TUACADEMIAFACIL

MEDIDAS DE DISPERSIÓN

Las medidas de tendencia central no siempre son fiables, a veces son un poco engañosas, por eso recurrimos a las medidas de dispersión. Para que la investigación acerca de una distribución quede completa es imprescindible saber si los datos numéricos están agrupados o no alrededor de los valores centrales. A los parámetros que miden estas desviaciones respecto a la media se les llama medidas o parámetros de dispersión.

- **Absolutas** (en unidades de x)
  - Recorrido (Re)
  - Recorrido Intercuartílico (RI)
  - Desviación media
    - Respecto a la media
    - Respecto a la mediana
  - Varianza
  - Desviación típica
- **Relativas** (adimensionales) → Permiten comparar
  - Coeficiente de apertura
  - Recorrido semi-intercuartílico
  - Coeficiente de variación de Pearson
  - Índice de dispersión respecto a la Mediana

• MEDIDAS DE DISPERSIÓN ABSOLUTAS

- RANGO O RECORRIDO (unidade de x) u

Es la diferencia entre el valor más grande y el más pequeño de x.

$$R = \text{máx}\{x_1, \ldots, x_k\} - \text{mín}\{x_1, \ldots, x_k\}$$

$$\boxed{RE = X_{max} - X_{min}}$$

- VARIANZA (unidade de x^2) u^2

Es la media aritmética de la suma de los cuadrados de las desviaciones respecto a la media.

	Varianza	Desviación	Media
Población	$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$	$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$
Muestra	$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

$$S_x^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} \geq 0$$

3, 3, 3, 3, 3, 3 $\quad \bar{x} = 3 \quad S_x^2 = 0$

Propiedades:

- La varianza no puede ser negativa V(x) ≥ 0.
- Si sumamos una constante a todas las observaciones (cambio de origen), la varianza no varía.
- Si multiplicamos todos los valores por una constante (cambio de escala), la varianza queda multiplicada por el cuadrado de dicha constante.
- La varianza de una constante es 0. Var(ctte) = Var(3) = 0
- El principal inconveniente de la varianza es que al ser una medida de dispersión elevada al cuadrado, hace que perdamos parte de su significado económico. Por ello utilizaremos otra medida de dispersión: la desviación estándar.

- DESVIACIÓN TÍPICA O ESTÁNDAR

Es la raíz cuadrada positiva de la varianza. $S_x^2 = 25€^2 \rightarrow S_x = \sqrt{25} = 5€$

	Varianza	Desviación
Población	$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$	$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$
Muestra	$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

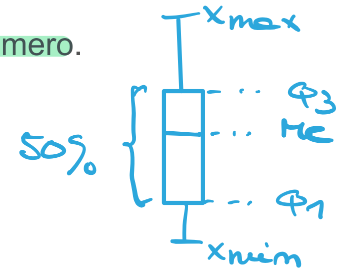
$S_x^2 = 0 \rightarrow S_x = 0$

Propiedades:

- Siempre es positiva.
- Es una medida de dispersión óptima.
- La desviación estándar de una constante es 0.
- Si sumamos a todos los valores de la variable una constante, la desviación estándar no varía.
- Si multiplicamos todos los valores de la variable por una constante, la desviación estándar queda multiplicada también por la constante.

- RANGO INTERCUARTÍLICO

Es la diferencia entre el tercer cuartil y el primero.



$$\boxed{IQR = RI = Q_3 - Q_1}$$

• MEDIDAS DE DISPERSIÓN RELATIVAS → Adimensional → Permiten comparar distribuciones

Éstas relacionan las medidas de tendencia central con las medidas de dispersión. Ponderan la dispersión respecto la magnitud global de la variable, y de esta manera permiten hacer comparaciones.

- COEFICIENTE DE APERTURA

Es el cociente entre el valor máximo y el mínimo de la variable.

$$A = \frac{Max\ X_i}{Min\ X_i} = \frac{X_{max}}{X_{min}} \longrightarrow \text{adimensional}$$

- COEFICIENTE DE VARIACIÓN DE PEARSON

La varianza y la desviación típica dependen de la unidad de medida que se emplea para medir la variable. Eso es un grave inconveniente cuando se quiere comparar la dispersión de dos poblaciones medidas con distintas escalas o se quieren comparar variables distintas. Para tener una medida invariante respecto de la unidad de medida empleada, se dispone del coeficiente de variación.

Se llama coeficiente de variación al cociente entre la desviación típica y el valor absoluto de la media. No se puede calcular cuando la media es 0, y no tiene mucho sentido calcularlo cuando la distribución de datos tiene valores negativos. Se representa por CV y su expresión es:

$$\boxed{CV = \frac{s}{|\bar{x}|}} \qquad \boxed{CV = \frac{S_x}{|\bar{x}|} \times 100}$$

Suele expresarse en forma de porcentaje, multiplicado por 100 el valor anterior. Expresa en porcentaje la magnitud de la desviación estándar en función de la media aritmética.

- Cuanto más grande sea el coeficiente de variación, más dispersados estarán los datos. Cuanto más pequeño sea, más concentrados estarán los datos alrededor de la media.

- El coeficiente de variación no siempre es positivo, ya que la media puede ser negativa.

- Cuando la media es un valor cercano a 0 no debe emplearse Pearson.

* CV bajo → datos con poca dispersión
  Distribución es más homogénea
  Media es más representativa

* CV elevado → datos con mucha dispersión
  Distribución menos homogénea
  Media es menos representativa