

Medidas de centralización

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

Mediana: Valor de x_i que ocupa la posición central (M o Me)

Moda: Valor de x_i cuya frecuencia f_i es la más alta (Mo o Mo)

Medidas de dispersión

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

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

Desviación media: $D_{\bar{x}} = \frac{\sum f_i \cdot |x_i - \bar{x}|}{\sum f_i}$

Coefficiente de variación: $CV = \frac{s}{\bar{x}}$

Medidas de posición

Cuartiles: Q_1 (primer cuartil) → Valor superior a 1/4 (25 %) de los valores de x_i e inferior a 3/4 (75 %) de dichos valores.

Q_3 (tercer cuartil) → Valor superior a 3/4 (75 %) de los valores de x_i e inferior a 1/4 (25 %) de dichos valores.

Rango intercuartílico → $Q_3 - Q_1$

Deciles: D_1 a D_9 → Valores que dividen la distribución en 10 grupos del mismo número de valores.

Percentiles: P_1 a P_{99} → Valores que dividen la distribución en 100 grupos del mismo número de valores.

Distribuciones bidimensionales

Covarianza: $s_{xy} = \frac{\sum f_i \cdot x_i \cdot y_i}{\sum f_i} - \bar{x} \cdot \bar{y}$

Coefficiente de correlación de Pearson: $r = \frac{s_{xy}}{s_x \cdot s_y}$

- r está comprendido entre -1 y 1.
- La correlación es más fuerte cuanto mayor sea el valor de $|r|$.

Rectas de regresión: De Y sobre X: $y - \bar{y} = \frac{s_{xy}}{s_x^2} (x - \bar{x})$

De X sobre Y: $x - \bar{x} = \frac{s_{xy}}{s_y^2} (y - \bar{y})$

- Ambas rectas se cortan en el punto $(\bar{x}, \bar{y})$.
- Si la correlación es fuerte, las dos rectas prácticamente coinciden y se ajustan a la nube de puntos.