

MEDIDAS DE CENTRALIZACIÓN

a) Media aritmética: \bar{x}

La media aritmética es el valor promedio aritmético de todas las muestras.

Si las tuviéramos una a una, las sumaríamos todas y las dividiríamos por el número de muestras.

Como tendremos los datos agrupados en una tabla de frecuencias, el cálculo equivalente

al anterior es:
$$\bar{x} = \frac{\sum x_i \cdot f_i}{N}$$

b) Moda: M_o

Es el valor que más se repite, o sea, el de mayor frecuencia.

c) Mediana: M_e

Con todas las muestras ordenadas de menor a mayor, la Mediana es el valor de la muestra que ocupa la posición central.

Si el nº de muestras es impar, M_e =valor de la muestra central.

Si el nº de muestras es par, no hay una muestra central, así que hay que hacer una media de los dos centrales: $M_e = \frac{\text{suma de los dos valores que están en las dos posiciones centrales}}{2}$

MEDIDAS DE DISPERSIÓN

Para entender la distribución de los datos con los que estamos trabajando, necesitamos estudiar además de la centralización de ellos, su dispersión, o sea, si se concentran alrededor de la media o están bastante alejados.

c) Rango o recorrido:

El rango es la diferencia entre el mayor y el menor valor de la variable. Indica la longitud del intervalo en el que se encuentran todas las muestras.

d) Varianza: V ó σ^2

La varianza es el promedio de los cuadrados de las distancias de las muestras a la media.

Se pueden aplicar dos fórmulas equivalentes para su cálculo:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{N} \quad \text{ó}$$

$$\sigma^2 = \frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2$$

Con las tablas de frecuencia es más sencillo calcularla con esta segunda fórmula.

e) Desviación típica: σ

Se nombra con la letra griega sigma.

Es la raíz cuadrada de la varianza: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2}$

A mayor desviación típica, mayor dispersión de las muestras.

COEFICIENTE DE VARIACIÓN: CV

Es el cociente entre la desviación típica y la media aritmética: $CV = \frac{\sigma}{\bar{x}}$

Es útil para poder comparar la dispersión entre datos con medias diferentes, pues relativizamos la dispersión al valor de la media. (No es lo mismo una misma desviación para una media pequeña que para una grande).

Cuanto más cercano a cero está, menor dispersión y las muestras estarán más concentradas alrededor de la media.

También se puede dar en porcentaje.

RESUMEN DE LAS FÓRMULAS

Media aritmética: $\bar{x} = \frac{\sum x_i \cdot f_i}{N}$

Moda: M_o = valor de la muestra que más se repite

Mediana: M_e = valor de la muestra central

Rango = valor mayor – valor menor

Varianza: $\sigma^2 = \frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2$

Desviación típica: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2}$

Coefficiente de Variación: $CV = \frac{\sigma}{\bar{x}}$