

# Relación de ejercicios 2 EDIP

Carlos García, Bora Goker, Javier Gómez,  
Ana Graciani, J.Alberto Hoces

2020/2021

**Ejercicio 1.** En una encuesta de familias sobre el número de individuos que la componen ( $X$ ) y el número de personas activas en ellas ( $Y$ ) se han obtenido los siguientes resultados:

$X/Y$	1	2	3	4
1	7	0	0	0
2	10	2	0	0
3	11	5	1	0
4	10	6	6	0
5	8	6	4	2
6	1	2	3	1
7	1	0	0	1
8	0	0	1	1

a) Calcular la recta de regresión de  $Y$  sobre  $X$ .

$X/Y$	1	2	3	4	$n_{i.}$	$n_{i.}x_i$	$n_{i.}x_i^2$
1	7	0	0	0	7	7	7
2	10	2	0	0	12	24	48
3	11	5	1	0	17	51	153
4	10	6	6	0	22	88	352
5	8	6	4	2	20	100	500
6	1	2	3	1	7	43	252
7	1	0	0	1	2	14	98
8	0	0	1	1	2	16	128
$n_{.j}$	48	21	15	5	89		
$n_{.j}y_j$	48	42	45	20			
$n_{.j}y_j^2$	48	84	135	80			

La recta de regresión lineal de  $Y$  sobre  $X$  viene dada por la expresión:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x}) \Rightarrow y = \frac{\sigma_{xy}}{\sigma_x^2}x - \frac{\sigma_{xy}}{\sigma_x^2}\bar{x} + \bar{y}$$

Por lo tanto, comencemos calculando las medias aritméticas y la varianza de  $x$  y la covarianza:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^8 x_i n_{i.} = 3,8427 \text{ individuos} \quad \bar{y} = \frac{1}{n} \sum_{j=1}^4 y_j n_{.j} = 1,7416 \text{ personas activas}$$

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^8 n_{i.} x_i^2 - \bar{x}^2 = 2,5146 \text{ individuos}^2$$

$$\sigma_{xy} = \frac{1}{n} \sum_{i=1}^8 \sum_{j=1}^4 n_{ij} x_i y_j - \bar{x} \bar{y} = 0,7907$$

Por lo tanto, la recta de regresión de  $Y$  sobre  $X$  quedaría:

$$y = \frac{\sigma_{xy}}{\sigma_x^2} x - \frac{\sigma_{xy}}{\sigma_x^2} \bar{x} + \bar{y} = 0,3144x + 0,5333$$

- b) ¿Es adecuado suponer una relación lineal para explicar el comportamiento de  $Y$  a partir de  $X$ ?

Para ver cómo de adecuado es suponer dicha relación calculamos el coeficiente de correlación lineal:

$$r^2 = \sqrt{\frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2}} \quad r = \sqrt{r^2}$$

Ahora calculamos la varianza de  $Y$ :

$$\sigma_y^2 = \frac{1}{n} \sum_{j=1}^4 n_{.j} y_j^2 - \bar{y}^2 = 0,8657 \text{ personas activas}^2$$

Por tanto:

$$r^2 = \frac{0,7907^2}{2,5146 \cdot 0,8657} = 0,2872 \quad r = 0,536$$

Observando estos resultados podemos afirmar que no es adecuado suponer esta relación lineal puesto que el coeficiente de correlación lineal está demasiado alejado de 1.

