MULTIPLE LINEAL REGRESSION

Ejercicio 1: Se realizó un estudio a 12 estudiantes para ver cómo influyen las calificaciones del examen y el número de clases que los estudiantes pierden en la calificación de la materia de estadística. Los datos son:

Estudiante	Calificación de estadística (y)	Calificación del examen (x_1)	Clases perdidas (x_2)	x_1^2	x_2^2	$x_1 \cdot x_2$	$x_1 \cdot y$	$x_2 \cdot y$
1	85	65	1	4.225	1	65	5.525	85
2	74	50	7	2.500	49	350	3.700	518
3	76	55	5	3.025	25	275	4.180	380
4	90	65	2	4.225	4	130	5.850	180
5	85	55	6	3.025	36	330	4.675	510
6	87	70	3	4.900	9	210	6.090	261
7	94	65	2	3.025	4	130	6.110	188
8	98	70	5	4.900	25	350	6.860	490
9	81	55	4	3.025	16	220	4.455	324
10	91	70	3	4.900	9	210	6.370	273
11	76	50	1	2.500	1	50	3.800	76
12	74	55	4	3.025	16	220	4.070	296
	1.011	**725**	**43**	**44.475**	**195**	**2.540**	**61.685**	**3.581**

$$\textit{Método matricial}: \begin{pmatrix} n & \Sigma x_1 & \Sigma x_2 \\ \Sigma x_1 & \Sigma x_1 \cdot \Sigma x_1 & \Sigma x_1 \cdot \Sigma x_2 \\ \Sigma x_2 & \Sigma x_2 \cdot \Sigma x_1 & \Sigma x_2 \cdot \Sigma x_2 \end{pmatrix} = \begin{pmatrix} n & \Sigma x_1 & \Sigma x_2 \\ \Sigma x_1 & \Sigma x_1^2 & \Sigma x_1 \cdot x_2 \\ \Sigma x_2 & \Sigma x_2 \cdot x_1 & \Sigma x_2^2 \end{pmatrix} = \begin{pmatrix} 12 & 725 & 43 \\ 725 & 44.475 & 2.540 \\ 43 & 2.540 & 195 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

COLUMN 1:

$$Row_1/_{12}: \begin{pmatrix} 1 & ^{725}/_{12} & ^{43}/_{12} \\ 725 & 44.475 & 2.540 \\ 43 & 2.540 & 195 \end{pmatrix} = \begin{pmatrix} ^{1}/_{12} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to Row_1 \cdot (-725) + Row_2: \begin{pmatrix} 1 & ^{725}/_{12} & ^{43}/_{12} \\ 0 & ^{8.075}/_{12} & ^{-695}/_{12} \\ 43 & 2.540 & 195 \end{pmatrix} = \begin{pmatrix} ^{1}/_{12} & 0 & 0 \\ ^{-725}/_{12} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to Row_1 \cdot (-43) + Row_3: \begin{pmatrix} 1 & ^{725}/_{12} & ^{43}/_{12} \\ 0 & ^{8.075}/_{12} & ^{-695}/_{12} \\ 0 & ^{-695}/_{12} & ^{491}/_{12} \end{pmatrix} = \begin{pmatrix} ^{1}/_{12} & 0 & 0 \\ ^{-725}/_{12} & 1 & 0 \\ ^{-43}/_{12} & 0 & 1 \end{pmatrix}$$

COLUMN 2:

$$Row_2/\left(^{8.075}/_{12}\right): \begin{pmatrix} 1 & ^{725}/_{12} & ^{43}/_{12} \\ 0 & 1 & ^{-139}/_{1.615} \\ 0 & ^{-695}/_{12} & ^{491}/_{12} \end{pmatrix} = \begin{pmatrix} ^{1}/_{12} & 0 & 0 \\ ^{-29}/_{323} & ^{12}/_{8.075} & 0 \\ ^{-43}/_{12} & 0 & 1 \end{pmatrix} \to Row_2 \cdot \left(^{-725}/_{12}\right) + Row_1: \begin{pmatrix} 1 & 0 & ^{2.837}/_{323} \\ 0 & 1 & ^{-139}/_{1.615} \\ 0 & ^{-695}/_{12} & ^{491}/_{12} \end{pmatrix} = \begin{pmatrix} ^{1.779}/_{323} & ^{-29}/_{323} & 0 \\ ^{-29}/_{323} & ^{12}/_{8.075} & 0 \\ ^{-43}/_{12} & 0 & 1 \end{pmatrix} \to Row_2 \cdot \left(^{695}/_{12}\right) + Row_3: \begin{pmatrix} 1 & 0 & ^{2.837}/_{323} \\ 0 & 1 & ^{-139}/_{1.615} \\ 0 & 0 & ^{11.606}/_{323} \end{pmatrix} = \begin{pmatrix} ^{1.779}/_{323} & ^{-29}/_{323} & 0 \\ ^{-29}/_{323} & ^{12}/_{8.075} & 0 \\ ^{-2.837}/_{323} & ^{139}/_{1.615} & 1 \end{pmatrix}$$

COLUMN 3:

$$Row_3/\left(^{11.606}/_{323}\right): \begin{pmatrix} 1 & 0 & ^{2.837}/_{323} \\ 0 & 1 & ^{-139}/_{1.615} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} ^{1.779}/_{323} & ^{-29}/_{323} & 0 \\ ^{-29}/_{323} & ^{12}/_{8.075} & 0 \\ ^{-2.837}/_{11.606} & ^{139}/_{58.030} & ^{323}/_{11.606} \end{pmatrix} \to Row_3 \cdot \left(^{-2.837}/_{323}\right) + Row_1: \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & ^{-139}/_{1.615} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 7{,}65474 & ^{-6.431}/_{58.030} & ^{-2.837}/_{11.606} \\ ^{-29}/_{323} & ^{12}/_{8.075} & 0 \\ ^{-2.837}/_{11.606} & ^{139}/_{58.030} & ^{323}/_{11.606} \end{pmatrix} \to Row_3 \cdot \left(^{139}/_{1.615}\right) + Row_2: \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 7{,}65474 & ^{-6.431}/_{58.030} & ^{-2.837}/_{11.606} \\ ^{-6.431}/_{58.030} & ^{491}/_{290.150} & ^{139}/_{58.030} \\ ^{-2.837}/_{11.606} & ^{139}/_{58.030} & ^{323}/_{11.606} \end{pmatrix}$$

$$\begin{pmatrix} 7{,}65474 & -6.431/58.030 & -2.837/11.606 \\ -6.431/58.030 & 491/290.150 & 139/58.030 \\ -2.837/11.606 & 139/58.030 & 323/11.606 \end{pmatrix} \cdot \begin{pmatrix} \Sigma y = 1.011 \\ \Sigma x_1 \cdot y = 61.685 \\ \Sigma x_2 \cdot y = 3.581 \end{pmatrix} \to \begin{pmatrix} 7.738{,}949767 & -6.836{,}054368 & -875{,}3486989 \\ -112{,}0410305 & 104{,}3850939 & 8{,}577615027 \\ -247{,}13113976 & 147{,}7548682 & 99{,}66077891 \end{pmatrix} \to \begin{matrix} \boldsymbol{\beta_0 = 27,5467001} \\ \boldsymbol{\beta_1 = 0,921678427} \\ \boldsymbol{\beta_2 = 0,2842495075} \end{matrix}$$

- *Ecuación múltiple*: $y = \beta_0 + \beta_1 \cdot x_1 + \cdots + \beta_n \cdot x_n \to \boldsymbol{y = 27,5467001 + 0.921678427 \cdot x_1 + 0,2842495075 \cdot x_2}$

- *Estimación*: $\begin{cases} x_1 = 60 \\ x_2 = 4 \end{cases} \to y = 27{,}5467001 + 0.921678427 \cdot 60 + 0{,}2842495075 \cdot 4 = \boldsymbol{83,9844}$

Estudiante	Calificación de estadística (y)	Calificación del examen (x_1)	Clases perdidas (x_2)	$\hat{y} = 27,54 + 0,92 \cdot x_1 + 0,28 \cdot x_2$	$e = y - \hat{y}$	e^2	$(\hat{y} - \bar{y})^2$
1	85	65	1	87,74004736	+2,74004736	7,507859535	12,18043058
2	74	50	7	75,620368	−1,620368	2,625592455	74,47054846
3	76	55	5	79,66026112	−3,66026112	13,39751147	21,065570299
4	90	65	2	88,02429687	+1,97570313	3,903402858	14,24531686
5	85	55	6	79,94451063	+5,05548937	25,5579277	18,53723872
6	87	70	3	92.91693851	−5,91693851	35,01016133	75,11582314
7	94	65	2	88,02429687	+5,97570313	35,7090279	14,24531686
8	98	70	5	93,48543753	+4,51456247	20,3812743	85,29330637
9	81	55	4	79,37601162	+1,62398838	2,637338258	23,75576273
10	91	70	3	92,91693851	−1,91693851	3,674653251	75,11582314
11	76	50	1	73,91487096	+2,08512904	4,347763113	106,8148923
12	74	55	4	79,37601162	−5,37601162	28,90150094	23,75576273
	1.011	725	43	−	−	$SCE = \sum e^2 = 183{,}6540$	$SCR = \sum = 544{,}5959$

$\to n = 12 = nº\ obs.muestra$ ### $k = 2 = nº\ variables\ independientes$ ### $\bar{y} = \frac{\Sigma y}{n} = \frac{1.011}{12} = 84{,}25 = Media$

- $SCE = Suma\ de\ cuadrados\ del\ error$
- $SCR = Suma\ de\ cuadrados\ de\ la\ regresión$

- *Error Estándar de la estimación Múltiple* $\left(S_{y_n} \dots k\right) = \sqrt{\frac{SCE}{n-(k+1)}} = \sqrt{\frac{183{,}6540}{12-(2+1)}} = \boldsymbol{4,517300794}$
- *Suma Total de Cuadrados* $(STC) = SCE + SCR = 183{,}6540 + 544{,}5959 = \boldsymbol{728,2499831}$
- *Coeficiente de determinación múltiple* $(R^2) = \frac{SCR}{STC} = \frac{544{,}5959}{728{,}2499} = 0{,}747814 \approx \boldsymbol{74,78\%}$
- *Coeficiente ajustado de determinación múltiple* $\left(R^2_{ajustado}\right) = 1 - \frac{(1-R^2)\cdot(n-1)}{n-k-1} = 1 - \frac{(1-0{,}747814)\cdot(12-1)}{12-2-1} = 0{,}6917733246 \approx \boldsymbol{69,18\%}$
- *Coeficiente de correlación múltiple* $(R) = \sqrt{R^2} = \sqrt{0{,}7478145383} = \boldsymbol{0,8647627064} \to$ *Cerca de* $+1 =$ ***correlación positiva fuerte*** $=$ *estrecha relación entre variables*