

ESTADÍSTICA

Definición de Estadística

La **Estadística** trata del recuento, ordenación y clasificación de los datos obtenidos por las observaciones, para poder hacer comparaciones y sacar conclusiones.

Conceptos de Estadística

Población

Una **población** es el conjunto de todos los elementos a los que se somete a un estudio estadístico.

Individuo

Un **individuo** o **unidad estadística** es cada uno de los elementos que componen la población.

Muestra

Una **muestra** es un conjunto representativo de la población de referencia, el número de individuos de una muestra es menor que el de la población.

Muestreo

El **muestreo** es la reunión de datos que se desea estudiar, obtenidos de una proporción reducida y representativa de la población.

Valor

Un **valor** es cada uno de los distintos resultados que se pueden obtener en un estudio estadístico. Si lanzamos una moneda al aire 5 veces obtenemos dos valores: cara y cruz.

Dato

Un **dato** es cada uno de los valores que se ha obtenido al realizar un estudio estadístico. Si lanzamos una moneda al aire 5 veces obtenemos 5 datos: cara, cara, cruz, cara, cruz.

Variables estadísticas

Variable cualitativa

Las **variables cualitativas** se refieren a **características o cualidades** que **no** pueden ser medidas con **números**. Podemos distinguir dos tipos:

Variable cualitativa nominal

Una **variable cualitativa nominal** presenta **modalidades no numéricas** que **no** admiten un **criterio de orden**.

Variable cualitativa ordinal o variable cuasicuantitativa

Una **variable cualitativa ordinal** presenta **modalidades no numéricas**, en las que existe un **orden**.

Variable cuantitativa

Una **variable cuantitativa** es la que se expresa mediante un **número**, por tanto se pueden realizar **operaciones aritméticas** con ella. Podemos distinguir dos tipos:

Variable discreta

Una **variable discreta** es aquella que toma **valores aislados**, es decir **no** admite **valores intermedios** entre dos valores específicos.

Variable continua

Una **variable continua** es aquella que puede tomar **valores comprendidos entre dos números**.

Distribución de frecuencias

La **distribución de frecuencias** o **tabla de frecuencias** es una **ordenación** en forma de **tabla** de los **datos estadísticos**, asignando a cada **dato** su **frecuencia correspondiente**.

Diagrama de barras

Un **diagrama de barras** se utiliza para de presentar **datos cualitativos** o **datos cuantitativos de tipo discreto**.

Los **datos** se representan mediante **barras** de una **altura proporcional** a la **frecuencia**.

Polígonos de frecuencias

Un **polígono de frecuencias** se forma uniendo los **extremos** de las **barras** mediante **segmentos**.

También se puede realizar trazando los **puntos** que representan las **frecuencias** y uniéndolos mediante **segmentos**.

Diagrama de sectores

Un **diagrama de sectores** se puede utilizar para todo tipo de *variables*, pero se usa frecuentemente para las **variables cualitativas**.

Los **datos** se representan en un **círculo**, de modo que el **ángulo** de cada **sector**

es **proporcional** a la **frecuencia absoluta** correspondiente.

$$\alpha = \frac{360^\circ}{N} \cdot f_i$$

Histograma

Un **histograma** es una **representación gráfica** de una **variable** en forma de **barras**.

Se utilizan para **variables continuas** o para **variables discretas**, con un gran número de datos, y que se han agrupado en **clases**.

En el **eje abscisas** se construyen unos **rectángulos** que tienen por **base la amplitud del intervalo**, y por **altura**, la **frecuencia absoluta** de cada **intervalo**.

Medidas de centralización

Moda

La **moda** es el **valor** que tiene **mayor frecuencia absoluta**.

Se representa por **M_o**.

Se puede hallar la **moda** para **variables cualitativas** y **cuantitativas**.

Cálculo de la moda para datos agrupados

1º Todos los intervalos tienen la misma amplitud.

$$M_o = L_{i-1} + \frac{f_{i+1}}{f_{i-1} + f_{i+1}} \cdot a_i$$

2º Los intervalos tienen amplitudes distintas.

En primer lugar tenemos que hallar las alturas.

$$h_i = \frac{f_i}{a_i}$$

La clase modal es la que tiene mayor altura.

$$M_o = L_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} \cdot a_i$$

Mediana

Es el **valor** que ocupa el **lugar central** de todos los **datos** cuando éstos están **ordenados de menor a mayor**.

La **mediana** se representa por **M_e**.

La **mediana** se puede **hallar** sólo para **variables cuantitativas**.

Cálculo de la mediana

1 Ordenamos los datos de menor a mayor.

2 Si la serie tiene un **número impar de medidas** la **mediana** es la **puntuación central** de la misma.

3 Si la serie tiene un **número par** de puntuaciones la **mediana** es la **media** entre las dos **puntuaciones centrales**.

Cálculo de la mediana para datos agrupados

La **mediana** se encuentra en el **intervalo** donde la **frecuencia acumulada** llega hasta la **mitad de la suma de las frecuencias absolutas**.

Es decir tenemos que buscar el intervalo en el que se encuentre.

$$Me = L_{i-1} + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$$

Media aritmética

La **media aritmética** es el **valor** obtenido al **sumar** todos los **datos** y **dividir** el resultado entre el **número** total de **datos**.

\bar{x} es el símbolo de la **media aritmética**.

$$\bar{x} = \frac{\sum_{i=1}^n X_i}{N}$$

Media aritmética para datos agrupados

Si los **datos** vienen **agrupados** en una tabla de frecuencias, la expresión de la **media** es:

$$\bar{x} = \frac{\sum_{i=1}^n X_i f_i}{N}$$

Medidas de dispersión

Desviación media

La **desviación media** es la **media aritmética** de los **valores absolutos de las desviaciones respecto a la media**.

La **desviación media** se representa por D_x

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

Desviación media para datos agrupados

Si los datos vienen agrupados en una **tabla de frecuencias**, la expresión de la **desviación media** es:

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

Varianza

La **desviación típica** es la **raíz cuadrada de la varianza**.

La **desviación típica** se representa por σ .

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

Desviación típica para datos agrupados

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

Para simplificar el cálculo vamos a utilizar las siguientes expresiones que son equivalentes a las anteriores.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n \frac{x_i^2 f_i}{N} - \bar{x}^2}$$

Desviación típica para datos agrupados

$$\sigma = \sqrt{\sum_{i=1}^n \frac{X_i^2}{N} - \bar{X}^2}$$

Desviación típica

La **desviación típica** es la **raíz cuadrada de la varianza**.

La **desviación típica** se representa por σ .

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{N}}$$

Desviación típica para datos agrupados

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 f_i}{N}}$$

Para simplificar el cálculo vamos a utilizar las siguientes expresiones que son equivalentes a las anteriores.

Desviación típica para datos agrupados

Coefficiente de variación

El **coeficiente de variación** es la relación entre la **desviación típica** de una muestra y su **media**.

$$C.V = \frac{\sigma}{\bar{X}} \cdot 100$$

Puntuaciones típicas

Las **puntuaciones típicas** son el resultado de **dividir** las **puntuaciones diferenciales** entre la **desviación típica**. Este proceso se llama **tipificación**.

$$Z = \frac{X_i - \bar{X}}{\sigma}$$

ESTADÍSTICA

1. Indica que **variables** son **cualitativas** y cuales **cuantitativas**:

- 1 Comida Favorita.
- 2 Profesión que te gusta.
- 3 Número de goles marcados por tu equipo favorito en la última temporada.
- 4 Número de alumnos de tu Instituto.
- 5 El color de los ojos de tus compañeros de clase.
- 6 Coeficiente intelectual de tus compañeros de clase.

2. De las siguientes **variables** indica cuáles son **discretas** y cuales **continuas**.

- 1 Número de acciones vendidas cada día en la Bolsa.
- 2 Temperaturas registradas cada hora en un observatorio.
- 3 Período de duración de un automóvil.
- 4 El diámetro de las ruedas de varios coches.
- 5 Número de hijos de 50 familias.
- 6 Censo anual de los españoles.

3. Clasificar las siguientes **variables** en **cualitativas** y **cuantitativas discretas** o **continuas**.

- 1 La nacionalidad de una persona.
- 2 Número de litros de agua contenidos en un depósito.
- 3 Número de libro en un estante de librería.
- 4 Suma de puntos tenidos en el lanzamiento de un par de dados.
- 5 La profesión de una persona.
- 6 El área de las distintas baldosas de un edificio.

4. Las puntuaciones obtenidas por un grupo de en una prueba han sido:

15, 20, 15, 18, 22, 13, 13, 16, 15, 19, 18, 15, 16, 20, 16, 15, 18, 16, 14, 13.

Construir la **tabla de distribución de frecuencias** y dibuja el **polígono de frecuencias**.

5. El número de estrellas de los hoteles de una ciudad viene dado por la siguiente serie:

3, 3, 4, 3, 4, 3, 1, 3, 4, 3, 3, 3, 2, 1, 3, 3, 3, 2, 3, 2, 2, 3, 3, 3, 2, 2, 2, 2, 3, 2, 1, 1, 1, 2, 2, 4, 1.

Construir la tabla de distribución de frecuencias y dibuja el diagrama de barras.

6. Las calificaciones de 50 alumnos en Matemáticas han sido las siguientes:

5, 2, 4, 9, 7, 4, 5, 6, 5, 7, 7, 5, 5, 2, 10, 5, 6, 5, 4, 5, 8, 8, 4, 0, 8, 4, 8, 6, 6, 3, 6, 7, 6, 6, 7, 6, 7, 3, 5, 6, 9, 6, 1, 4, 6, 3, 5, 5, 6, 7.

Construir la **tabla de distribución de frecuencias** y dibuja el **diagrama de barras**.

7. Los pesos de los 65 empleados de una fábrica vienen dados por la siguiente tabla:

Peso	[50, 60)	[60, 70)	[70, 80)	[80,90)	[90, 100)	[100, 110)	[110, 120)
f_i	8	10	16	14	10	5	2

1 Construir la **tabla de frecuencias**.

2 Representar el **histograma** y el **polígono de frecuencias**.

8. Los 40 alumnos de una clase han obtenido las siguientes puntuaciones, sobre 50, en un examen de Física.

3, 15, 24, 28, 33, 35, 38, 42, 23, 38, 36, 34, 29, 25, 17, 7, 34, 36, 39, 44, 31, 26, 20, 11, 13, 22, 27, 47, 39, 37, 34, 32, 35, 28, 38, 41, 48, 15, 32, 13.

1 Construir la **tabla de frecuencias**.

2 Dibujar el **histograma** y el **polígono de frecuencias**.

9. Sea una distribución estadística que viene dada por la siguiente tabla:

x_i	61	64	67	70	73
f_i	5	18	42	27	8

Calcular:

1 La **moda**, **mediana** y **media**.

2 El **rango**, **desviación media**, **varianza** y **desviación típica**.

10. Calcular la **media**, la **mediana** y la **moda** de la siguiente serie de números: 5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4.

11 Hallar la **varianza** y la **desviación típica** de la siguiente serie de datos:
12, 6, 7, 3, 15, 10, 18, 5.

12 Hallar la **media**, **mediana** y **moda** de la siguiente serie de números:
3, 5, 2, 6, 5, 9, 5, 2, 8, 6.

13. Hallar la **desviación media**, la **varianza** y la **desviación típica** de la series de números siguientes:

2, 3, 6, 8, 11.

12, 6, 7, 3, 15, 10, 18, 5.

14 Se ha aplicado test a los empleados de una fábrica, obteniéndose las siete tabla:

	f_i
[38, 44)	7
[44, 50)	8
[50, 56)	15
[56, 62)	25
[62, 68)	18
[68, 74)	9
[74, 80)	6

Dibujar el **histograma** y el **polígono de frecuencias acumuladas**.

15. Dadas las series estadísticas:

3, 5, 2, 7, 6, 4, 9.

3, 5, 2, 7, 6, 4, 9, 1.

Calcular:

La **moda**, la **mediana** y la **media**.

La **desviación media**, la **varianza** y la **desviación típica**.