

2. COEFICIENTES DE CORRELACIÓN

2.1. Coeficiente de correlación por el producto de los valores tipificados

$$r = \frac{\sum Z_x Z_y}{n}$$

Est.	MAT X	EST Y	d _x	d _y	d ² _x	d ² _y	Z _x	Z _y	Z _x Z _y
C	20	18	6,76	3,88	45,76	15,07	1,64	1,04	1,71
B	18	19	4,76	4,88	22,66	23,81	1,15	1,31	1,51
O	18	19	4,76	4,88	22,66	23,81	1,15	1,31	1,51
M	17	18	3,76	3,88	14,14	15,05	0,91	1,04	0,95
I	16	15	2,76	0,88	7,62	0,77	0,67	0,24	0,16
N	15	17	1,76	2,88	3,10	8,29	0,43	0,77	0,33
A	15	10	1,76	-4,12	3,10	16,97	0,43	-1,11	-0,47
K	14	16	0,76	1,88	0,58	3,53	0,18	0,50	0,09
L	14	16	0,76	1,88	0,58	3,53	0,18	0,50	0,09
D	13	16	-0,24	1,88	0,06	3,53	-0,06	0,50	-0,03
J	13	14	-0,24	-0,12	0,06	0,01	-0,06	-0,03	0,00
Q	12	7	-1,24	-7,12	1,54	50,69	-0,30	-1,91	0,57
P	11	13	-2,24	-1,12	5,02	1,25	-0,54	-0,30	0,16
E	10	8	-3,24	-6,12	10,50	37,45	-0,79	-1,64	1,29
H	8	13	-5,24	-1,12	27,46	1,25	-1,27	-0,30	0,38
J	7	12	-6,24	-2,12	38,94	4,49	-1,51	-0,57	0,86
G	4	9	-9,24	-5,12	85,38	26,21	-2,24	-1,37	3,08
	225	240			289,12	235,78			12,21
	ΣX	ΣY			Σd ² _x	Σd ² _y			ΣZ _x Z _y

$$\bar{X} = \frac{\sum X_i}{n} = \frac{225}{17} = 13.24$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{240}{17} = 14.12$$

$$\sigma_x = \sqrt{\frac{\sum d_x^2}{n}} = \sqrt{\frac{289.12}{17}} = 4.12$$

$$\sigma_y = \sqrt{\frac{\sum d_y^2}{n}} = \sqrt{\frac{235.78}{17}} = 3.72$$

$$Z_x = \frac{X - \bar{x}}{\sigma_x}$$

$$Z_y = \frac{y - \bar{y}}{\sigma_y}$$

$$r = \frac{\sum Z_x Z_y}{n} = \frac{12.21}{17} = 0.72$$

De acuerdo al coeficiente de correlación por el producto de los valores tipificados, existe una correlación alta y positiva entre las calificaciones de matemática y estadística obtenida por el grupo de estudiantes.

2.2. Coeficiente de correlación por los cuadrados de las diferencias de los valores tipificados

$$r = 1 - \frac{\sum (Z_x - Z_y)^2}{2n}$$

Est.	MAT X	EST Y	d _x	d _y	d ² _x	d ² _y	Z _x	Z _y	Z _x -Z _y	(Z _x -Z _y) ²
C	20	18	6,76	3,88	45,70	15,05	1,65	1,04	0,61	0,37
B	18	19	4,76	4,88	22,66	23,81	1,16	1,31	-0,15	0,02
O	18	19	4,76	4,88	22,66	23,81	1,16	1,31	-0,15	0,02
M	17	18	3,76	3,88	14,14	15,05	0,92	1,04	-0,13	0,02
I	16	15	2,76	0,88	7,62	0,77	0,67	0,24	0,44	0,19
N	15	17	1,76	2,88	3,10	8,29	0,43	0,77	-0,34	0,12
A	15	10	1,76	-4,12	3,10	16,97	0,43	-1,11	1,54	2,36
K	14	16	0,76	1,88	0,58	3,53	0,19	0,51	-0,32	0,10
L	14	16	0,76	1,88	0,58	3,53	0,19	0,51	-0,32	0,10
D	13	16	-0,24	1,88	0,06	3,53	-0,06	0,51	-0,56	0,32
J	13	14	-0,24	-0,12	0,06	0,01	-0,06	-0,03	-0,03	0,00
Q	12	7	-1,24	-7,12	1,54	50,69	-0,30	-1,91	1,61	2,60
P	11	13	-2,24	-1,12	5,02	1,25	-0,55	-0,30	-0,25	0,06
E	10	8	-3,24	-6,12	10,50	37,45	-0,79	-1,65	0,85	0,73
H	8	13	-5,24	-1,12	27,46	1,25	-1,28	-0,30	-0,98	0,95
J	7	12	-6,24	-2,12	38,94	4,49	-1,52	-0,57	-0,95	0,91
G	4	9	-9,24	-5,12	85,38	26,21	-2,25	-1,38	-0,88	0,77
	225	240			289,06	235,76			-0,01	9,58
	ΣX	ΣY			Σd ² _x	Σd ² _y			ΣZ _x -Z _y	Σ(Z _x -Z _y) ²

$$r = 1 - \frac{\sum (Z_x - Z_y)^2}{2n}$$

$$r = 1 - \frac{9,58}{34} = 0,72$$

De acuerdo al coeficiente de correlación por el cuadrado de la diferencia de los valores tipificados, existe una correlación alta y positiva entre las calificaciones de matemática y estadística obtenida por el grupo de estudiantes

2.3. Coeficiente de correlación por las desviaciones desde las medias

$$r = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}}$$

Est.	MAT X	EST Y	dx	dy	d ² _x	d ² _y	dx dy
C	20	18	6,76	3,88	45,70	15,05	26,23
B	18	19	4,76	4,88	22,66	23,81	23,23
O	18	19	4,76	4,88	22,66	23,81	23,23
M	17	18	3,76	3,88	14,14	15,05	14,59
I	16	15	2,76	0,88	7,62	0,77	2,43
N	15	17	1,76	2,88	3,10	8,29	5,07
A	15	10	1,76	-4,12	3,10	16,97	-7,25
K	14	16	0,76	1,88	0,58	3,53	1,43
L	14	16	0,76	1,88	0,58	3,53	1,43
D	13	16	-0,24	1,88	0,06	3,53	-0,45
J	13	14	-0,24	-0,12	0,06	0,01	0,03
Q	12	7	-1,24	-7,12	1,54	50,69	8,83
P	11	13	-2,24	-1,12	5,02	1,25	2,51
E	10	8	-3,24	-6,12	10,50	37,45	19,83
H	8	13	-5,24	-1,12	27,46	1,25	5,87
J	7	12	-6,24	-2,12	38,94	4,49	13,23
G	4	9	-9,24	-5,12	85,38	26,21	47,31
	225	240			289,06	235,76	187,53
	ΣX	ΣY			Σd ² X	Σd ² Y	Σdx dy

$$r = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}}$$

$$r = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}}$$

$$r = \frac{187.53}{\sqrt{289.06 * 235.76}}$$

$$r = 0.72$$

De acuerdo al coeficiente de correlación por las desviaciones desde las medias, existe una correlación alta y positiva entre las calificaciones de matemática y estadística obtenida por el grupo de estudiantes

2.4. Coeficiente de correlación por medio de las desviaciones típicas

$$r = \frac{\sum d_x d_y}{n \sigma_x \sigma_y}$$

Est.	MAT X	EST Y	dx	dy	d ² _x	d ² _y	dx dy
C	20	18	6,76	3,88	45,70	15,05	26,23
B	18	19	4,76	4,88	22,66	23,81	23,23
O	18	19	4,76	4,88	22,66	23,81	23,23
M	17	18	3,76	3,88	14,14	15,05	14,59
I	16	15	2,76	0,88	7,62	0,77	2,43
N	15	17	1,76	2,88	3,10	8,29	5,07
A	15	10	1,76	-4,12	3,10	16,97	-7,25
K	14	16	0,76	1,88	0,58	3,53	1,43
L	14	16	0,76	1,88	0,58	3,53	1,43
D	13	16	-0,24	1,88	0,06	3,53	-0,45
J	13	14	-0,24	-0,12	0,06	0,01	0,03
Q	12	7	-1,24	-7,12	1,54	50,69	8,83
P	11	13	-2,24	-1,12	5,02	1,25	2,51
E	10	8	-3,24	-6,12	10,50	37,45	19,83
H	8	13	-5,24	-1,12	27,46	1,25	5,87
J	7	12	-6,24	-2,12	38,94	4,49	13,23
G	4	9	-9,24	-5,12	85,38	26,21	47,31
	225	240			289,06	235,76	187,53
	$\sum X$	$\sum Y$			$\sum d^2 X$	$\sum d^2 Y$	$\sum dx dy$

$$\sigma_x = \sqrt{\frac{\sum d_x^2}{n}} = \sqrt{\frac{289.12}{17}} = 4.12 \quad \sigma_y = \sqrt{\frac{\sum d_y^2}{n}} = \sqrt{\frac{235.78}{17}} = 3.72$$

$$r = \frac{\sum d_x d_y}{n \sigma_x \sigma_y}$$

$$r = \frac{187.53}{17(4.12)(3.72)} = 0.72$$

De acuerdo al coeficiente de correlación por las desviaciones típicas, existe una correlación alta y positiva entre las calificaciones de matemática y estadística obtenida por el grupo de estudiantes

2.5. Series estadísticas mínimos cuadrados o datos originales

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

Est.	MAT X	EST Y	X ²	Y ²	XY
C	20	18	400	324	360
B	18	19	324	361	342
O	18	19	324	361	342
M	17	18	289	324	306
I	16	15	256	225	240
N	15	17	225	289	255
A	15	10	225	100	150
K	14	16	196	256	224
L	14	16	196	256	224
D	13	16	169	256	208
J	13	14	169	196	182
Q	12	7	144	49	84
P	11	13	121	169	143
E	10	8	100	64	80
H	8	13	64	169	104
J	7	12	49	144	84
G	4	9	16	81	36
	225	240	3267	3624	3364
	$\sum X$	$\sum Y$	$\sum X^2$	$\sum Y^2$	$\sum XY$

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{17(3364) - (225)(240)}{\sqrt{[17(3267) - 50625][17(3624) - 57600]}}$$

$$r = \frac{3188}{\sqrt{19695312}}$$

$$r = 0.72$$

De acuerdo al coeficiente de correlación con valores originales o de los mínimos cuadrados, existe una correlación alta y positiva entre las calificaciones de matemática y estadística obtenida por el grupo de estudiantes

2.6. Coeficiente de correlación para datos agrupados

Cuando los elementos que se van a relacionar son muchos debemos aplicar la correlación a través de los intervalos para lo cual necesitamos marcar una matriz de doble entrada.

EJERCICIO

$$Ax = X_M - X_m$$

$$Ax = 20 - 3$$

$$Ax = 17$$

$$Ay = Y_M - Y_m$$

$$Ay = 20 - 5$$

$$Ay = 15$$

$$C = \frac{\Delta + 1}{20} = \frac{18}{20} = 0.9$$

$$\frac{16}{20} = 0.8$$

$$C = \frac{\Delta + 1}{5} = \frac{18}{5} = 3.6$$

$$\frac{16}{5} = 3.2$$

Para el número de intervalos.

$$\frac{18}{3} = \textcircled{6} \text{ de } X$$

$$\frac{16}{3} = 5.33 \quad \textcircled{6} \text{ de } Y$$

X	Y	X	Y
13 – 14		15 – 13	
18 – 19		12 – 18	
15 – 13		11 – 16	
6 – 10		10 – 10	
3 – 5		13 – 9	
12 – 12		18 – 13	
11 – 11		18 – 19	
10 – 14		17 – 19	
9 – 8		16 – 18	
11 – 13		10 – 14	
15 – 16		11 – 10	
18 – 19		12 – 12	
13 – 15		16 – 13	
16 – 13		14 – 16	
20 – 18		15 – 10	
17 – 20		13 – 11	
18 – 19		10 – 15	
7 – 10		7 – 17	
4 – 8		9 – 10	
6 – 8		11 – 10	
3 – 5		8 – 13	
5 – 10		13 – 10	

Y \ X	VARIABLE X												n _{iy}	U _y	Σn _{iy} U _y	Σn _{iy} U _y ²	ΣU _y n U _x
	3	5	6	8	9	11	12	14	15	17	18	20					
20 22									1	18			1	3	3	9	6
17 19			1	-2			1	4	2	4	5	30	9	2	18	36	38
14 16					4	0	3	5	1	8			8	1	8	8	5
11 13		0	1	0	2	0	3	0	4	0	1	0	11	0	0	0	0
8 10	2	2	3	1	5	0	2	-2	1	-2			13	-1	-13	13	3
5 7	2	8											2	-2	-4	8	8
n _{ix}	4		5		11		9		9		6		44		12	74	60
U _x	-2		-1		0		1		2		3						
Σn _{ix} U _x	-8		-5		0		9		18		18		32				
n _{ix} U _x ²	16		5		0		9		36		54		120				
ΣU _x n _i U _y	12		1		0		3		14		30		60				

$$r = \frac{n \sum U_x n_{ix} U_y - \sum n_{ix} U_x - \sum n_{iy} U_y}{\sqrt{\left[n \sum n_{ix} U_x^2 - \left(\sum n_{ix} U_x \right)^2 \right] \left[n \sum n_{iy} U_y^2 - \left(\sum n_{iy} U_y \right)^2 \right]}}$$

$$r = \frac{44(60) - (32)(12)}{\sqrt{\left[44(120) - (32)^2 \right] \left[44(74) - (12)^2 \right]}}$$

$$r = 0.62$$

$$\bar{X} = \text{otx} + \left(\frac{\sum U_x n_{ix}}{n} \right) C_x$$

$$\bar{X} = 10 + \left(\frac{32}{44} \right) 3$$

$$\bar{X} = 12.18$$

$$\sigma_x = \frac{3}{44} \sqrt{44(120) - (32)^2}$$

$$\sigma_x = 4.45$$

$$C.Vx = \left(\frac{4.45}{12.18} \right) 100$$

$$CVx = 36.53\% \text{ heterogéneo}$$

g

$$\bar{Y} = \text{oty} + \left(\frac{\sum U_y n_{iy}}{n} \right) C_y$$

$$\bar{Y} = 10 + \left(\frac{44}{44} \right) 3$$

$$\bar{Y} = 13.$$

$$\sigma_y = \frac{3}{44} \sqrt{44(120) - (46)^2}$$

$$\sigma_y = 3.84$$

$$C.Vy = \left(\frac{3.84}{13} \right) 100$$

$$CVx = 29.22\% \text{ homogéneo}$$