

Ejercicio 01

Un profesor de matemáticas pregunta a 6 alumnos y alumnas el número de horas que estudian diariamente y anota las calificaciones que ha obtenido cada uno de ellos, obteniendo los siguientes resultados:

X (horas de estudio)	1	3	4	2	1	0
Y (calificación)	4	8	9	6	5	2

Se pide:

- Calcular las medias y desviaciones típicas marginales.
- Calcular el coeficiente de correlación lineal e interpretarlo

Calculamos las medias de las dos variables:

$$\bar{x} = \frac{\sum_{i=1}^6 x_i}{n} = \frac{1+3+4+2+1+0}{6} = \frac{11}{6} = 1,83$$

$$\bar{y} = \frac{\sum_{i=1}^6 y_i}{n} = \frac{4+8+9+6+5+2}{6} = \frac{34}{6} = 5,67$$

Construimos la tabla que nos va a permitir calcular los elementos pedidos:

x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	4	-0,83	-1,67	0,69	2,78
3	8	1,17	2,33	1,36	5,44
4	9	2,17	3,33	4,69	11,11
2	6	0,17	0,33	0,03	0,11
1	5	-0,83	-0,67	0,69	0,44
0	2	-1,83	-3,67	3,36	13,44
Suma				10,83	33,33

Calculamos las desviaciones típicas marginales:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^6 (x_i - \bar{x})^2}{6}} = \sqrt{\frac{10,83}{6}} = \sqrt{1,805} = 1,343$$

$$\sigma_y = \sqrt{\frac{\sum_1^6 (y_i - \bar{y})^2}{6}} = \sqrt{\frac{33,33}{6}} = \sqrt{5,556} = 2,357$$

Para calcular el coeficiente de correlación lineal o de Pearson, aplicamos la fórmula:

$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$, siendo σ_{xy} el momento producto. Calculamos pues este valor utilizando

los resultados de nuestra tabla:

$$\sigma_{xy} = \frac{\sum_1^6 (x_i - \bar{x})(y_i - \bar{y})}{6} = \frac{1,39 + 2,72 + 7,22 + 0,06 + 0,56 + 6,72}{6} = \frac{18,67}{6} = 3,11$$

Ahora sustituimos en la fórmula los valores obtenidos:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{3,11}{1,343 \cdot 2,357} = \frac{3,11}{3,165} = 0,983$$

Recuerda que el coeficiente de correlación lineal es un valor comprendido entre -1 y 1. Valores muy cercanos a 0 indicarán que no hay correlación lineal entre las variables que estamos midiendo. En nuestro caso, $r = 0,983$ que es un valor muy próximo a 1, lo que indica una gran correlación entre las horas de estudio invertidas y la calificación obtenida en la prueba de mates, es decir, cuantas más horas se han invertido en el estudio, tanto mayor es la calificación lograda.