Resuelve tú (Pág ❸⓿❾)

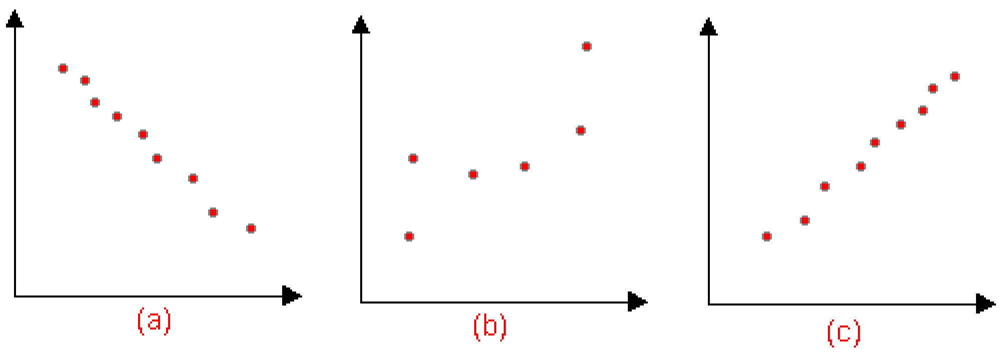
Dibuja nubes de puntos que describan aceptablemente bien las relaciones entre las variables que se indican:

(a) Libros leídos y faltas de ortografía cometidas (inversa y muy fuerte).

(b) Afición al ciclismo en TV y práctica del mismo (muy débil).

(c) Consumo de grasas y colesterol (directa y fuerte).

◊◊◊▣▣⊙▣▣◊◊◊



❄❄❄❁❁✪❁❁❄❄❄

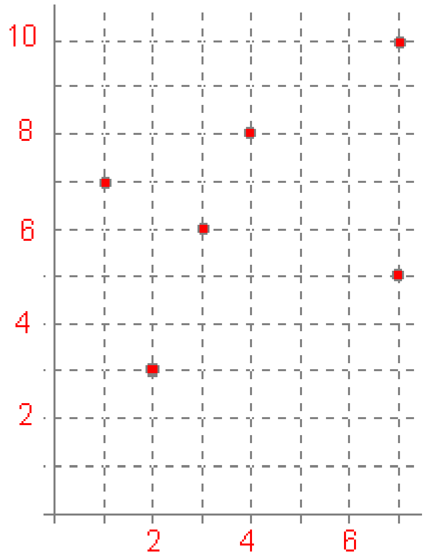
Resuelve tú (Pág ❸❶❶)

El número de problemas resueltos antes de una prueba y los aciertos obtenidos en ella, por 6 alumnos escogidos al azar, fueron:

Problemas resueltos, x:	2	3	1	7	4	7
Aciertos, y:	3	6	7	5	8	10

Representa gráficamente estos puntos. Halla el centro de gravedad de la nube de puntos obtenida y las desviaciones típicas marginales.

◊◊◊▣▣⊙▣▣◊◊◊



Para hallar las medias marginales (centro de gravedad) y las desviaciones típicas elaboramos la tabla base de los cálculos :

x_i	y_i	x_i^2	y_i^2
2	3	4	9
3	6	9	36
1	7	1	49
7	5	49	25
4	8	16	64
7	10	49	100
$\sum x_i = 24$	$\sum y_i = 39$	$\sum x_i^2 = 128$	$\sum y_i^2 = 283$

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{24}{6} = 4 \; ; \; \bar{y} = \frac{\sum_{i=1}^{6} y_i}{n} = \frac{39}{6} = 6'5$$

Centro de gravedad = ($\bar{x}$, $\bar{y}$) = (4, 6'5)

$$s_x = \sqrt{\frac{\sum_{i=1}^{6} x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{128}{6} - 4^2} = \sqrt{5'\widehat{3}} = 2'31 ; s_y = \sqrt{\frac{\sum_{i=1}^{6} y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{283}{6} - 6'5^2} = \sqrt{8'1\widehat{6}} = 2'86$$

Resuelve tú (Pág ❸❶❸)

Hallar el valor de la covarinza de la siguiente distribución, si la altura viniera expresada en metros. Comprobaras que es 100 veces menor.

Peso (kg), x:	45	55	43	47	51	60	63	58
Estatura (cm), y:	1,58	1,65	1,55	1,61	1,54	1,67	1,62	1,71

Usamos un programa de Estadística y obtenemos los resultados :

Media aritmética

X : 53.25 Y: 1.6175

Desviación típica

X : 6.456586 Y: 0.055396

Covariancia: 0.255625

Resuelve tú (Pág ❸❶❻)

A los alumnos del ejercicio de aplicación 4 se les preguntó también las horas que dedican a dormir diariamente. Contestaron :

7,5 8,5 8 9 7,5 8,5

calcula el coeficiente de correlación entre las horas dedicadas a ver TV y a dormir.

Elaboramos la tabla base para los cálculos :

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
12	7'5	144	56'25	90
10	8'5	100	72'25	85
15	8	225	64	120
20	9	400	81	180
28	7'5	784	56'25	210
8	8'5	64	72'25	68
$\sum x_i = 93$	$\sum y_i = 49$	$\sum x_i^2 = 1717$	$\sum y_i^2 = 402$	$\sum x_i \cdot y_i = 753$

Medias :

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{93}{6} = 15'5 \; ; \bar{y} = \frac{\sum_{i=1}^{6} y_i}{n} = \frac{49}{6} = 8'1\widehat{6}$$

Desviaciones típicas:

$$s_x = \sqrt{\frac{\sum_{i=1}^{6} x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{1717}{6} - 15'5^2} = \sqrt{45'91\widehat{6}} = 6'78 \; ; s_y = \sqrt{\frac{\sum_{i=1}^{6} y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{402}{6} - 8'1\widehat{6}^2} = \sqrt{0'31} = 0'55$$

Covarianza y coeficiente de correlación :

$$s_{xy} = \frac{\sum_{i=1}^{6} x_i y_i}{n} - \bar{x} \cdot \bar{y} = \frac{753}{6} - 15'5 \cdot 8'1\widehat{6} = -1'083 \qquad r = \frac{s_{xy}}{s_x \cdot s_y} = \frac{-1'083}{6'78 \cdot 0'55} = -0'29$$

Resuelve tú (Pág ❸❶❾)

Halla la recta que mejor se ajusta, según el criterio de los mínimos cuadrados, a los datos :

x	1	3	4	5	5	7
y	2	3	5	5	7	8

Necesitamos hallar las medias, la desviación típica y la covarianza, luego elaboramos primero la tabla auxiliar :

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
1	2	1	4	2
3	3	9	9	9
4	5	16	25	20
5	5	25	25	25
5	7	25	49	35
6	8	36	64	48
$\sum x_i = 24$	$\sum y_i = 30$	$\sum x_i^2 = 112$	$\sum y_i^2 = 176$	$\sum x_i \cdot y_i = 139$

Medias :

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{24}{6} = 4 \; ; \bar{y} = \frac{\sum_{i=1}^{6} y_i}{n} = \frac{30}{6} = 5$$

Desviación típica:

$$s_x = \sqrt{\frac{\sum_{i=1}^{6} x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{112}{6} - 4^2} = \sqrt{2'\widehat{6}} = 1'63$$

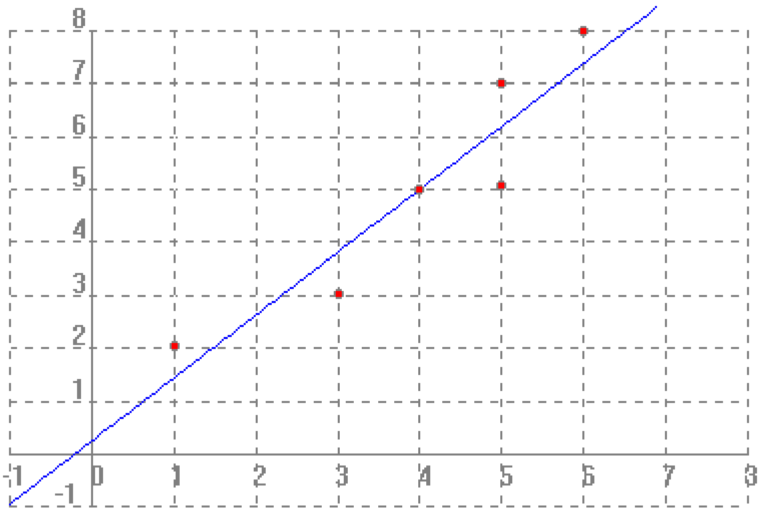
Covarianza :

$$s_{xy} = \frac{\sum_{i=1}^{6} x_i y_i}{n} - \bar{x} \cdot \bar{y} = \frac{139}{6} - 4 \cdot 5 = 3'1\widehat{6}$$

Ecuación de la recta :

$$y-\overline{y}=\frac{s_{xy}}{s_x^2}(x-\overline{x}) \rightarrow y-5=\frac{3'1\widehat{6}}{2'\widehat{6}}(x-4) \Leftrightarrow y-5=1'1875(x-4) \Leftrightarrow y=1'1875x+0'25$$

En el gráfico siguiente se muestra la nube de puntos y la representación de la ecuación de la recta hallada :



❄❄❄❁❁✪❁❁❄❄❄

Resuelve tú (Pág ❸❷❶)

Con los datos del Ejercicio de aplicación 4, halla la recta de regresión quc permite estimar las horas de estudio de un alumno sabiendo las hora; que dedica a ver la TV. ¿Cuánto estudiará, por término medio, ur alumno que ve la TV 14 h semanales?

◊◊◊▣▣⊙▣▣◊◊◊

Se trata de la recta de regresión de y sobre x :

$$y-\overline{y}=\frac{s_{xy}}{s_x^2}(x-\overline{x}) \rightarrow y-5'83=\frac{-16'53}{45'97}(x-15'5) \Leftrightarrow y-5'83=-0'36(x-15'5) \Leftrightarrow y=-0'36x+11'41$$

Sustituimos ahora en la ecuación obtenida la x = 14 y hallamos la y :

Si x = 14 h , y = -0’36 ·14 + 11’41 = 6’37 horas de estudio semanales.

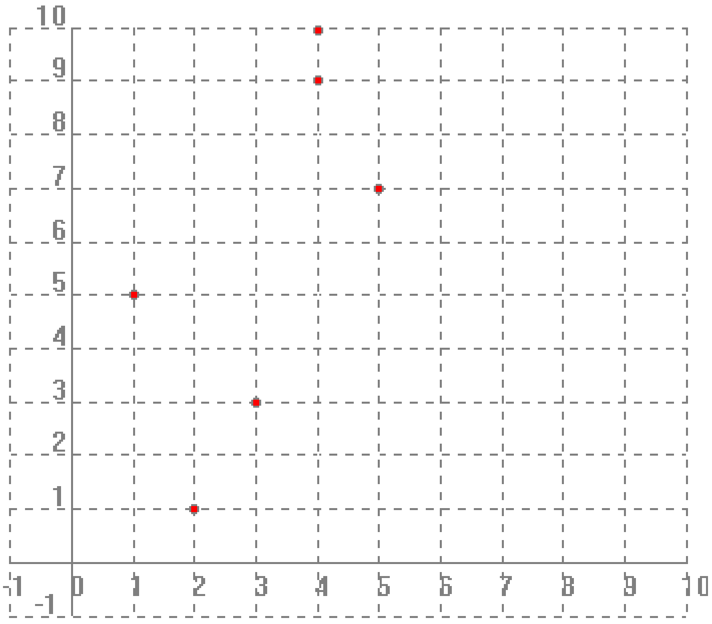
❄❄❄❁❁✪❁❁❄❄❄

PROBLEMAS PROPUESTOS

1 Representa la nube de puntos asociada a los datos:

x	3	2	5	4	1	9
y	3	1	7	10	5	4

¿Se observa correlación?



Poca correlación, hay gran dispersión.

2 Halla el centro de gravedad de la distribución anterior.

El centro de gravedad es el punto formado por las dos medias :

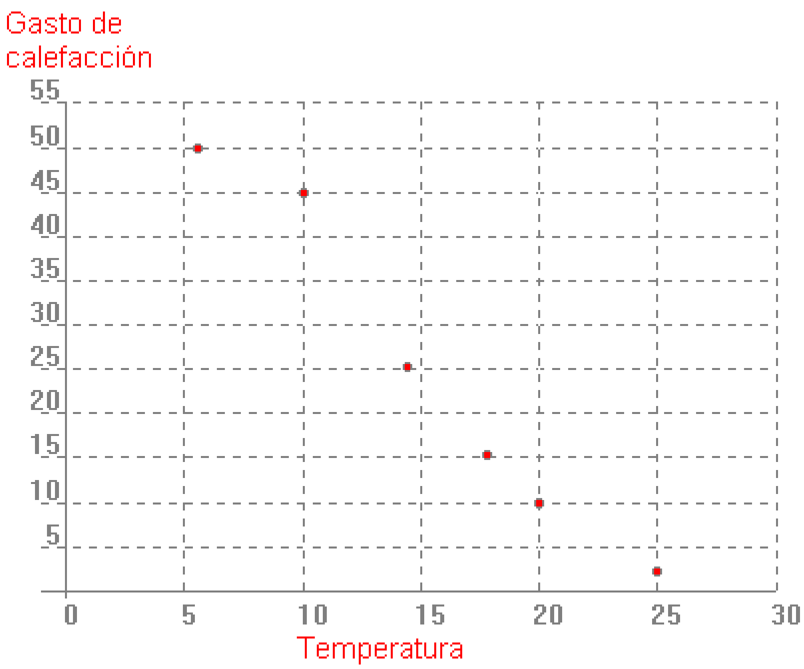
$$\bar{x} = \frac{3+2+5+4+1+9}{3} = \frac{24}{3} = 8 \qquad \bar{y} = \frac{3+1+7+10+5+4}{3} = \frac{30}{3} = 10$$

Centro de gravedad = ($\bar{x}$, $\bar{y}$) = (8, 10)

3 Representa un diagrama de dispersión en el que se observe una fuerte correlación negativa. Pon un ejemplo real que se ajuste aceptablemente bien a esa nube de puntos.

Representamos la temperatura media de varias ciudades y el gasto medio anual en calefacción por habitante (en miles de pesetas , que viene dado por la tabla :

Temperatura (ºC) x	6	10	14	18	20	25
Gasto en calefacción, y	50	45	25	15	10	2



❄❄❄❁❁✪❁❁❄❄❄

4 Dos conjuntos de datos bidimensionales tienen coeficientes de correlación r = 0,5 y r = -0,8. Se pide:

(a) ¿En cuál de los dos conjuntos se puede hacer una mejor estimación, mediante una recta, de una variable a partir de la otra?

(b) Representa dos conjuntos de puntos cuyas correlaciones se correspondan aproximadamente con los dados; sobre esos gráficos traza las rectas de regresión correspondientes.

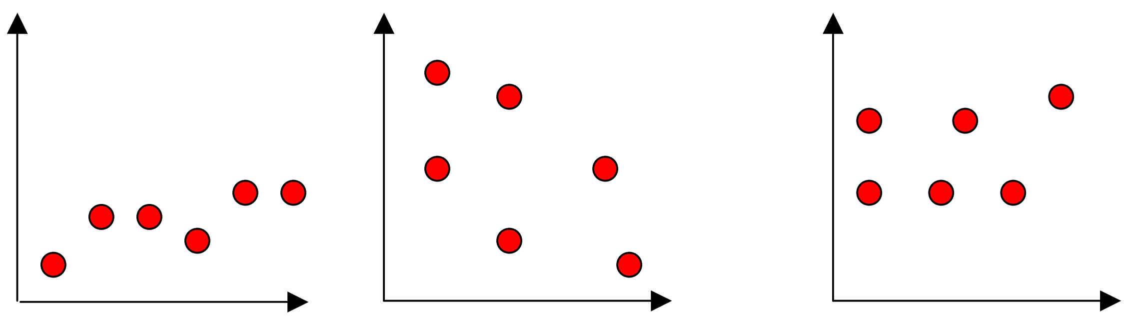
◊◊◊▣▣⊙▣▣◊◊◊

(a) En el segundo caso en el que el coeficiente de correlación es mayor en valor absoluto

(b) Las correspondientes a los ejercicios (1) y (3) anteriores por ejemplo.

❄❄❄❁❁✪❁❁❄❄❄

5 Asocia las rectas de regresión y=-x+6, y=x+5, y=0,5x+1 a las nubes de puntos siguientes:



(a) (b) (c)

◊◊◊▣▣⊙▣▣◊◊◊

y = - x + 6 que tiene pendiente negativa (con ángulo de 135º) ha de ser la **(b)** que es la única que tuiene esa pendiente.

De las dos de pendiente positiva una tiene el doble de pendiente que otra y mayor ordenada en el orígen, luego y = x + 5 ha de corresponder con la distribución **(c)** . y = 0'5x +1 es la **(a)** por exclusión.

❄❄❄❀❀✪❀❀❄❄❄

6 Asigna los coeficientes de correlación lineal r = 0,7, r = - 0,5 y r = 0,9 a las nubes de puntos del problema anterior.

◊◊◊▣▣⊙▣▣◊◊◊

No hay duda que el valor negativo r = - 0'5 ha de ser la **(b)** que tiene correlación inversa. De los dos valores positivos el mayor r = 0'9 se corresponde con el diagrama **(a)** y r = 0'7 con el **(c).**

❄❄❄❀❀✪❀❀❄❄❄

7 En la siguiente tabla se muestran las calificaciones de 8 alumnos en la asignatura de Física en dos pruebas:

Prueba 1	5	6	4	7	6	7	6	3
Prueba 2	7	5	4	9	8	8	9	3

(a) Representa la nube de puntos. Traza a ojo la recta de regresión. Estima su pendiente.

(b) Halla la recta de regresión. Compárala con tu estimación inicial.

◊◊◊▣▣⊙▣▣◊◊◊

(a) Ver figuar del final del apartad **(b)** . Recta estimada, en rojo.
(b) Para hallar los datos que necesitamos eleboramos al tabla auxiliar :

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
5	7	25	49	35
6	5	36	25	30
4	4	16	16	16
7	9	49	81	63
6	8	36	64	48
7	8	49	64	56
6	9	36	81	54
3	3	9	9	9
$\sum x_i = 44$	$\sum y_i = 53$	$\sum x_i^2 = 256$	$\sum y_i^2 = 389$	$\sum x_i \cdot y_i = 311$

❂ Medias :

$$\bar{x}=\frac{\sum_{i=1}^{6}x_i}{n}=\frac{44}{8}=5'5\,;\bar{y}=\frac{\sum_{i=1}^{6}y_i}{n}=\frac{53}{8}=6'625$$

❂ Desviaciones típicas:

$$s_x=\sqrt{\frac{\sum_{i=1}^{6}x_i^2}{n}-\bar{x}^2}=\sqrt{\frac{256}{8}-5'5^2}=\sqrt{1'75}=1'32\,;s_y=\sqrt{\frac{\sum_{i=1}^{6}y_i^2}{n}-\bar{y}^2}=\sqrt{\frac{389}{8}-6'625^2}=\sqrt{4'734}=2'18$$

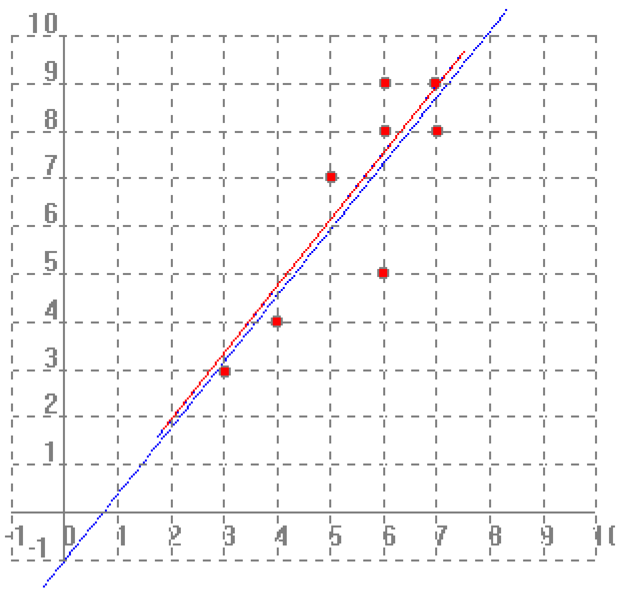
❂ Covarianza :

$$s_{xy}=\frac{\sum_{i=1}^{6}x_iy_i}{n}-\bar{x}\cdot\bar{y}=\frac{311}{8}-5'5\cdot 6'625=2'4375$$

❂ Recta de regresión :

$$y-\bar{y}=\frac{s_{xy}}{s_x^2}(x-\bar{x})\rightarrow y-6'625=\frac{2'4375}{1'75}(x-5'5)\Leftrightarrow y-6'625=1'39(x-5'5)\Leftrightarrow y=1'39x-1'02$$

Comparamos con la estimada (en rojo) la calculada (en azul) :



❈ ❈ ❈ ❁ ❁ ❂ ❁ ❁ ❈ ❈ ❈

8 Con los datos del problema anterior halla la recta de regresión de X sobre Y. ¿Qué nota habría sacado un alumno en el primer examen si obtuvo un 10 en el segundo?

◊◊◊▣▣⊙▣▣◊◊◊

❂ Recta de regresión de X sobre Y:

$$x-\bar{x}=\frac{s_{xy}}{s_y^2}(y-\bar{y}) \rightarrow x-5'5=\frac{2'4375}{4'734}(y-6'625) \Leftrightarrow x-5'5=0'515(y-6'625) \Leftrightarrow x=0'515y+2'09$$

La nota del primer examen estimada mediante la recta de regresión de X sobre Y sería :

x = 0’515 ·10 + 2’09 = 7’24

❄❄❄❀❀❂❀❀❄❄❄

9 Halla el coeficiente de correlación lineal y las rectas de regresión, de Y sobre X y de X sobre Y, del peso y la estatura de las alumnas consideradas en el Ejercicio de aplicación 3. ¿ Cuánto medirá una alumna si pesa 65 kg ? ¿ Cuánto pesará una alumna que mide 162 cm?

◊◊◊▣▣⊙▣▣◊◊◊

Peso (kg), x:	45	55	43	47	51	60	63	58
Estatura (cm), y:	1,58	1,65	1,55	1,61	1,54	1,67	1,62	1,71

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
45	1’58	2025	2’4964	71’1
55	1’65	3025	2’7225	90’75
43	1’55	1849	2’4025	66’65
47	1’61	2209	2’5921	75’67
51	1’54	2601	2’3716	78’54
60	1’67	3600	2’7889	100’2
63	1’62	3969	2’6244	102’06
58	1’71	3364	2’9241	99’18
$\sum x_i=422$	$\sum y_i=12'93$	$\sum x_i^2=22642$	$\sum y_i^2=20'9225$	$\sum x_i \cdot y_i=684'15$

❂ Medias :

$$\bar{x}=\frac{\sum_{i=1}^{6}x_i}{n}=\frac{422}{8}=52'75\,;\bar{y}=\frac{\sum_{i=1}^{6}y_i}{n}=\frac{12'93}{8}=1'62$$

❂ Desviaciones típicas:

$$s_x=\sqrt{\frac{\sum_{i=1}^{6}x_i^2}{n}-\bar{x}^2}=\sqrt{\frac{22642}{8}-52'75^2}=\sqrt{47'69}=6'91\,;s_y=\sqrt{\frac{\sum_{i=1}^{6}y_i^2}{n}-\bar{y}^2}=\sqrt{\frac{20'9225}{8}-1'62^2}=\sqrt{0'0031}=0'055$$

❂ Covarianza y coeficiente de correlación :

$$s_{xy}=\frac{\sum_{i=1}^{6}x_iy_i}{n}-\overline{x}\cdot\overline{y}=\frac{684'15}{8}-52'75\cdot1'616=0'275 \qquad r=\frac{s_{xy}}{s_x\cdot s_y}=\frac{0'275}{6'91\cdot0'055}=0'723$$

❂ **Recta de regresión de Y sobre X :**

$$y-\overline{y}=\frac{s_{xy}}{s_x^2}(x-\overline{x})\rightarrow y-1'616=\frac{0'275}{47'69}(x-52'75)\Leftrightarrow y-1'616=0'00577(x-52'75)\Leftrightarrow y=0'00577x+1'312$$

❂ **Recta de regresión de X sobre Y :**

$$x-\overline{x}=\frac{s_{xy}}{s_y^2}(y-\overline{y})\rightarrow x-52'75=\frac{0'275}{0'0031}(y-1'616)\Leftrightarrow x-52'75=88'71(y-1'616)\Leftrightarrow x=88'7y-90'60$$

❂ La alumna que pesa 65 kg medirá :

y = 0’00577·65 +1’312 = 1’69 m.

❂ La alumna que mida 1’62 m :

x = 88’7 · 1’62 – 90’6 = 53’09 kg.

❄❄❄❁❁❂❁❁❄❄❄

10 Se ha experimentado sobre la distancia de frenada de un coche dependiendo de su velocidad, obteniéndose los siguientes datos:

Velocidad (km/h) (x)	70	50	45	120	85	65
Distancia (m) (y)	32	18	19	43	35	34

Según estos datos, ¿qué distancia de seguridad habría que mantener con el vehículo que va delante si circulamos a 100 km/h?

◈◈◈▣▣⊙▣▣◈◈◈

Hay que hallar la recta de regresión de Y sobre X :

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
70	32	4900	1024	2240
50	18	2500	324	900
45	19	2025	361	855
120	43	14400	1849	5160
85	35	7225	1225	2975
65	34	4225	1156	2210
$\sum x_i=435$	$\sum y_i=181$	$\sum x_i^2=35275$	$\sum y_i^2=5939$	$\sum x_i\cdot y_i=14340$

❂ **Medias :**

$$\overline{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{435}{6} = 72'5 \; ; \; \overline{y} = \frac{\sum_{i=1}^{6} y_i}{n} = \frac{181}{6} = 30'1\widehat{6}$$

❂ **Desviación típica :**

$$s_x = \sqrt{\frac{\sum_{i=1}^{6} x_i^2}{n} - \overline{x}^2} = \sqrt{\frac{35275}{6} - 72'5^2} = \sqrt{622'91\widehat{6}} = 24'9583$$

❂ **Covarianza :**

$$s_{xy} = \frac{\sum_{i=1}^{6} x_i y_i}{n} - \overline{x} \cdot \overline{y} = \frac{14340}{6} - 72'5 \cdot 30'1\widehat{6} = 202'92$$

❂ **Recta de regresión de Y sobre X :**

$$y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x}) \to y - 30'1\widehat{6} = \frac{202'92}{622'91\widehat{6}}(x - 72'5) \Leftrightarrow y - 30'1\widehat{6} = 0'3257(x - 72'5) \Leftrightarrow y = 0'325x + 6'549$$

La distancia de frenado esperada será :

y = 0’325 · 100 + 6’549 = 32’5 + 6’549 = 39’05 m

❄❄❄❁❁❂❁❁❄❄❄

11 A un grupo de 12 alumnos de ESO se le han realizado pruebas para determinar su: nivel de vocabulario (NV); nivel intelectual (NI); velocidad lectora (VL); agilidad de cálculo (AC). Los resultados han sido los siguientes:

NV	28	27	14	17	18	14	23	24	14	12	16	10
NI	43	30	18	21	24	20	23	19	22	14	10	18
VL	69	68	38	37	48	50	50	57	33	17	42	35
AC	28	22	15	19	20	14	25	19	20	9	7	10

Halla la correlación entre NV y NI. Interpreta el resultado

◊◊◊▣▣⊙▣▣◊◊◊

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
28	43	784	1849	1204
27	30	729	900	810
14	18	196	324	252
17	21	289	441	357
18	24	324	576	432
14	20	196	400	280
23	23	529	529	529
24	19	576	361	456
14	22	196	484	308
12	14	144	196	168
16	10	256	100	160
10	18	100	324	180
$\sum x_i = 217$	$\sum y_i = 262$	$\sum x_i^2 = 4319$	$\sum y_i^2 = 6484$	$\sum x_i \cdot y_i = 5136$

❂ **Medias :**

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{217}{12} = 18'08\widehat{3}\ ;\ \bar{y} = \frac{\sum_{i=1}^{6} y_i}{n} = \frac{262}{12} = 21'8\widehat{3}$$

❂ **Desviaciones típicas:**

$$s_x = \sqrt{\frac{\sum_{i=1}^{6} x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{4319}{12} - 18'08\widehat{3}^2} = \sqrt{32'91} = 5'737\ ;\ s_y = \sqrt{\frac{\sum_{i=1}^{6} y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{6484}{12} - 21'8\widehat{3}^2} = \sqrt{63'64} = 7'977$$

❂ **Covarianza y coeficiente de correlación :**

$$s_{xy} = \frac{\sum_{i=1}^{6} x_i y_i}{n} - \bar{x}\cdot\bar{y} = \frac{5136}{12} - 18'08\widehat{3}\cdot 21'8\widehat{3} = 33'1806 \qquad r = \frac{s_{xy}}{s_x \cdot s_y} = \frac{33'18056}{5'737\cdot 7'977} = 0'725$$

❄ ❄ ❄ ❁ ❁ ❂ ❁ ❁ ❄ ❄ ❄