

X: ANTIGÜEDAD LABORAL [AÑOS] (j)

Y: \$ DEL BONO RECIBIDO [M.\$] (i)

Y/X	[0;5)	[5;10)	[10;15)	[15;20)	TOTAL	m_i
[3;3,8)	22	27	21	0	70	3,4
[3,8;4,6)	15	21	13	0	49	4,2
[4,6;5,4)	55	12	15	23	105	5
[5,4;6,2)	0	0	25	45	70	5,8
[6,2;7]	0	14	14	27	55	6,6
TOTAL	92	74	88	95	349 = n	
m_j	2,5	7,5	12,5	17,5		

a) [0;5)

j=1

m_i	m_j	f_i	N_i	F_i
3,4	22	22/92	22	22/92
4,2	15	15/92	37	37/92
5	55	55/92	92	1
5,8	0	0	92	1
6,6	0	0	92	1
TOTAL	92			

$$\bar{y}_1 = \sum_{i=1}^5 f_i \cdot m_i = 4,487 \quad \text{Media}$$

$$s_{0,1}^2 = \left(\sum_{i=1}^5 f_i \cdot m_i^2 \right) - \bar{y}_1^2 = 0,453 \quad \text{Varianza}$$

$$\Rightarrow \text{DESV. STD} \quad S = \sqrt{s^2} = \sqrt{0,453} = 0,673$$

• [5; 10]
j=2

m_i	n_i	f_i
3,4	27	27/74
4,2	21	21/74
5	12	12/74
5,8	0	0
6,6	14	14/74
TOTAL	74	

$$\mu \bar{Y}_2 = \sum_{i=1}^5 f_i \cdot m_i = \underline{4,492} \quad (2)$$

$$S_{0,2}^2 = \left(\sum_{i=1}^5 f_i \cdot m_i^2 \right) - \bar{Y}_2^2 = 1,342 \Rightarrow S_{0,2} = \sqrt{S_{0,2}^2} = \underline{1,158}$$

• [10; 15]
j=3

m_i	n_i	f_i
3,4	21	21/88
4,2	13	13/88
5	15	15/88
5,8	25	25/88
6,6	14	14/88
TOTAL	88	

$$\mu \bar{Y}_3 = \sum_{i=1}^5 f_i \cdot m_i = \underline{4,982}$$

$$S_{0,3}^2 = \left(\sum_{i=1}^5 f_i \cdot m_i^2 \right) - \bar{Y}_3^2 = 1,294 \Rightarrow S_{0,3} = \sqrt{S_{0,3}^2} = \underline{1,138}$$

• [15; 20]
j=4

m_i	n_i	f_i
3,4	0	0
4,2	0	0
5	23	23/95
5,8	45	45/95
6,6	27	27/95
TOTAL	95	

$$\mu \bar{Y}_4 = \sum_{i=1}^5 f_i \cdot m_i = \underline{5,834}$$

$$S_{0,4}^2 = \left(\sum_{i=1}^5 f_i \cdot m_i^2 \right) - \bar{Y}_4^2 = 0,336 \Rightarrow S_{0,4} = \sqrt{S_{0,4}^2} = \underline{0,579}$$

• El rango que presenta un promedio mayor en el bono es [15; 20] con $\bar{Y}_4 = 5,834$ [m\$]
 y el que presenta la menor desviación estándar es el mismo con $S_{0,4} = 0,579$

③

b) Para generar un M.R.L., primero se definen las Variables, donde:

- Variable independiente o predictora sea $X = \text{"ANTIGÜEDAD LA BOMBA"}$
- Variable dependiente o respuesta sea $Y = \text{"MONTO DEL BONO"}$

donde $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \Rightarrow \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$

$$\hat{\beta}_1 = \frac{\text{COV}(X, Y)}{S_X^2}$$

- Promedio MARGINAL DE X y Y se obtiene utilizando los TOTALES DE CADA RANGO

$$\bar{X} = \sum_{j=1}^4 f_{.j} \cdot m_j = 10,165 \quad \bar{Y} = \sum_{i=1}^5 f_{i.} \cdot m_i = 4,979$$

- AL MISMO TIEMPO y con los mismos datos; se obtiene la varianza marginal de X

$$S_X^2 = \left(\sum_{j=1}^4 f_{.j} \cdot m_j^2 \right) - \bar{X}^2 = 33,014$$

$$\Rightarrow \text{si } ① \text{COV}(X, Y) = \left[\sum_{i=1}^5 \sum_{j=1}^4 f_{ij} \cdot m_i \cdot m_j \right] - \bar{X} \cdot \bar{Y}$$

$$\Rightarrow \text{si } ② \left| f_{ij} = \frac{m_{ij}}{m} \right| \quad y \quad ③ \left| m = 349 \right|$$

$\Rightarrow$ ~~m_{ij}~~ Reemplazando ② y ③ en ①

$$\text{COV}(X, Y) = \left[\frac{1}{349} \sum_{i=1}^5 \sum_{j=1}^4 m_{ij} \cdot m_i \cdot m_j \right] - \bar{X} \cdot \bar{Y} = 2,98$$

$$\Rightarrow \hat{\beta}_1 = \frac{\text{COV}(X, Y)}{S_x^2} = \frac{2,98}{33,014} = \underline{0,09}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{X} = 4,979 - (0,09)(10,165) = \underline{4,06}$$

$$\Rightarrow \hat{Y}_1 = 4,06 + 0,09 X_1$$

$\therefore$ si $X_1 = 7 \Rightarrow Y_1 = 4,69$ $\therefore$ si un trabajador lleva 7 años debe recibir 4,69 [M\$]

Ahora, si $X_2 = 25$; Matemáticamente, NO se puede determinar porque la predicción debe estar dentro de la "nube" de puntos con la que se hizo la regresión; por lo tanto no sería correcto estimar con $X = 25$.

c) Para determinar la calidad del modelo se debe calcular el coeficiente de determinación del modelo de regresión lineal (R^2)

donde
$$R^2 = r^2 = \frac{[\text{COV}(X, Y)]^2}{S_x^2 S_y^2}$$

$$\Rightarrow S_y^2 = \left(\sum_{i=1}^5 f_i \cdot m_i^2 \right) - \bar{Y}^2 = \underline{1,135}$$

$$\Rightarrow R^2 = \frac{(2,98)^2}{(33,014)(1,135)} = \underline{0,24}$$

$\therefore$ Considerando que $R^2 \in [0, 1]$ donde si:

$R^2 = 0 \rightarrow$ No es confiable la estimación y no explica el modelo

$R^2 = 1 \rightarrow$ Es totalmente confiable y explica perfectamente el modelo

$\Rightarrow$ si $R^2 = 0,24$, entonces quiere decir que la estimación para 7 años es poco confiable ya que su coeficiente de determinación es bajo y cercano a "0".