EJERCICIOS DE ESTADÍSTICA BIDIMENSIONAL

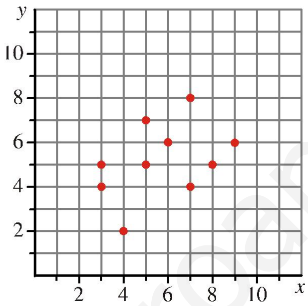
EJERCICIO 1

Considera la siguiente distribución bidimensional:

x	**9**	**5**	**4**	**3**	**7**	**5**	**6**	**7**	**8**	**3**
y	**6**	**7**	**2**	**4**	**8**	**5**	**6**	**4**	**5**	**5**

Representa los datos mediante una nube de puntos y di si crees que existe una correlación fuerte entre ellos y si es directa

Solución:



Vemos que hay una relación positiva y mediana entre las variables

EJERCICIO 2

Se ha estudiado en distintas marcas de yogures naturales el porcentaje de grasa que contenían, así como las kilocalorías por envase. Estos son los resultados obtenidos en seis de ellos:

X: **Grasa (%)**	**2,2**	**2**	**1,9**	**3,1**	**3**	**2**
Y: **Kcal/envase**	**64**	**55**	**58**	**79**	**65**	**52**

A) Halla la recta de regresión de *Y* sobre *X*

B) Estima las kcal que le corresponderán a un yogur con 2.2% de grasa

Solución:

· Medias:

$$\bar{x} = \frac{14,2}{6} = 2,37\% \text{ de grasa}$$

$$\bar{y} = \frac{373}{6} = 62,17 \text{ Kcal/envase}$$

· Varianza de *x*:

$$\sigma_x^2 = \frac{35,06}{6} - 2,37^2 = 0,23$$

· Covarianza:

$$\sigma_{xy} = \frac{904,9}{6} - 2,37 \cdot 62,17 = 3,47$$

· Ecuación de la recta de regresión de *Y* sobre *X*:

$$m_{yx} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{3,47}{0,23} = 15,1$$

La recta es $y - \bar{y} = \frac{\sigma_{xy}}{\sigma_{xy}^2}(x - \bar{x})$

Y - 62.17 = 15,1 (x - 2,37)

$$y = 62,17 + 15,1(x - 2,37) \quad \rightarrow \quad y = 15,1x + 26,38$$

Para estimar las kcal que le corresponderán a un yogur con 2.2% de grasa, basta sustituir en la recta de regresión:

b) $\hat{y}(2,5) = 15,1 \cdot 2,5 + 26,38 = 64,13$ kcal

EJERCICIO 3

Con los datos de la tabla siguiente:

X	2	3	6	10	12
Y	150	130	125	120	100

a) Dibuja la nube de puntos

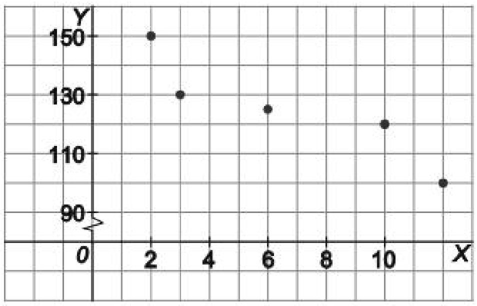
b) Halla la media y la varianza de las distribuciones *X* e *Y*.

c) Obtén la recta de regresión de *Y* sobre *X* y valora la bondad del ajuste.

d) ¿Qué porcentaje de la variabilidad de *Y* viene explicado por la variabilidad de *X*?

e) ¿Qué valor se espera en la variable *Y* si la variable *X* toma el valor *x*=5?

Solución:



a) Se representa el diagrama de dispersión a la derecha.

b)

x_i	y_i	x_i^2	y_i^2	x_iy_i
2	150	4	22500	300
3	130	9	16900	390
6	125	36	15625	750
10	120	100	14400	1200
12	100	144	10000	1200
33	625	293	79425	3840

$$\overline{X} = \frac{33}{5} = 6,6 \quad ; \quad s_X^2 = \frac{293}{5} - 6,6^2 = 15,04$$

$$\overline{Y} = \frac{625}{5} = 125 ; \quad s_Y^2 = \frac{79425}{5} - 125^2 = 260$$

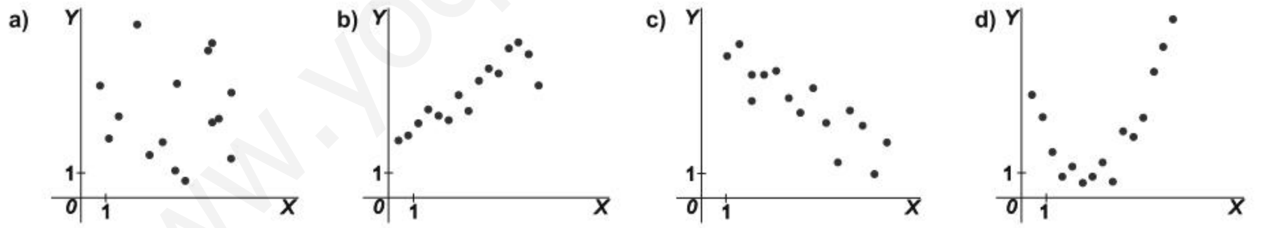
c) Se obtiene, en primer lugar, la covarianza: $s_{XY} = \frac{3840}{5} - 125 \cdot 6,6 = -57$, que junto con los resultados del apartado b) proporcionan los coeficientes de la recta de regresión de *Y* sobre *X*.

$$b = \frac{s_{XY}}{s_X^2} = \frac{-57}{15,04} = -3,79 \quad ; \quad a = \overline{Y} - b\,\overline{X} = 125 + 3,79 \cdot 6,6 = 150,01$$

La recta de regresión de *Y* sobre *X* viene dada por la ecuación: $y = 150,01 - 3,79x$

EJERCICIO 4

En los siguientes casos, se representa la nube de puntos de una variable bidimensional.



En cada caso indica si existe relación lineal entre las variables y, en caso afirmativo, ¿cuál es el signo de la covarianza y del coeficiente de correlación?

Solución:

a) En este caso no parece que exista relación lineal entre las variables representadas.

b) Se puede observar una relación lineal directa, cuando una variable crece la otra también,

c) Puede observarse una relación lineal con tendencia inversa (cuando una variable crece la otra decrece).

d) Parece que existe relación entre ambas variables, pero que esta no es lineal.