

# GLOSARIO

## Media aritmética

Es la medida de tendencia central más utilizada y conocida. Se calcula sumando todos los valores del conjunto de datos y dividiendo el resultado entre el número total de observaciones ( $\frac{\sum x}{n}$ ). Representa el "centro de gravedad" de la distribución, con la particularidad de ser altamente sensible a la presencia de valores atípicos.

## Mediana

Es el valor que ocupa la posición central en un conjunto de datos una vez que estos han sido ordenados de menor a mayor. Divide la distribución exactamente en dos partes iguales, dejando el 50% de los datos por debajo de ella y el otro 50% por encima. A diferencia de la media, la mediana es una medida robusta que no se ve afectada por valores extremos.

## Moda

Es el valor o categoría que se presenta con la mayor frecuencia absoluta en un conjunto de datos; es decir, el dato que más se repite. Una distribución puede carecer de moda (amodal), tener una sola (unimodal) o presentar varias modas (bimodal o multimodal).

## Tendencia central

Concepto que agrupa a los indicadores estadísticos (como la media, mediana y moda), cuyo objetivo es resumir en un solo valor numérico el centro de una distribución de datos. Sirven como punto de referencia para ubicar hacia dónde se inclinan o agrupan la mayoría de las observaciones de la muestra.

## Cuartil

Medidas de posición que dividen a un conjunto de datos ordenados en **cuatro partes porcentualmente iguales** (cada una correspondiente al 25% de las observaciones).

Existen tres cuartiles principales:

- **Q<sub>1</sub>**: Deja por debajo el 25% de los datos.
- **Q<sub>2</sub>**: Coincide exactamente con el valor de la mediana (50%).
- **Q<sub>3</sub>**: Deja por debajo el 75% de los datos.

## Decil

Medidas de posición no central que dividen la distribución de los datos ordenados en **diez partes iguales**, donde cada intervalo representa el 10% del total de las observaciones. Se denotan desde  $D_1$  (que acumula el 10% inferior) hasta  $D_9$  (que acumula el 90%).

## Percentil

Medidas de posición que dividen el conjunto de datos ordenados en **cien partes porcentualmente iguales**. El percentil  $k$  ( $P_k$ ) indica el valor por debajo del cual se encuentra el  $k\%$  de las observaciones de la muestra. Por ejemplo, el percentil 85 ( $P_{85}$ ) supera al 85% de los datos del grupo.

## Valor atípico (Outlier)

*(Nota: Es sinónimo de "Valor extremo").* Hace referencia a cualquier observación numérica que se aleja de manera desproporcionada y sospechosa del patrón general del resto de los datos. Su presencia suele distorsionar el cálculo de la media aritmética y las medidas de dispersión.

## Distribución

Es la representación (gráfica, tabular o matemática) que muestra cómo se organizan, reparten o distribuyen las frecuencias de los distintos valores que toma una variable estadística a lo largo de su rango de variación.

## Posición relativa

Se refiere a la ubicación que ocupa un valor específico dentro de un conjunto de datos en comparación con el resto de las observaciones. Herramientas como los cuartiles, percentiles y las puntuaciones estandarizadas son las que permiten determinar con exactitud dicha jerarquía o estatus de un dato frente al grupo.