

RESUMEN DE CONTENIDOS



ÍNDICE

Llegar a 100... ¡Malo será!.....	2
Población y muestra.....	2
Variables estadísticas.....	2
Tipos de variables.....	2
Frecuencias.....	3
Tipos de frecuencias.....	3
Gráficos estadísticos.....	4
Medidas o parámetros estadísticos.....	5
Tipos de parámetros.....	5
Atribución de los recursos incorporados al documento.....	6

Llegar a 100... ¡Malo será!

Población y muestra

Se entiende por **población** un conjunto de unidades del que se desea obtener cierta información y, por **muestra**, un subconjunto representativo de la misma en la que se desean estudiar al menos una variable estadística.

Ejemplo: en caso de querer estudiar los hábitos de los ancianos gallegos, la población podrían ser todos los mayores de ochenta años residentes en Galicia y la muestra podría ser un grupo reducido de doscientos de ellos escogidos de manera aleatoria.

Variables estadísticas

Una variable estadística es una característica a analizar.

Tipos de variables

Variables cualitativas

Los valores que toma son cualidades. No son medibles.

- **Ejemplo:** color de ojos, comida favorita, estado emocional...

Variables cuantitativas

Los valores que toma son números. Son medibles.

- **Ejemplo:** Edad, peso, talla...

Variables cuantitativas discretas

Toman un número finito o infinito numerable de valores.

- **Ejemplo:** número de hermanos, cantidad de goles, número de alumnos...

Variables cuantitativas continuas

Describen características medibles que pueden tomar cualquier valor dentro de un determinado intervalo.

- **Ejemplo:** peso, temperatura, altura...

Si la **variable** es **continua** se utilizan intervalos para agrupar los distintos valores que toma la variable con el objeto de resumir la información:

Clase: Intervalo en el que se agrupan los datos de una variable estadística cuando el número de datos es grande y pocos datos se repiten.

Marca de clase x_i : Punto medio de la clase: suma de sus extremos y dividida entre 2.

Frecuencias

La frecuencia de un dato es el número de veces que se repite, bien en total o comparado con el total.

Tipos de frecuencias

Frecuencia absoluta

Número de veces que se recoge un dato. La suma de las frecuencias absolutas coincide con el número total de datos N .

$$\sum_{i=1}^n f_i = f_1 + f_2 + \dots + f_n = N$$

Frecuencia relativa

Es el cociente entre la frecuencia absoluta y el número total de datos.

$$h_i = \frac{f_i}{N} \qquad \sum_{i=1}^n h_i = h_1 + h_2 + \dots + h_n = \frac{f_1 + f_2 + \dots + f_n}{N} = \frac{N}{N} = 1$$

Frecuencia absoluta acumulada

Suma de las frecuencias absolutas de un dato y de las de sus anteriores.

$$F_i = \sum_{i=1}^i f_i = f_1 + f_2 + \dots + f_i \qquad F_n = \sum_{i=1}^n f_i = N$$

Frecuencia relativa acumulada

Suma de las frecuencias relativas de un dato y de las de sus anteriores.

$$H_i = \sum_{i=1}^i h_i = h_1 + h_2 + \dots + h_i \qquad H_n = \sum_{i=1}^n h_i = 1$$

Ejemplo de tabla de frecuencias:

Días por semana x_i	Frecuencia absoluta f_i	Frecuencia relativa h_i	Frecuencia abs. acumulada F_i	Frecuencia rel. acumulada H_i
1-2 días	4	$4/10 = 0,4$	4	0,4
3-4 días	8	$8/10 = 0,8$	12	1,20

Gráficos estadísticos

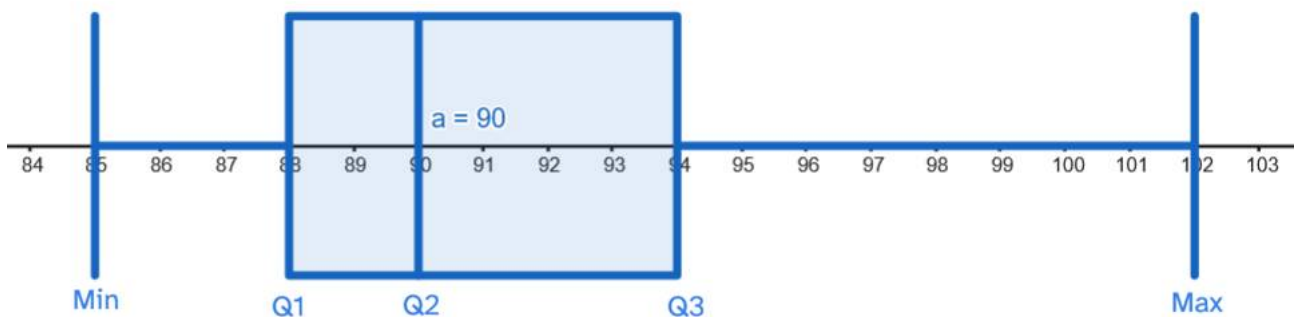
Tipos de gráficos estadísticos

Diagrama de caja (boxplot)

Este gráfico está formado por unas líneas laterales (los bigotes) y una caja entre ellas.

- La primera línea va desde el valor mínimo hasta Q1.
- La caja va desde Q1 hasta Q3.
- La línea final va desde Q3 hasta el valor máximo.

En la imagen puedes ver un ejemplo de gráfico de caja. Las líneas laterales suelen llamarse bigotes.



La altura o grosor de la caja no tiene significado estadístico; su propósito es sólo visual, para facilitar la lectura del gráfico.

Diagrama de barras:

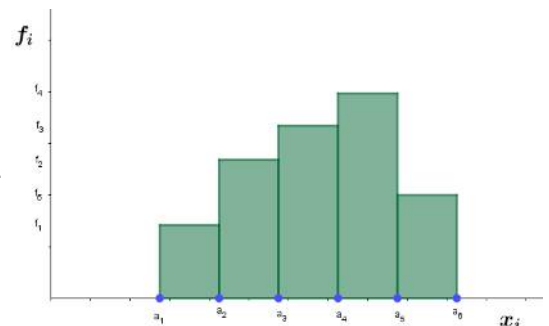
En este gráfico, se representan los valores de la variable en el eje X, y su frecuencia absoluta en el eje Y.

Diagrama de sectores:

En este gráfico se representan los datos en un círculo dividido en sectores cuya amplitud es proporcional a la frecuencia relativa de cada dato.

Histograma:

Los histogramas se usan para representar datos agrupados en clases o intervalos. Es el equivalente al diagrama de barras, con base el tamaño del intervalo, altura su frecuencia y pegadas entre sí.



A veces se unen los puntos superiores de las barras formando una línea llamada polígono de frecuencias.

Medidas o parámetros estadísticos

A partir de los datos de la muestra, se calculan unos valores llamados parámetros o medidas estadísticas.

Sirven para resumir los datos, aportando información sobre su forma, su posición, etc.

Tipos de parámetros

Parámetros de posición

Medidas cuyos valores indican una posición dentro del conjunto de datos.

Los más importantes son los de centralización.

Parámetros de centralización

Media \bar{x}

Cantidad que resulta de dividir la suma de los datos por sus frecuencias absolutas entre el

total de los datos. Su fórmula es: $\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{N}$

Mediana *Me*

Valor que divide al conjunto de datos ordenados en dos partes iguales.

Se representa como **Me**.

Moda *Mo*

Valor o valores de la variable que presentan una mayor frecuencia.

Se representa como **Mo**.

Otras medidas de posición

Cuartiles *Q*

El cuartil (*Q*) divide al conjunto de datos en cuatro partes iguales.

- Q1**: el primer cuartil indica el valor que deja la cuarta parte de los datos (el 25 %) por debajo.

- Q2**: el segundo cuartil indica el valor que deja la mitad de los datos (el 50 %) por debajo. Es la mediana.

- Q3**: el tercer cuartil indica el valor que deja las tres cuartas partes de los datos (el 75 %) por debajo.

Percentiles

Los percentiles dividen al conjunto de datos en cien partes iguales.

Parámetros de dispersión

Rango o recorrido

Es la diferencia entre el mayor y el menor valor de una variable: $R = x_{max} - x_{min}$

Si los datos están agrupados, se calcula restando el extremo superior del último intervalo menos el extremo inferior del primero: $R = L_{max} - L_{min}$

Recorrido intercuartílico

Es la diferencia entre Q3 - Q2.

Desviación

Las desviaciones son los resultados de las diferencias entre cada uno de los valores que toma la variable y la media: $x - \bar{x}$

Desviación absoluta media

La desviación absoluta media es la media de los **valores absolutos** de las desviaciones

con respecto a la media: $D_m = \frac{\sum_{i=1}^n |x_i - \bar{x}| \cdot f_i}{N}$

Varianza

Es la media de los **cuadrados** de las desviaciones de cada uno de los datos con respecto

a la media: $\sigma^2 = Var = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot f_i}{N}$

Desviación típica

También conocida como desviación estándar, es la raíz cuadrada de la varianza: $\sigma = \sqrt{Var}$

Coefficiente de variación

Es una medida de dispersión relativa, ya que se calcula como el cociente entre la desviación típica y la media (siempre que esta última sea no nula), por lo que no tiene

unidades: $CV = \frac{\sigma}{\bar{x}}$

Atribución de los recursos incorporados al documento

Las imágenes que figuran en este documento son de elaboración propia (proyecto cREAgal).



“Resumen de contenidos: Llegar a los 100... ¡Malo será!”, se publica con la [Licencia Creative Commons Reconocimiento No-comercial Compartir igual 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)