

**Definición del coeficiente de correlación**

Dada una distribución estadística bidimensional, se define su coeficiente de correlación como el cociente entre la covarianza y el producto de las desviaciones típicas de cada variable.

**Definición simbólica**

Si las variables estadísticas se llaman X e Y, el coeficiente de correlación es:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

**Enunciado**

Calcula con dos cifras significativas el coeficiente de correlación de la siguiente distribución:

X	1	2	3	3	4	5	5	6	7	8
Y	9	9	7	6	2	4	2	3	2	2

**Resolución**

Tenemos diez datos de la forma  $(x_i, y_i)$ . Los colocamos en una tabla junto con algunas filas y columnas auxiliares.

											↓ Sumas ↓
$x_i$	1	2	3	3	4	5	5	6	7	8	44
$y_i$	9	9	7	6	2	4	2	3	2	2	46
$x_i^2$	1	4	9	9	16	25	25	36	49	64	238
$y_i^2$	81	81	49	36	4	16	4	9	4	4	288
$x_i y_i$	9	18	21	18	8	20	10	18	14	16	152

De la tabla obtenemos:  $\Sigma x_i = 44$ ,  $\Sigma y_i = 46$ ,  $\Sigma x_i^2 = 238$ ,  $\Sigma y_i^2 = 288$ ,  $\Sigma x_i y_i = 152$

Operaciones:

\* Medias:  $\bar{x} = \frac{\Sigma x_i}{10} = \frac{44}{10} = 4,4$ ;  $\bar{y} = \frac{\Sigma y_i}{10} = \frac{46}{10} = 4,6$

\* Covarianza:  $\sigma_{xy} = \frac{\Sigma x_i y_i}{10} - \bar{x} \bar{y} = \frac{152}{10} - 4,4 \cdot 4,6 = 15,2 - 20,24 = -5,04$

\* Desv. típica de las «x»:  $\sigma_x = \sqrt{\frac{\Sigma x_i^2}{10} - \bar{x}^2} = \sqrt{\frac{238}{10} - 4,4^2} = \sqrt{23,8 - 19,36} = \sqrt{4,44}$

\* Desv. típica de las «y»:  $\sigma_y = \sqrt{\frac{\Sigma y_i^2}{10} - \bar{y}^2} = \sqrt{\frac{288}{10} - 4,6^2} = \sqrt{28,8 - 21,16} = \sqrt{7,64}$

\* Coeficiente de correlación:  $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{-5,04}{\sqrt{4,4} \sqrt{7,64}} = -0,87$

\* Calculadora:

$(-) 5 \cdot 0 4 \div ( \sqrt{4 \cdot 4 4} \times \sqrt{7 \cdot 6 4} ) = \Rightarrow -0,865351143$

Solución:  $-0,87$