÍSTICA**

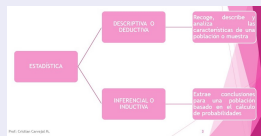
Contenido

- 1 Preliminares
- 2 Resumen de la Información
- 3 Representaciones Gráficas
- 4 Medidas Estadísticas
 - Medidas de Tendencia Central
 - Medidas de Variabilidad
 - Medidas de Posición
 - Medidas de Forma



ESTADÍSTICA

La Estadística es la ciencia que aporta los métodos científicos por medio de los cuales es posible recolectar, organizar, resumir, presentar y analizar datos relativos a un conjunto de individuos u observaciones y que nos permiten extraer conclusiones válidas y efectuar decisiones lógicas basadas en dicho análisis.



Población Estadística:

Es un conjunto P de personas, animales o cosas que son objeto de estudio de una investigación.

- Estudiantes Universitarios
- Pacientes
- Puntajes PSU
- Merluzas en la costa de Valparaíso
- Celulares Samsung.

Técnicas de Muestreo Probabilístico

Muestreo Aleatorio simple (MAS):

Es una técnica en que cada individuo o elemento de la población tiene la misma oportunidad de ser seleccionada para pertenecer a la muestra. Este muestreo puede ser **Sin Reposición** o **Con Reposición**. Suponga una Población de tamaño N donde se extrae una muestra de tamaño n , $(x_1, x_2, \dots, x_n)$

Cálculo del número de muestras posibles		
	Teniendo en cuenta el orden	Sin tener en cuenta el orden
Sin Reposición	$n! \binom{N}{n}$	$\binom{N}{n}$
Con Reposición	N^n	$\binom{N+n-1}{n}$

Técnicas de Muestreo Probabilístico

Ejemplo: Cálculo del número de muestras posibles

Suponga una población P de tamaño $N = 3$ y se quiere extraer una MAS de tamaño $n = 2$. Determine todas las posibles muestras.

Resp: Suponga $P = \{Leo, Nico, Pilar\}$

	Con Orden	Sin Orden
Sin Reposición	$2!(\binom{3}{2}) = 6$	$(\binom{3}{2}) = 3$
Con Reposición	$3^2 = 9$	$(\binom{3+2-1}{2}) = 6$

Técnicas de Muestreo Probabilístico

Muestreo Aleatorio sistemático:

- Como en general, N no es un múltiplo entero de k , las diferentes muestras sistemáticas de la misma población finita pueden variar su tamaño en 1 unidad
- Este hecho produce una perturbación despreciable si $n \geq 30$



Técnicas de Muestreo Probabilístico

Ejemplo:

Suponga $N = 23$ y $k = 5$

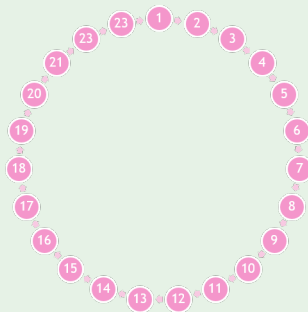
I	II	III	IV	V
1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23		



Técnicas de Muestreo Probabilístico

Ejemplo:

Suponga que se desea $n = 5$ con $N = 23$



Técnicas de Muestreo Probabilístico

Ejemplo:

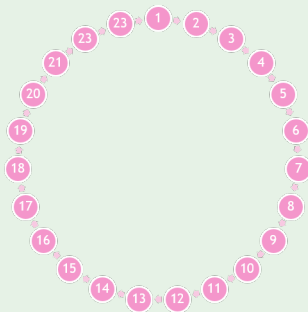
Se tiene una comunidad de 20 familias, y se quiere estudiar la responsabilidad en la tenencia de mascotas. Sea y_i el número de mascotas en las viviendas de la comunidad. Para el estudio se requiere una muestra sistemática de tamaño 4. Determinar todas las posibles muestras sistemáticas

u_i	1	2	3	4	5	6	7	8	9	10
y_i	0	4	2	4	5	3	1	2	6	0
u_i	11	12	13	14	15	16	17	18	19	20
y_i	7	5	1	5	0	10	8	5	5	3

Técnicas de Muestreo Probabilístico

Ejemplo:

Suponga que se desea $n = 5$ con $N = 23$



Técnicas de Muestreo Probabilístico

Muestreo Aleatorio Estratificado:

- La población de N unidades se divide en subpoblaciones $N_1, N_2, \dots, N_L$ unidades,
- Estas subpoblaciones son disjuntas, es decir no se intersecan.
- $N_1 + N_2 + \dots + N_L = N$
- Las subpoblaciones se denominan estratos
- Los tamaños de las muestras de cada estrato se denotan $n_1, n_2, \dots, n_L$, respectivamente.
- Si se toma una muestra aleatoria simple en cada estrato, el procedimiento total se describe como un Muestreo Aleatorio Estratificado (MAE).

Técnicas de Muestreo Probabilístico

Muestreo Aleatorio Estratificado: Notación

El subíndice h denota el estrato, e i la unidad dentro del estrato. Todos los símbolos siguientes se refieren al estrato h .

Símbolo	Descripción
N_h	Número total de unidades
n_h	Número de unidades en la muestra
y_{hi}	Valor obtenido para la i -ésima unidad
$W_h = \frac{N_h}{N}$	Ponderación del estrato
$f_h = \frac{n_h}{N_h}$	Fracción de muestreo en el estrato

Técnicas de Muestreo Probabilístico

Asignación optima de n_h

En el MAE la varianza se minimiza para un tamaño de muestral total fijo n si

$$n_{h,\text{opt}}(\text{A.N}) = n \frac{N_h s_h}{\sum N_h s_h}$$

Esta asignación se llama ASIGNACIÓN DE NEYMAN, donde

$$s_h = \sqrt{\frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}$$

Universidad de Playa Ancha

Técnicas de Muestreo Probabilístico

Ejemplo:

En una población con $N = 6$ y $L = 2$, los valores de y_{hi} son 0, 1, 2 en el estrato 1 y 4, 6, 11 en el estrato 2. Se va a tomar una muestra de tamaño $n = 4$,

- Obtenga n_h óptimo bajo la asignación de Neyman $[n_{h,\text{opt}}(\text{A.N})]$
- Obtenga n_h óptimo bajo la asignación Proporcional $[n_{h,\text{opt}}(\text{A.P})]$
- Obtenga todas las muestras posibles bajo estas dos asignaciones.



Técnicas de Muestreo Probabilístico

Tamaño de la Muestra para la Proporción Poblacional P

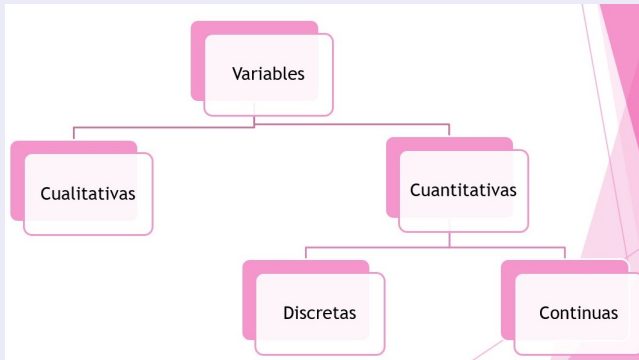
Cuando N es desconocida	$n = \frac{Z_{1-\alpha/2}^2 \cdot p \cdot q}{e^2}$
Cuando N es conocida	$n = \frac{N \cdot Z_{1-\alpha/2}^2 \cdot p \cdot q}{e^2(N-1) + Z_{1-\alpha/2}^2 \cdot p \cdot q}$

Donde:

- p proporción de elementos que poseen la característica deseada
- σ Desviación Estándar
- $Z_{1-\alpha/2}$ Cuantil Normal
- e Error Muestral Admitido

Preliminares

Variables



Preliminares

Escala de Medición de las Variables

Escala intervalar: En esta escala la distancia entre las unidades de medida son uniformes. Permite realizar operaciones matemáticas, como suma, resta, multiplicación o división. El cero es arbitrario, no indica la ausencia de atributo. Ejemplo: grados centígrados.

Escala de razón: Similar a la de intervalo, con la única diferencia que el cero en esta escala sí indica la ausencia de atributo, es cero absoluto. Ejemplo: altura



Preliminares

Ejemplo:

Variable	Tipo de variable	Escala
Edad		
Sexo		
Promedio de Notas		
Quintil de ingreso		
Educación de la Madre		
Tiempo de Estudio		
Número de libros leídos		
Número de partes policiales		
Puntaje PSU Matemática		

Tabla de Frecuencias

Frecuencia relativa f_i :

f_i es el cociente entre la frecuencia absoluta de la clase C_i y el número total de observaciones, es decir:

$$f_i = \frac{n_i}{N}$$

Frecuencia relativa porcentual $f_i \%$:

Es la frecuencia relativa multiplicada por 100, es decir:

$$f_i \% = \frac{n_i}{N} \cdot 100$$

Universidad de Playa Ancha

Tabla de Frecuencias

Frecuencia relativa acumulada F_i :

Es el número F_i , definido como el tanto por uno de los elementos de la población que están en alguna de las clases y que presentan un «valor» inferior o igual a la clase c_i , es decir

$$F_i = \frac{N_i}{n} = \frac{n_1 + \dots + n_i}{n} = f_1 + \dots + f_i = \sum_{j=1}^i f_j.$$



Tabla de Frecuencias

En síntesis se tiene:

Clase	Frec. Abs	Frec. Rel	Frec. Acum	Frec.Rel.Acum
C	n_i	f_i	N_i	F_i
c_1	n_1	$f_1 = \frac{n_1}{n}$	$N_1 = n_1$	$F_1 = \frac{N_1}{n} = f_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
c_j	n_j	$f_j = \frac{n_j}{n}$	$N_j = n_1 + \dots + n_j$	$F_j = \frac{N_j}{n} = f_1 + \dots + f_j$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
c_k	n_k	$f_k = \frac{n_k}{n}$	$N_k = n$	$F_k = 1$
	n	1		

Tablas de Frecuencias

Ejemplo:

Construir una tabla de frecuencia INTERVALAR, dado los datos que representan los puntajes de un examen de Estadística

77	76	75	62	63	68	81	75	67	68
74	85	76	71	78	60	57	88	80	73
75	63	65	95	85	74	97	75	62	61
93	68	79	69	95	94	61	82	78	66
71	83	74	60	77	75	75	89	78	96



Tablas de Frecuencias

Ejemplo:

Suponga que el número de observaciones negativas de un grupo de estudiantes son los siguientes:

0	3	1	1	2	2	3	2	2	3
1	0	2	3	3	1	2	3	0	3
0	3	1	1	2	2	3	2	2	3
1	0	2	3	3	1	2	3	0	3

Construir una tabla de frecuencia SIMPLE.



Representaciones Gráficas

Hemos visto que la tabla estadística resume los datos que disponemos de una población, de forma que ésta se puede analizar de una manera más sistemática y resumida. Para esclarecer de mejor forma nuestros datos haremos uso de gráficos y diagramas, cuya construcción abordamos en esta sección.



Universidad de
Playa Ancha
Facultad de Ciencias
Naturales y Exactas

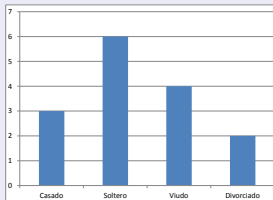
**DEPARTAMENTO
DISCIPLINARIO
DE MATEMÁTICA
Y ESTADÍSTICA**

Representaciones Gráficas

Gráficos para variables cualitativas

Los gráficos más usuales para representar variables de tipo nominal son los siguientes:

Gráfico de barras



Representaciones Gráficas

Gráficos para variables cualitativas: Barras

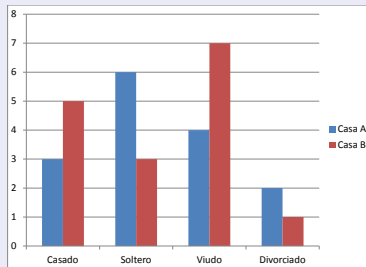
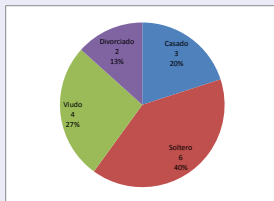


Figura: Gráfico de barras para comparar una variable cualitativa en

Representaciones Gráficas

Gráficos para variables cualitativas: Sectores

Se divide un círculo en tantas porciones como clases existan, de modo que a cada clase le corresponde un arco de círculo proporcional a su frecuencia absoluta o relativa.



Representaciones Gráficas

Gráficos para variables cuantitativas

Se utilizarán dos tipos de gráficos, en función de las frecuencias utilizadas (absolutas o relativas o acumuladas):

Diagramas diferenciales: Son aquellos en los que se representan frecuencias absolutas o relativas.

Diagramas integrales: Son aquellos en los que se representan el número de elementos que presentan una modalidad inferior o igual a una dada. Se realizan a partir de las frecuencias acumuladas, lo que da lugar a gráficos crecientes, y es obvio que este tipo de gráficos no tiene sentido para variables cualitativas.

Representaciones Gráficas

Gráficos para variables cuantitativas: **Discretas**

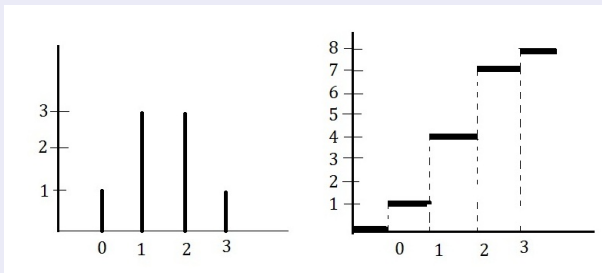


Figura: Diagrama diferencial (barras) e integral para una variable discreta. Obsérvese que el diagrama integral (creciente) contabiliza el número de observaciones de la variable inferiores o iguales a cada punto del eje de

Representaciones Gráficas

Gráficos para variables cuantitativas: **Continuas**

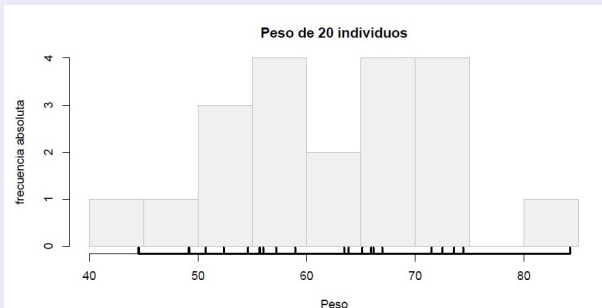


Figura: Histograma.

Representaciones Gráficas

Gráficos para variables cuantitativas: **Continuas**

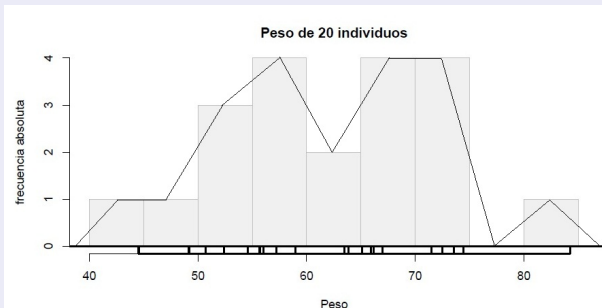
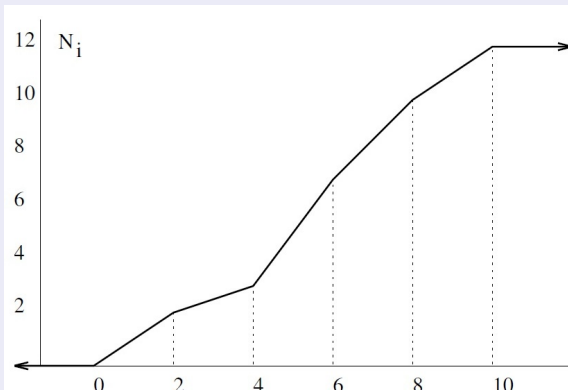


Figura: Polígono de Frecuencia.

Representaciones Gráficas

Gráficos para variables cuantitativas: **Continuas**



Representaciones Gráficas

Otros Gráficos:Box-Plot

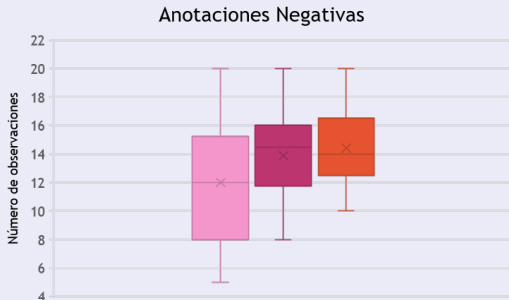


Figura: Box-Plot.

Representaciones Gráficas

Otros Gráficos: Tallo-Hoja

Tallo	Hoja
4	4 5 9
5	0 2 3 3 4 4 6 7 7 7 8
6	1 2 2 3 4 7 8 9
7	0 1 1 2 3 4 4 5 6 6 8 9
8	0 1 3 5

Figura: Tallo-Hoja.

Medidas Estadísticas

Dado un grupo de datos organizados en una distribución de frecuencias, pretendemos describirlos mediante dos o tres Medidas Estadísticas, a saber:

- La tendencia central de los datos;
- La dispersión o variación con respecto a este centro;
- Los datos que ocupan ciertas posiciones.
- La simetría de los datos.
- La forma en la que los datos se agrupan.

Medidas Estadísticas

Medidas Estadísticas

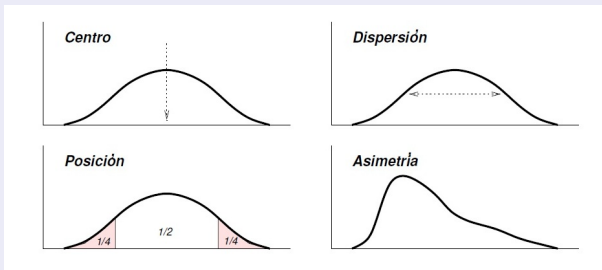


Figura: Medidas representativas de un conjunto de datos estadísticos.

Medidas Estadísticas: Medidas de Tendencia Central

Medidas de Tendencia Central

Además de presentar la información usando tablas o gráficas es de mucha utilidad el poder presentar esta información en términos de algún valor (numérico) que la resuma en forma eficiente.

Las medidas de tendencia central son muy útiles para este propósito pues describen una distribución donde se encuentran la acumulación más alta de datos, y ello ocurre con frecuencia en el “centro” de ésta.

Las tres medidas más usadas de tendencia central son:

- La Media,
- La Mediana,
- La Moda.

Medidas Estadísticas: *Tendencia Central*

Media Arítmética $\bar{X}$

La **media aritmética** de una variable estadística es la suma de todos sus posibles valores, ponderada por las frecuencias de los mismos. Es decir, si la tabla de valores de una variable X es

X_i	n_i	f_i
x_1	n_1	f_1
...	...	...
x_k	n_k	f



Entonces la media aritmética, la definimos

$$\begin{aligned}\bar{X} &= \frac{1}{n}(x_1n_1 + \dots + x_kn_k) \\ &= \frac{1}{n} \sum_{i=1}^k x_i n_i.\end{aligned}$$

Medidas Estadísticas: *Tendencia Central*

Media Arítmética $\bar{X}$

Si los datos se presentan en una tabla de distribuciones de frecuencias, todos los valores que contenga una clase se consideran iguales a la marca de clase del intervalo, por lo tanto si la tabla es de la forma

Clase C	Marca de Clase	Frec. Abs
c_1	$M.C_1$	n_1
...	...	...
c_j	$M.C_j$	n_j
...	...	...
c_k	$M.C_k$	n_k

Medidas Estadísticas: *Tendencia Central*

Media Arítmética $\bar{X}$

Entonces la Media Aritmética la definimos

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k M.C_i \times n_i.$$



Medidas Estadísticas: *Tendencia Central*

Media Arítmética $\bar{X}$

Por último si los datos no están ordenados en una tabla, es decir están sueltos o a granel, entonces la media aritmética la definimos

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$= \frac{1}{n} \sum_{i=1}^n x_i.$$

Medidas Estadísticas: *Tendencia Central*

Media Arítmética $\bar{X}$

La media presenta inconvenientes en algunas situaciones:

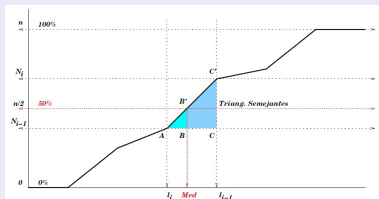
- Uno de ellos es que es muy sensible a los valores extremos de la variable, ya que todas las observaciones intervienen en el cálculo de la media, la aparición de una observación extrema, hará que la media se desplace en esa dirección.
- no es recomendable usar la media como medida central en las distribuciones muy asimétricas;



Medidas Estadísticas: *Tendencia Central*

Mediana $\tilde{X}$

Consideramos una variable discreta X cuyas observaciones en una tabla estadística han sido ordenadas de menor a mayor. Llamaremos mediana, al primer valor de la variable que deja por debajo de sí al 50 % de las observaciones.



Medidas Estadísticas: *Tendencia Central*

Mediana $\tilde{X}$

Sean $x_{(1)}, \dots, x_{(n)}$ una muestra de n datos ordenados (creciente o decreciente), entonces la mediana se define:

$$\tilde{X} = \begin{cases} X\left(\frac{n+1}{2}\right) & \text{si } n \text{ es impar} \\ \frac{X\left(\frac{n}{2}\right) + X\left(\frac{n}{2}+1\right)}{2} & \text{si } n \text{ es par} \end{cases}$$

Medidas Estadísticas: *Tendencia Central*

Mediana $\tilde{X}$

Si los datos están agrupados en una tabla de frecuencia intervalar entonces:

$$\tilde{X} = L_{i-1} + \left(\frac{\frac{n}{2} - N_{i-1}}{n_i} \right) \cdot a_i$$

Donde:

- L_{i-1} Límite inferior de la clase mediana.
- N_{i-1} Frec. absoluta acumulada de la clase anterior a la clase mediana.
- n_i Frecuencia absoluta de la clase mediana.

Medidas Estadísticas: *Tendencia Central*

Moda $\mathring{X}$

Llamaremos **moda** a cualquier máximo relativo de la distribución de frecuencias, es decir, cualquier valor de la variable que posea una frecuencia mayor que su anterior y su posterior.

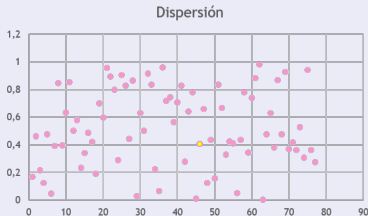
Para datos en tablas de distribución de frecuencias intervalares definiremos la moda de la siguiente manera:

$$\mathring{X} = L_{i-1} + \left(\frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \right) \cdot a_i$$

Un conjunto de datos puede: *No tener Moda (Amodal)*; *Tener más de una Moda (Polimodal)*

Medidas Estadísticas: *Variabilidad*

Las medidas de tendencia central nos indican donde se sitúa un grupo de puntuaciones. Los de variabilidad o dispersión nos indican si esas puntuaciones o valores están próximas entre sí o si por el contrario están muy dispersas.



Medidas Estadísticas: *Variabilidad*

Rango

Una medida razonable de la variabilidad podría ser la amplitud o rango, que se obtiene restando el valor más bajo de un conjunto de observaciones del valor más alto, es decir:

$$\text{Rango} = X_{\text{máx}} - X_{\text{mín}}$$

- Si $\text{Rango} \rightarrow \infty \Rightarrow \text{Heterogeneo}$
- Si $\text{Rango} \rightarrow 0 \Rightarrow \text{Homogeneo}$

Medidas Estadísticas: *Variabilidad*

Varianza S^2

La varianza, S^2 , se define como la media de las diferencias cuadráticas de n puntuaciones con respecto a su media aritmética, es decir

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

Esta medida es siempre una cantidad positiva, con propiedades interesante para la realización de inferencia estadística.



Si los datos se presentan en una tabla de distribuciones de frecuencias, entonces la varianza se define:

$$S^2 = \frac{1}{n} \sum_{i=1}^k (MC_i - \bar{x})^2 \cdot n_i$$

Medidas Estadísticas: *Variabilidad*

Desviación Estándar S

La varianza no tiene la misma magnitud que las observaciones (ej. si las observaciones se miden en metros, la varianza lo hace en metros cuadrados). Si queremos que la medida de dispersión sea de la misma dimensionalidad que las observaciones bastará con tomar su raíz cuadrada. Por ello se define la desviación típica, S , como

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

Medidas Estadísticas: *Variabilidad*

Coeficiente de Variación

El coeficiente de variación permite comparar la dispersión de dos poblaciones, pues elimina la dimensionalidad de las variables y tiene en cuenta la proporción existente entre medias y desviación típica. Se define del siguiente modo:

$$C_{var} = \frac{S_x}{\bar{x}}$$



Medidas Estadísticas: *Variabilidad*

Tipificación o Estandarización

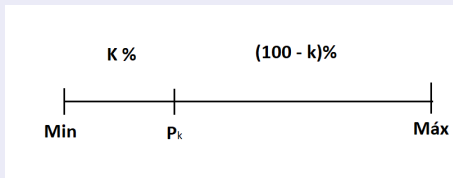
Se conoce por tipificación al proceso de restar la media y dividir por su Desviación estándar a una variable X . De este modo se obtiene una nueva variable

$$Z = \frac{X - \bar{x}}{S}$$

donde la media de $Z = 0$ y la desviación estándar o típica de $Z = 1$, lo que denominamos **variable tipificada**.

Universidad de Playa Ancha

Para una variable discreta, se define el percentil de orden k , como la observación, P_k , que deja por debajo de si el $k\%$ de la población.



Medidas Estadísticas: *Posición*

Percentiles

Sean $x_{(1)}, \dots, x_{(n)}$ una muestra de tamaño n , ordenada de forma creciente, entonces para determinar el percentil k -ésimo P_k seguimos los siguientes pasos:

- 1 Se determina la posición: $P = \frac{(n-1)k}{100} + 1$
- 2 Se buscan los datos X_i y X_j , tal que $X_i \leq P \leq X_j$.
- 3 Se calcula la distancia entre estos datos, $d = X_j - X_i$.
- 4 Se multiplica la parte decimal de P por d .
- 5 El resultado anterior se le suma al dato X_i .

Medidas Estadísticas: *Posición*

Percentiles

En el caso de una variable continua, el intervalo donde se encuentra $P_k \in (L_{i-1}, L_i]$, se calcula buscando el que deja debajo de si al $k\%$ de las observaciones.

$$P_k = L_{i-1} + \left(\frac{n \frac{k}{100} - N_{i-1}}{n_i} \right) \cdot a_i$$



Medidas Estadísticas: *Posición*

Cuartiles

Los cuartiles, Q_l , son un caso particular de los percentiles. Hay 3, y se definen como:

$$Q_1 = P_{25}$$

$$Q_2 = P_{50} = \tilde{X}$$

$$Q_3 = P_{75}$$



Se definen los deciles como los valores de la variable que dividen a las observaciones en 10 grupos de igual tamaño, más precisamente definiremos $D_1, D_2, \dots, D_9$ como:

$$D_i = P_{10:i} \quad i = 1, \dots, 9$$

Medidas Estadísticas: *Posición*

Quintiles

Se definen los quintiles como los valores de la variable que dividen a las observaciones en 5 grupos de igual tamaño, más precisamente definimos $K_1, K_2, \dots, K_4$ como:

$$K_i = P_{20 \cdot i} \quad i = 1, \dots, 4$$



Medidas Estadísticas: *Forma*

- Simetría:** Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución con respecto a un eje.
- Curtosis:** La curtosis es una medida estadística que determina el grado de concentración que presentan los valores de una variable alrededor de la zona central de la distribución de frecuencias. También es conocida como medida de apuntamiento.



Medidas Estadísticas: *Forma*

Simetría

Una distribución es simétrica cuando a la derecha y a la izquierda de la media existe el mismo número de valores, equidistantes dos a dos de la media, y además con la misma frecuencia. Una distribución es Simétrica si:

$$\overline{X} = \widetilde{X} = \dot{X}$$

En caso contrario, decimos que la distribución es Asimétrica, y entonces puede ser de dos tipos:

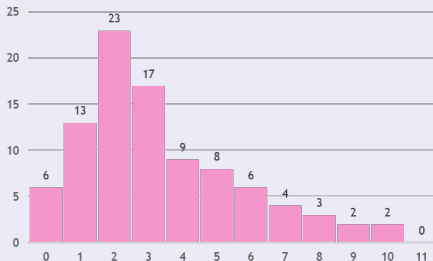


Universidad de
Playa Ancha
Facultad de Ciencias
Naturales y Exactas

DEPARTAMENTO
DISCIPLINARIO
DE MATEMÁTICA
Y ESTADÍSTICA

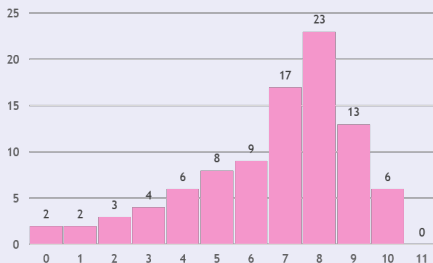
Medidas Estadísticas: *Forma*

Asimetría Positiva



Medidas Estadísticas: *Forma*

Asimetría Negativa



Medidas Estadísticas: *Forma*

Coeficientes de Asimetría de Bowley

Se basa en las distancias entre los cuartiles a fin de establecer un resumen de la asimetría de la distribución. La fórmula es la siguiente:

$$AsB = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

- Si $AsB < 0 \rightarrow$ Asimetría Negativa
- Si $AsB = 0 \rightarrow$ Simetría
- Si $AsB > 0 \rightarrow$ Asimetría Positiva

Medidas Estadísticas: *Forma*

Coefficientes de Asimetría de Pearson

El coeficiente de asimetría de Pearson mide la diferencia entre la media y la moda respecto a la dispersión del conjunto. La fórmula es la siguiente:

$$AsP = \frac{\bar{X} - \hat{X}}{S}$$

- Si $AsP < 0 \rightarrow$ Asimetría Negativa
- Si $AsP = 0 \rightarrow$ Simetría
- Si $AsP > 0 \rightarrow$ Asimetría Positiva

Medidas Estadísticas: *Forma*

Coefficientes de Asimetría de Fisher

El coeficiente de asimetría de Fisher evalúa la proximidad de los datos a su media. La fórmula de la asimetría de Fisher es:

$$AsF = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{S^3}$$

- Si $AsF < 0 \rightarrow$ Asimetría Negativa
- Si $AsF = 0 \rightarrow$ Simetría
- Si $AsF > 0 \rightarrow$ Asimetría Positiva

Medidas Estadísticas: *Forma*

Curtosis

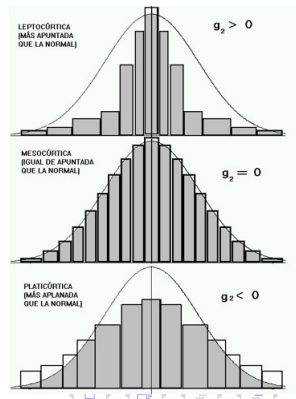
La Curtosis expresa el grado de apuntamiento que presenta la distribución en comparación con la Distribución Normal y se agrupan en 3 grandes categorías de apuntamiento:

- Platicúrtica
- Mesocúrtica
- Leptocúrtica



Medidas Estadísticas: *Forma*

- 1 **Distribución Leptocúrtica:** indica que en las colas hay menos casos acumulados que en las colas de la distribución Normal.
- 2 **Distribución Mesocúrtica:** como en la distribución normal.
- 3 **Distribución Platicúrtica:** indica que en las colas hay más casos acumulados que en las colas de la distribución Normal.



Medidas Estadísticas: *Forma*

Coeficientes de Curtosis de Fisher

$$CcF = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{S^4} - 3$$

- Si $CcF < 0 \rightarrow$ Platicúrtica
- Si $CcF = 0 \rightarrow$ Mesocúrtica
- Si $CcF > 0 \rightarrow$ Leptocúrtica