

Bioestadística

5^{to} Semestre

Lic. En Terapia Física

Profesora:

Dra. en C. Nallely P. Jiménez Mancilla

Descriptiva

Parte de la estadística que trata del análisis y la representación de los datos

Características

- Medidas de tendencia central (promedio, media, mediana, moda, desviación estándar)
- Representación gráfica.

Propósitos

- Utilización de medidas para describir los datos obtenidos.
- Análisis de los datos obtenidos

Situaciones de aplicación (ejemplos)

- Resultados de encuestas.
- Resultados deportivos.
- Porcentajes.

Inferencial

Parte de la estadística por medio de la cual se deduce información o relaciones entre variables

Características

- Relaciones entre variables.
- Generalización a partir de una muestra.
- Utilización de la probabilidad.

Propósitos

- Obtener relaciones entre variables.
- Deducir información.

Situaciones de aplicación

- Predicciones y pronósticos.
- Relaciones entre variables.

ESTADÍSTICA BÁSICA

DISEÑO DE EXPERIMENTOS

ESTADÍSTICA DESCRIPTIVA

TIPOS DE VARIABLES

TABLAS Y GRÁFICAS

MEDIDAS DE POSICIÓN CENTRAL Y DE DISPERSIÓN

ESTADÍSTICA INFERENCIAL

ESTIMACIÓN

CONTRASTE DE HIPÓTESIS

PUNTUAL

POR INTERVALOS

MÉTODOS PARAMÉTRICOS

MÉTODOS NO PARAMÉTRICOS

T-STUDENT

ANOVA

CORRELACIÓN PEARSON

REGRESION LINEAL

Ji-CUADRADA

CORRELACION SPEARMAN

TABLAS DE CONTINGENCIA

ESTADÍSTICA DESCRIPTIVA

Descriptiva

Parte de la estadística que trata del análisis y la representación de los datos

Características

- Medidas de tendencia central (promedio, media, mediana, moda, desviación estándar)
- Representación gráfica.

Propósitos

- Utilización de medidas para describir los datos obtenidos.
- Análisis de los datos obtenidos

Situaciones de aplicación (ejemplos)

- Resultados de encuestas.
- Resultados deportivos.
- Porcentajes.

Gráficos estadísticos

Se denominan **gráficos** a aquellas imágenes que, combinando la utilización de: sombreado, colores, puntos, líneas, símbolos, números, texto y un sistema de referencia (coordenadas), **permiten presentar información cuantitativa o cualitativa.**

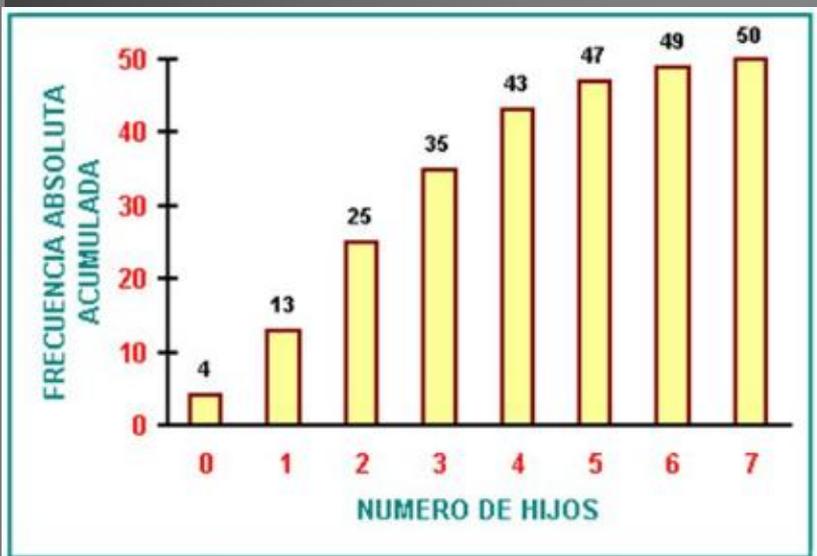
La utilidad de los gráficos:

- Pueden servir como sustituto a las tablas.
- También constituyen por sí mismos una poderosa herramienta para el análisis de los datos.

En ocasiones los gráficos suelen ser el medio más efectivo **no sólo para describir y resumir la información, sino también para analizarla.**

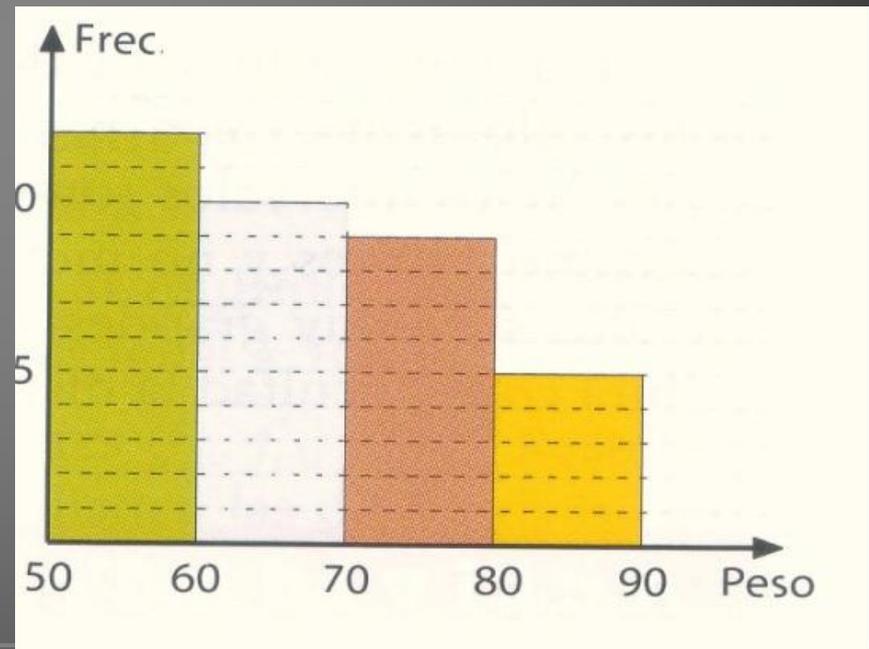
Diagrama de barras

En él se asocia a cada valor de la variable una barra, cuya longitud es igual o proporcional a su frecuencia.



Histograma

Está formado por rectángulos, cuyas bases corresponden con los intervalos de clase y sus áreas son iguales o proporcionales a sus frecuencias.



Polígono de frecuencias

Es una línea poligonal que une los vértices superiores de las barras de un diagrama de barras, o los puntos medios de las bases superiores de los rectángulos de un histograma.

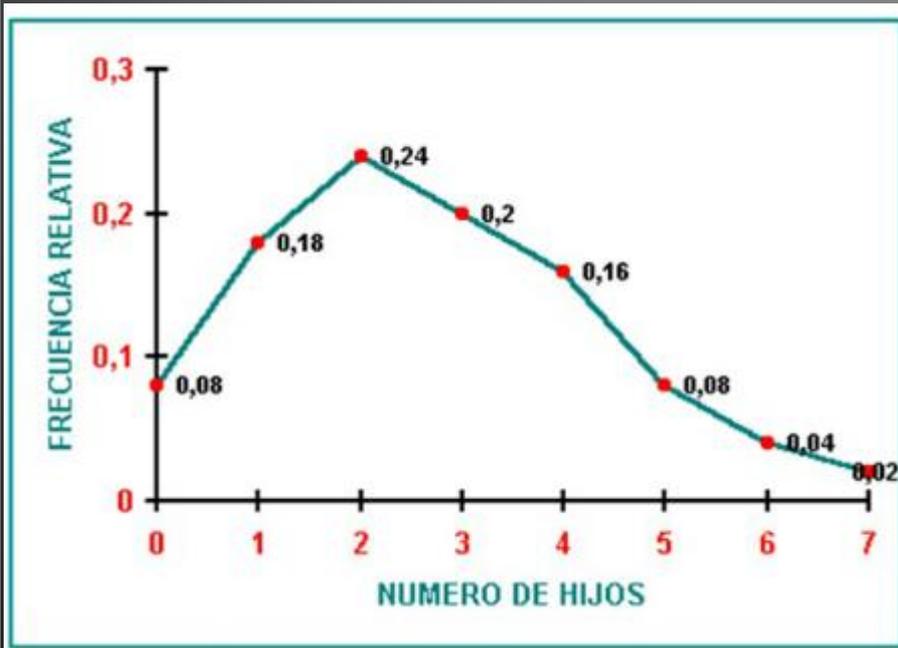
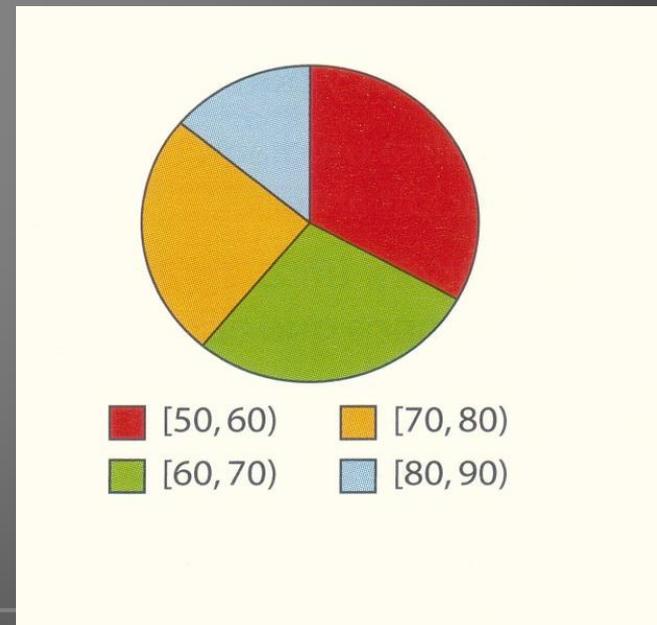


Diagrama de sectores

Es un gráfico formado por un círculo dividido en sectores circulares cuyas amplitudes son proporcionales a las frecuencias de los datos representados.



Pictogramas

Los pictogramas son gráficos similares a los gráficos de barras, pero empleando un dibujo en una determinada escala para expresar la unidad de medida de los datos. Generalmente este dibujo debe cortarse para representar los datos. Se usan para lograr el interés masivo del público.

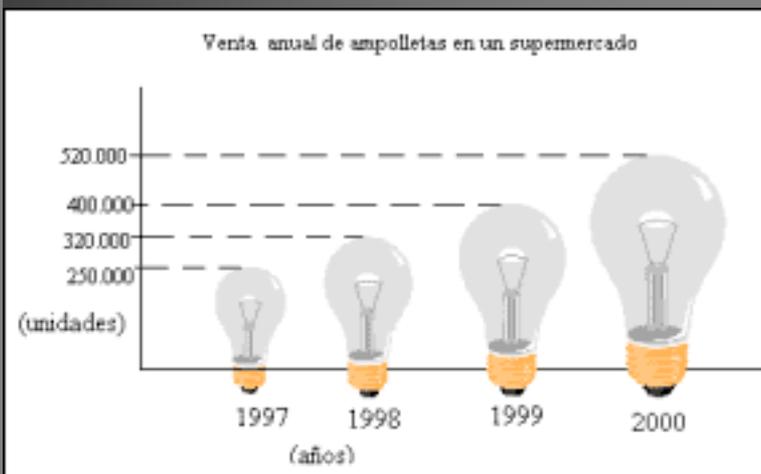
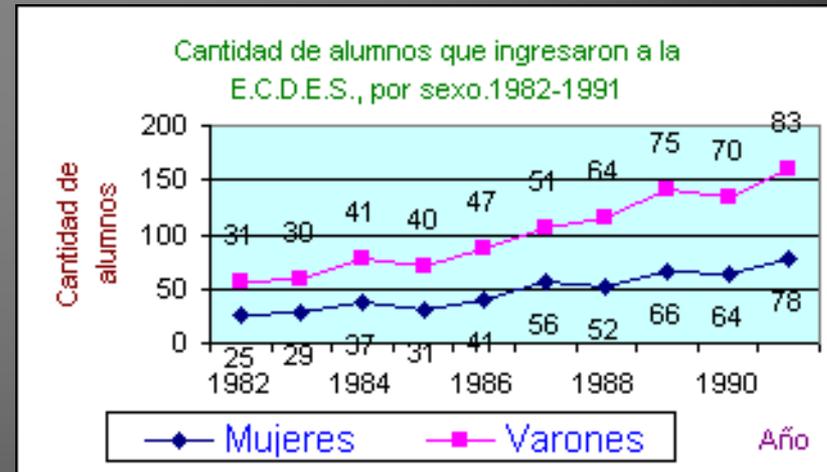


Gráfico de líneas

En este tipo de gráfico se representan los valores de los datos en dos ejes cartesianos ortogonales entre sí.

Se pueden usar para representar: una serie o más series



Gráficos estadísticos

La información contenida en las **tablas de frecuencias** resulta más accesible y fácil de interpretar si se representan por medio de gráficos estadísticos.

Una Tabla de frecuencia es un formato tabular en la que se organizan **los datos en clases**, es decir, en grupos de valores que describen una característica de los datos y **muestra el número de observaciones del conjunto de datos que caen en cada una de las clases.**

- Frecuencia Absoluta
- Frecuencia Absoluta Acumulada
- Frecuencia Relativa
- Frecuencia Relativa Acumulada

Frecuencia Absoluta f_i

- La frecuencia absoluta es el número de veces que aparece un valor (x_i) en los datos obtenidos.
- En nuestro ejemplo, la frecuencia absoluta indica el número de familias que tienen esa cantidad de hijos:

x_i	f_i
0	4
1	9
2	12
3	10
4	8
5	4
6	2
7	1

Numero total de datos n

Frecuencia Absoluta Acumulada F_i

- La frecuencia absoluta acumulada indica cuantos elementos de la lista de datos son menores o iguales a un valor dado. Es la suma de las frecuencias absolutas desde la primera fila hasta la fila elegida.

x_i	f_i	F_i
0	4	4
1	9	13
2	12	25
3	10	35
4	8	43
5	4	47
6	2	49
7	1	50

Frecuencia Relativa h_i

Es el cociente entre la frecuencia Absoluta (f_i) y el número total de datos (n).

X_i	f_i	F_i	h_i
0	4	4	0.08
1	9	13	0.18
2	12	25	0.24
3	10	35	0.2
4	8	43	0.16
5	4	47	0.08
6	2	49	0.04
7	1	50	0.02
	n= 50		

Frecuencia Relativa Acumulada H_i

Es el cociente de la frecuencia absoluta acumulada (F_i) y el número total de datos (n).

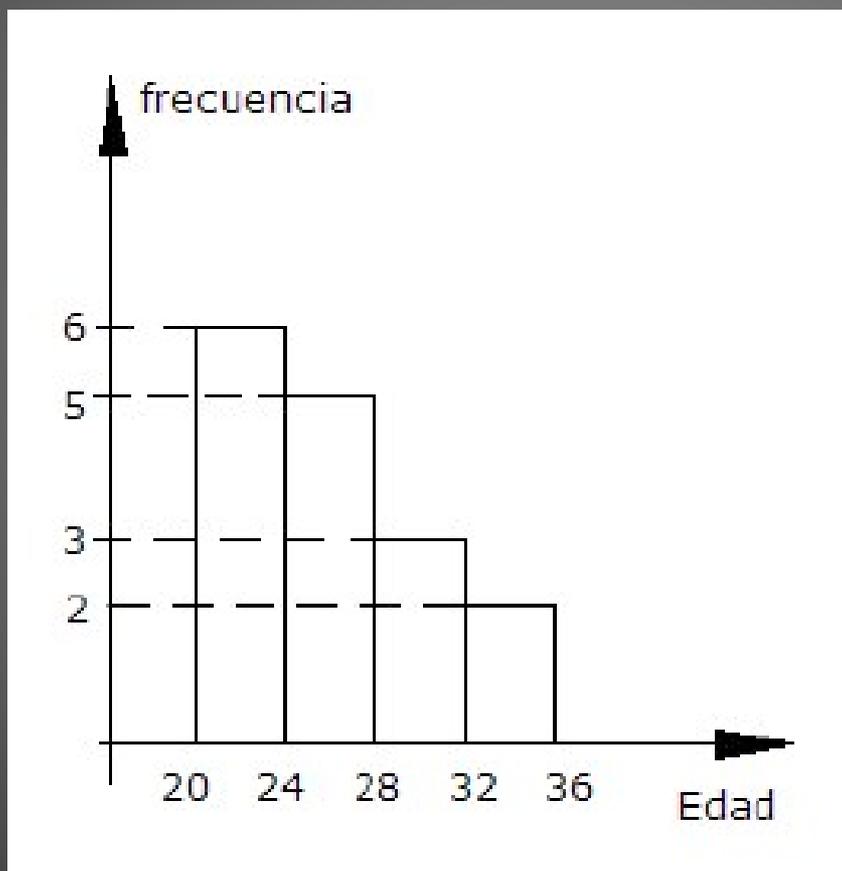
X_i	f_i	F_i	h_i	H_i
0	4	4	0.08	0.08
1	9	13	0.18	0.26
2	12	25	0.24	0.5
3	10	35	0.2	0.7
4	8	43	0.16	0.86
5	4	47	0.08	0.94
6	2	49	0.04	0.98
7	1	50	0.02	1
	n= 50			

Tabla de frecuencias

Valor (X_i)	Frecuencia Absoluta (f_i)	Frecuencia Absoluta Acumulada (F_i)	Frecuencia Relativa (h_i)	Frecuencia Relativa Acumulada (H_i)
X_1	f_1	$F_1=f_1$	$h_1=f_1/N$	$H_1=h_1$
X_2	f_2	$F_2=F_1+f_2$	$h_2=f_2/N$	$H_2=H_1+h_2$
X_3	f_3	$F_3=F_2+f_3$	$h_3=f_3/N$	$H_3=H_2+h_3$
.
.
.
X_K	f_K	$F_K=N$	$h_K=f_K/N$	$H_K=1$
Total de datos	N	-	1	-

Ejercicio:

Construya una tabla de frecuencia de los siguientes gráficos.



2) Construya una gráfica de barras

i) En una caja hay 10 bolitas marcadas con los números del 1 al 4. En la siguiente tabla se muestra la distribución de frecuencias de cada uno de los números:

Número	Frecuencia
1	3
2	1
3	1
4	5



Medidas de Tendencia Central o Posición

Las medidas de tendencia central más importantes son:

- **Media Aritmética o Promedio.**
- **Mediana.**
- **Moda.**

Media Aritmética o Promedio

- Es la suma de todas las observaciones dividida entre el número total de observaciones (n).

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

Mediana

- Es el valor que ocupa la posición central de un conjunto de observaciones, una vez que han sido ordenados en forma ascendente o descendente.
- Divide al conjunto de datos en dos partes iguales.

Calculo de la Mediana:

- Si n es impar: posición donde se ubica la mediana es igual a $(n+1)/2$.
- Si n es par: $(n+1)/2$ no es entero, por lo tanto la mediana será igual al promedio de las dos posiciones centrales.

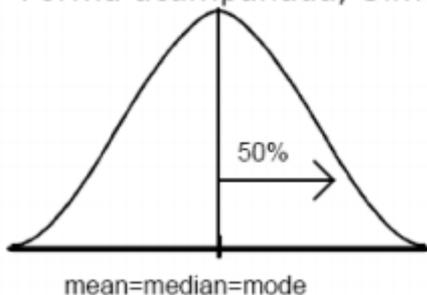
Moda

- Observación o clase que tiene la mayor frecuencia en un conjunto de observaciones.
- Un conjunto de datos puede ser unimodal, bimodal o multimodal.
- Es la única medida de tendencia central que se puede determinar para datos de tipo cualitativo.

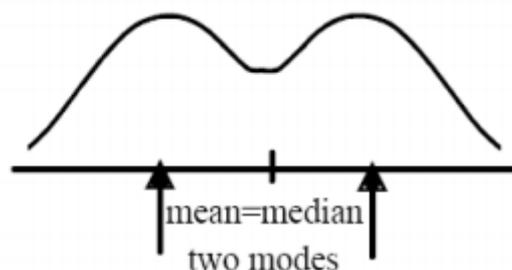
Relación entre Media, Mediana y Moda

- Si **media=moda=mediana**, la distribución es simétrica
- Si **media > mediana**, la distribución es asimétrica con cola a la derecha (sesgada a la derecha).
- Si **media < mediana**, la distribución es asimétrica con cola a la izquierda (sesgada a la izquierda).

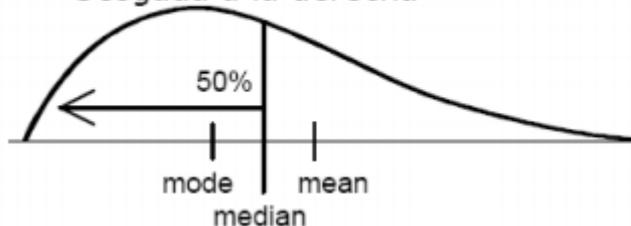
Forma acampanada, Simétrica



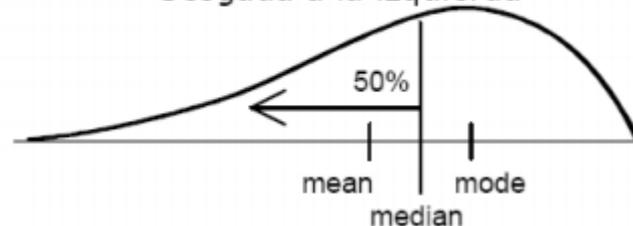
Bimodal



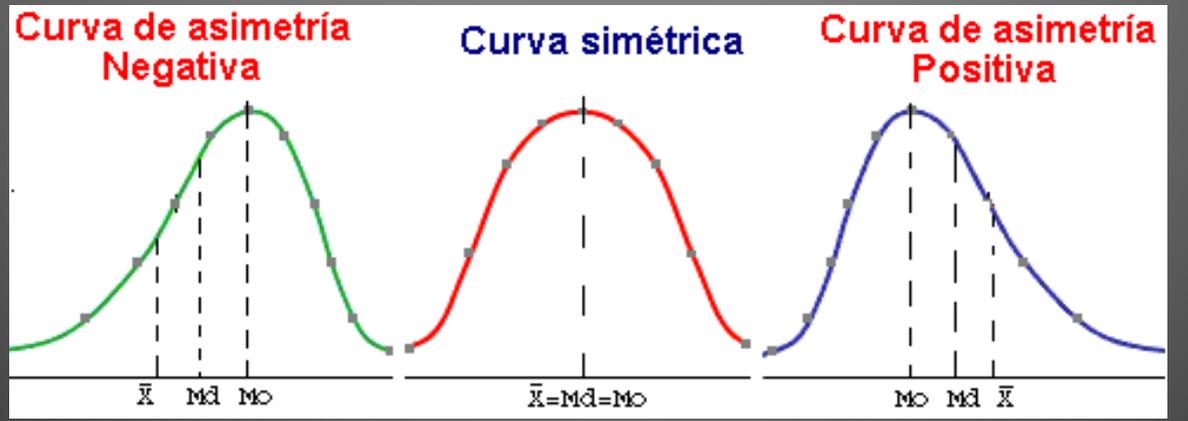
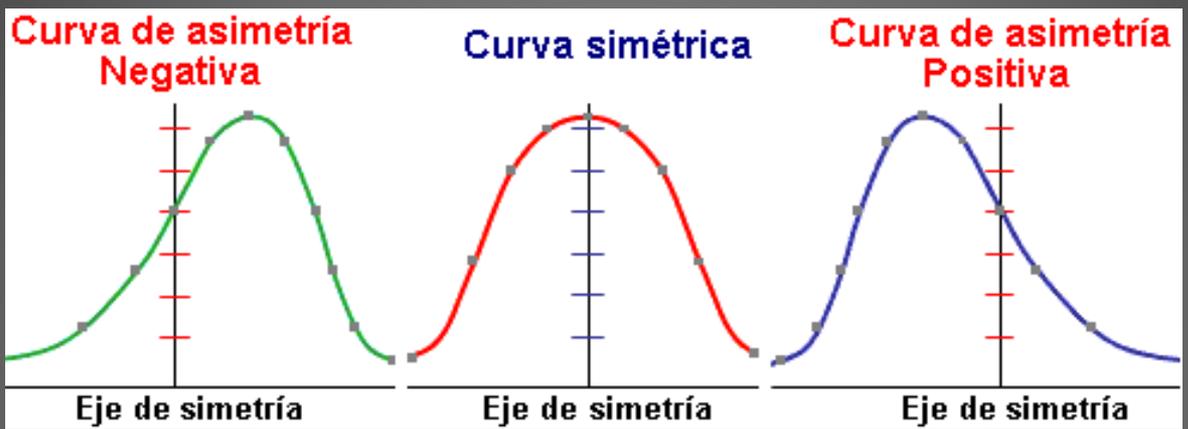
Sesgada a la derecha



Sesgada a la izquierda

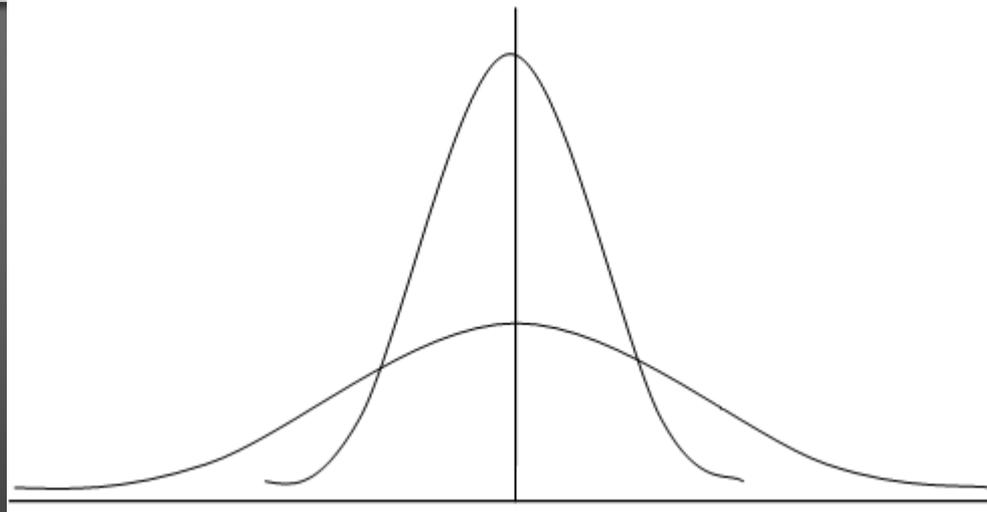


Relación entre Media, Mediana y Moda



Medidas de dispersión, variación o variabilidad

- Son valores numéricos que indican o describen la forma en que las observaciones están dispersas o diseminadas, con respecto al valor central.
- Son importantes debido a que dos muestras de observaciones con el mismo valor central pueden tener una variabilidad muy distinta.



**Medidas de dispersión, variación o
variabilidad**

- **Rango.**
- **Varianza.**
- **Desviación Típica o
Estándar.**
- **Coeficiente de variación.**

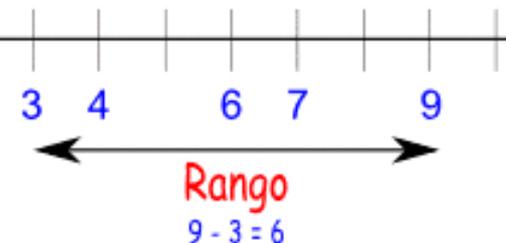
Medidas de dispersión, variación o variabilidad

➤ Rango

Una manera de medir la dispersión es calcular la amplitud de la muestra, es decir, la diferencia entre las observaciones máxima y mínima.

$$Rango = X_{max} - X_{min}$$

Su mayor ventaja es que se puede calcular fácilmente, sin embargo, no brinda información sobre la dispersión existente entre ambos valores extremos.



➤ Varianza (s^2)

Mide la distancia entre los valores de la muestra y la media (promedio \bar{X})

La **varianza** s^2 se calcula como el promedio de los cuadrados de la desviaciones de las observaciones respecto a su media. Formalmente:

$$s^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n - 1}$$

De forma compacta:

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

 Varianza

La varianza siempre será mayor que **CERO**.

Mientras la varianza mas se aproxime a CERO, nos indica que los valores de la muestra se aproximan mas a la media. Y mientras mas se alejen de CERO, nos indica que las medidas de la muestra están mas dispersas.

➤ Desviación Estándar (s)

Nos indica cuanto tienden a alejarse los valores puntuales de la muestra del promedio.

Si $s = 0$ solo ocurre cuando no hay dispersión: todas las observaciones toman el mismo valor. De lo contrario $s > 0$.

Cuanto más dispersión hay entre las observaciones, mayor es s .

Matemáticamente:

$$s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

➤ **Coeficiente de Variación (*CV*)**

Es una medida que describe la variabilidad relativa con respecto a la media aritmética.

Se usa principalmente para comparar las desviaciones estándar de dos muestras con diferentes medias.

$$CV = \frac{S}{\bar{X}}$$

Ejercicio

Calcule las medidas centrales y las medidas de dispersión de la siguiente tabla de frecuencias:

X_i	f_i
0	4
1	9
2	12
3	10
4	8
5	4
6	2
7	1

Variable

Frecuencias

Datos totales:

$n = 50$

El gráfico circular de la figura muestra las preferencias de 30 alumnos en actividades deportivas. ¿Cuál(es) de las siguientes afirmaciones es(son) correcta(s) ?

- I) La frecuencia relativa del grupo de fútbol es de 40%.
- II) La frecuencia relativa del grupo de básquetbol es de 30%.
- III) La mitad del grupo no prefirió fútbol ni tenis.



Observación: Construye la tabla frecuencias correspondiente

ESTADÍSTICA DESCRIPTIVA



Deporte	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa %
Fútbol	12	0.40	40
Tenis	3	0.10	10
Atletismo	6	0.20	20
Básquetbol	9	0.30	30
Total	30	1	100

- I) La frecuencia relativa del grupo de fútbol es de 40%. (Correcta)
- II) La frecuencia relativa del grupo de básquetbol es de 30%. (Correcta)
- III) La mitad del grupo no prefirió fútbol ni tenis. (Correcta)