

FÓRMULAS TEMA 2

MEDIA ARITMÉTICA: Suma de los valores de los datos dividida por el número de observaciones.

Poblacional: $\mu = \frac{\sum_{i=1}^N X_i}{N} = \frac{X_1 + X_2 + \dots + X_N}{N}$

Muestral: $\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$

MEDIANA: Es la observación que ocupa el lugar central de un conjunto de observaciones ordenadas.

$$\text{Posición_mediana} = \frac{1}{2} (n + 1)$$

MODA: Si existe, es el valor que aparece con más frecuencia.

RANGO: Diferencia entre la mayor y menor observación

$$\text{Rango} = \text{Max} (X_i) - \text{Min} (X_i)$$

CUARTILES

➤ **PRIMER CUARTIL (Q₁):** $\text{posición_Q}_1 = \frac{1}{4} (n + 1)$

Llamando m a la parte entera de $\frac{1}{4} (n + 1)$, se llama **primer cuartil muestral** (o **cuartil inferior muestral**) a:

$$Q_1(X) = \begin{cases} X_{(m)} & \text{si } \frac{1}{4} (n + 1) = m \\ \frac{3}{4} X_{(m)} + \frac{1}{4} X_{(m+1)} & \text{si } \frac{1}{4} (n + 1) = m + 0.25 \\ \frac{X_{(m)} + X_{(m+1)}}{2} & \text{si } \frac{1}{4} (n + 1) = m + 0.5 \\ \frac{1}{4} X_{(m)} + \frac{3}{4} X_{(m+1)} & \text{si } \frac{1}{4} (n + 1) = m + 0.75 \end{cases}$$

➤ **TERCER CUARTIL (Q₃):** $\text{posición_Q}_3 = \frac{3}{4} (n + 1)$

Llamando m a la parte entera de $\frac{3}{4} (n + 1)$, se llama **tercer cuartil muestral** (o **cuartil superior muestral**) a:

$$Q_3(X) = \begin{cases} X_{(m)} & \text{si } \frac{3}{4} (n + 1) = m \\ \frac{3}{4} X_{(m)} + \frac{1}{4} X_{(m+1)} & \text{si } \frac{3}{4} (n + 1) = m + 0.25 \\ \frac{X_{(m)} + X_{(m+1)}}{2} & \text{si } \frac{3}{4} (n + 1) = m + 0.5 \\ \frac{1}{4} X_{(m)} + \frac{3}{4} X_{(m+1)} & \text{si } \frac{3}{4} (n + 1) = m + 0.75 \end{cases}$$

RANGO INTERCUARTÍLICO (RIQ): Mide la dispersión que hay en el 50% central de los datos.

$$RIQ(X) = Q_3(X) - Q_1(X)$$

VARIANZA: Suma de los cuadrados de las diferencia entre cada observación y la media dividida por el tamaño de la población o el de la muestra, menos 1.

Poblacional: $\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$

Muestral: $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$

DESVIACIÓN TÍPICA: Raíz cuadrada de la varianza.

Poblacional: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$

Muestral: $S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$

COEFICIENTE DE VARIACIÓN (CV): Medida de la dispersión relativa a la media.

Poblacional: $CV = \frac{\sigma}{\mu}$

Muestral: $CV = \frac{S}{\bar{X}}$

MEDIA PONDERADA

Poblacional: $\mu = w_1X_1 + w_2X_2 + \dots + w_NX_N$

Muestral: $\bar{X} = w_1X_1 + w_2X_2 + \dots + w_nX_n$

DATOS AGRUPADOS:

MEDIA: $\bar{x} = \frac{\sum_{i=1}^K n_i c_i}{n} = \sum_{i=1}^K f_i c_i$

VARIANZA: $S^2 = \frac{\sum_{i=1}^K n_i (c_i - \bar{x})^2}{n-1}$

COVARIANZA: Medida del sentido de la relación lineal entre dos variables.

Poblacional: $Cov(x,y) = \sigma_{xy} = \frac{\sum_{i=1}^N (X_i - \mu_x) (Y_i - \mu_y)}{N}$

Muestral: $Cov(x,y) = S_{xy} = \frac{\sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y})}{n-1}$

COEFICIENTE DE CORRELACIÓN: Coeficiente de la covarianza con el producto de las desviaciones típicas.

Poblacional: $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

Muestral: $r_{xy} = \frac{S_{xy}}{S_x S_y}$