DISTRIBUCIONES BIDIMENSIONALES

Las numeraciones indicadas entre páginas se refieren a las páginas del libro de matemáticas aplicadas a las ciencias sociales I, de primero de bachillerato de la editorial Anaya, Andalucía, cuyos autores son J. Colera, R. García y M.J.Oliveira.

Página 228

EJERCICIOS Y PROBLEMAS PROPUESTOS

PARA PRACTICAR

1 **Para cada uno de los siguientes casos indica:**

- **Cuáles son las variables que se relacionan.**
- **Cuál es el colectivo de individuos que se estudia.**
- **Si se trata de una relación funcional o de una relación estadística.**
- **El signo de la correlación.**

a) Familias: estatura media de los padres – estatura media de los hijos mayores de 17 años.

b) Entre los países europeos: volumen de exportación – volumen de importación (con España).

c) Entre los países del mundo: índice de mortalidad infantil – número de médicos por cada 1 000 habitantes.

d) kW · h consumidos en cada casa de una ciudad durante el mes de enero – coste del recibo de la luz.

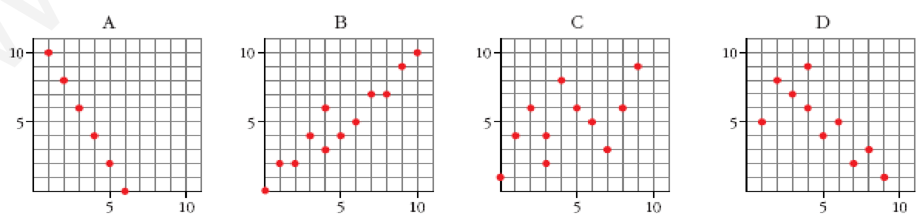
e) Coste del recibo de la luz – número de personas que viven en cada casa.

Las variables que se relacionan están claras en todos los casos. El colectivo sobre el que se hace el estudio también está claro salvo, acaso, en los apartados d) y e), en qué es un grupo de casas (todas las de una barriada, una ciudad, un país…).

Solo hay relación funcional en d), el resto son relaciones estadísticas.

La correlación es positiva en a), d) y e), y es negativa en b) y c).

2 **a) Traza, a ojo, la recta de regresión en cada una de estas distribuciones bidimensionales:**



b) ¿Cuáles de ellas tienen correlación positiva y cuáles tienen correlación negativa?

c) Una de ellas presenta relación funcional. ¿Cuál es? ¿Cuál es la expresión analítica de la función que relaciona las dos variables?

d) Ordena de menor a mayor las correlaciones.

b) B y C tienen correlación positiva; A y D, negativa.

c) La A es relación funcional: $y = 12 - 2x$.

d) C, D, B, A (prescindiendo del signo).

3 **Una distribución bidimensional en la que los valores de x son 12, 15, 17, 21, 22 y 25, tiene una correlación $r = 0,99$ y su recta de regresión e $y = 10,5 + 3,2x$. Calcula $\hat{y}(13)$, $\hat{y}(20)$, $\hat{y}(30)$, $\hat{y}(100)$.**

¿Cuáles de las estimaciones anteriores son fiables, cuál poco fiable y cuál no se debe hacer?

Expresa los resultados en términos adecuados. (Por ejemplo: $\hat{y}(13) = 52,1$. Para $x = 13$ es muy probable que el valor correspondiente de y sea próximo a 52.)

$$\hat{y}(13) = 52,1; \quad \hat{y}(20) = 74,5; \quad \hat{y}(30) = 106,5; \quad \hat{y}(100) = 330,5$$

Son fiables $\hat{y}(13)$ e $\hat{y}(20)$, porque 13 y 20 están en el intervalo de valores utilizados para obtener la recta de regresión.

$\hat{y}(30)$ es menos fiable, pues 30 está fuera del intervalo, aunque cerca de él.

$\hat{y}(100)$ es una estimación nada fiable, pues 100 está muy lejos del intervalo [12, 25]

4 **Los parámetros correspondientes a esta distribución bidimensional son:**

x	0	1	2	3	3	4	5	6	7	8	9
y	1	4	6	2	4	8	6	5	3	6	9

$\bar{x} = 4,4$ $\quad \bar{y} = 4,9$ $\quad \sigma_{xy} = 3,67$

$\sigma_x = 2,77$ $\quad \sigma_y = 2,31$ $\quad r = 0,57$

Halla las ecuaciones de las dos rectas de regresión, X sobre Y e Y sobre X, y represéntalas junto con la nube de puntos.

$$m_{yx} = \frac{\sigma_{xy}}{\sigma_x^2} = 0,48$$

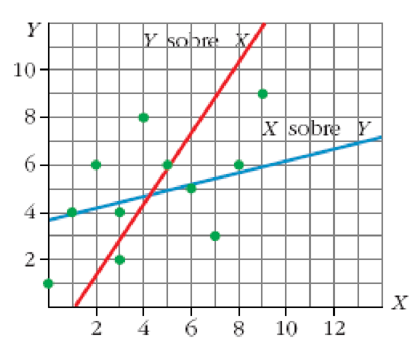
Recta de regresión de Y sobre X:

$$y = 4,9 + 0,48(x - 4,4) \;\rightarrow\; y = 0,48x + 2,79$$

$$m_{xy} = \frac{\sigma_{xy}}{\sigma_y^2} = 0,69$$

Recta de regresión de X sobre Y:

$$x = 4,4 + 0,69(y - 4,9) \;\rightarrow\; y = 1,45x - 1,48$$



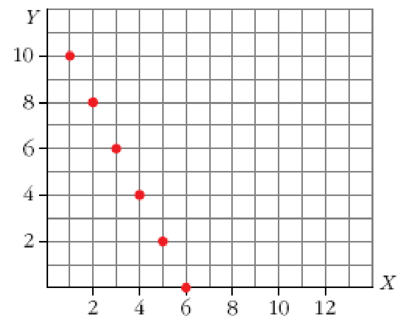
5 **Representa estos puntos y, sin efectuar cálculos, contesta las siguientes preguntas:**

x	1	2	3	4	5	6
y	10	8	6	4	2	0

a) ¿Cuánto vale el coeficiente de correlación?

b) ¿Cómo son las dos rectas de regresión? Escribe su ecuación.

c) A la vista de la respuesta anterior, da el valor de m_{yx} y el de m_{xy}.



a) Los puntos están alineados todos ellos sobre la recta $y = 12 - 2x$. Por tanto, el coeficiente de correlación es –1: $r = -1$.

b) Las dos rectas de regresión son coincidentes. Su ecuación es $y = 12 - 2x$.

c) $m_{yx} = -2$ (pendiente de la recta de regresión de Y sobre X).

$m_{xy} = -1/2$

6 **Calcula el coeficiente de correlación entre estas dos variables:**

x: ALTITUD	365	450	350	220	150
y: LITROS DE LLUVIA	240	362	121	145	225

$r = 0,5$

7 **La media de los pesos de los individuos de una población es de 65 kg y la de sus estaturas, 170 cm.**

Las desviaciones típicas son 5 kg y 10 cm, respectivamente, y la covarianza de ambas variables es 40.

a) ¿Cuál es el coeficiente de correlación?

b) Calcula la recta de regresión de los pesos respecto de las estaturas.

c) ¿Cuánto estimas que pesará un individuo de 180 cm de estatura?

a) $r = 0,8$

b) $y = 65 + 0,4(x - 170) = 0,4x - 3 \rightarrow \begin{cases} x: \text{ estaturas en cm} \\ y: \text{ pesos en kg} \end{cases}$

c) $\hat{y}(180) = 69$ kg

Página 229

PARA RESOLVER

8 **Estudia la correlación entre estas dos variables y explica el resultado:**

	Esp.	Hol.	Gre.	Ita.	Irl.	Fra.	Din.	Bél.	Lux.	Al.	R.U.
ÍNDICE MORTALIDAD	7,4	8,2	8,7	9,4	9,4	10	10,8	11,1	11,3	11,6	11,8
MAYORES 64 AÑOS	11,3	11,6	13,2	13,6	10,7	15,4	14,5	14,4	13,5	15,3	15,3

$r = 0,77$

Hay una clara relación entre las dos variables.

9 **De un muelle se cuelgan pesas y se obtienen los siguientes alargamientos:**

x: MASA DE LA PESA (g)	0	10	30	60	90	120	150	200	250	350
y: ALARGAMIENTO (cm)	0	0,5	1	3	5	6,5	8	10,2	12,5	18

Halla la recta de regresión de Y sobre X y estima el alargamiento que se conseguirá con pesos de 100 g y de 500 g. ¿Cuál de las dos estimaciones es más fiable?

$r = 0{,}999$; $y = -0{,}01 + 0{,}051x$

100 g $\rightarrow$ 5,09 cm

500 g $\rightarrow$ 25,49 cm (esta última es menos fiable).

10 **La siguiente tabla muestra el número de gérmenes patógenos por centímetro cúbico de un determinado cultivo según el tiempo transcurrido:**

Nº DE HORAS	0	1	2	3	4	5
Nº DE GÉRMENES	20	26	33	41	47	53

a) Calcula la recta de regresión para predecir el número de gérmenes por cm^3 en función del tiempo.

b) ¿Qué cantidad de gérmenes por cm^3 es predecible encontrar cuando hayan transcurrido 6 horas? ¿Es buena esa predicción?

a) $y = 19{,}81 + 6{,}74x$, donde: $x \rightarrow$ número horas, $y \rightarrow$ número de gérmenes

b) $\hat{y}(6) = 60{,}25 \approx 60$ gérmenes.

Es una buena predicción, puesto que $r = 0{,}999$ (y 6 está cercano al intervalo de valores considerado).

11 **En un depósito cilíndrico, la altura del agua que contiene varía conforme pasa el tiempo según la siguiente tabla:**

TIEMPO (h)	8	22	27	33	50
ALTURA (m)	17	14	12	11	6

a) Halla el coeficiente de correlación lineal entre el tiempo y la altura e interprétalo.

b) ¿Cuál será la altura del agua cuando hayan transcurrido 40 horas?

c) Cuando la altura del agua es de 2 m, suena una alarma. ¿Qué tiempo ha de pasar para que avise la alarma?

a) $r = -0{,}997$. Hay una relación muy fuerte entre las dos variables, y negativa. A medida que pasa el tiempo, la altura va bajando (se va consumiendo el agua).

b) La recta de regresión es $y = 19{,}37 - 0{,}26x$, donde: $x \rightarrow$ tiempo, $y \rightarrow$ altura.

$\hat{y}(40) = 8{,}97$ m

c) $2 = 19{,}37 - 0{,}26x \rightarrow x = 66{,}8$ h

12 **En una cofradía de pescadores, las capturas registradas de cierta variedad de pescados, en kilogramos, y el precio de subasta en lonja, en euros/kg, fueron los siguientes:**

x (kg)	2 000	2 400	2 500	3 000	2 900	2 800	3 160
y (euros/kg)	1,80	1,68	1,65	1,32	1,44	1,50	1,20

a) ¿Cuál es el precio medio registrado?

b) Halla el coeficiente de correlación lineal e interprétalo.

c) Estima el precio que alcanzaría en lonja el kilo de esa especie si se pescasen 2 600 kg.

a) $\bar{y} = 1{,}51$ euros

b) $r = -0{,}97$. La relación entre las variables es fuerte y negativa. A mayor cantidad de pescado, menor es el precio por kilo.

c) La recta de regresión es $y = 2{,}89 - 0{,}0005x$

$\hat{y}(2\,600) = 1{,}59$ euros

Página 230

13 **Durante 10 días, hemos realizado mediciones sobre el consumo de un coche (litros consumidos y kilómetros recorridos). Los datos obtenidos han sido los siguientes:**

x (km)	100	80	50	100	10	100	70	120	150	220
y (l)	6,5	6	3	6	1	7	5,5	7,5	10	15

a) Halla el coeficiente de correlación lineal y la recta de regresión de Y sobre X.

b) Si queremos hacer un viaje de 190 km, ¿qué cantidad de combustible debemos poner?

a) $r = 0{,}99$; $y = 0{,}157 + 0{,}066x$

b) $\hat{y}(190) = 12{,}697$ litros. Debemos poner, como mínimo, unos 13 litros.

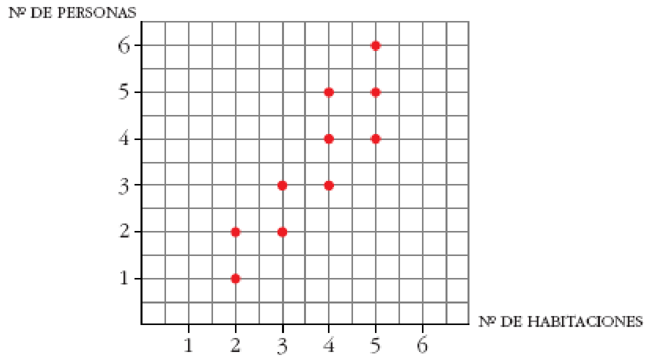
14 **En una zona de una ciudad se ha tomado una muestra para estudiar el número de habitaciones de que dispone un piso y el de personas que viven en él, obteniéndose estos datos:**

Nº DE HABITACIONES	2	2	3	3	4	4	4	5	5	5
Nº DE PERSONAS	1	2	2	3	3	4	5	4	5	6

a) Representa la nube de puntos.

b) Calcula e interpreta el coeficiente de correlación.

a)



b) r = 0,88. Hay una correlación alta entre las dos variables.

15 **El consumo de energía "per cápita" en miles de kWh y la renta "per cápita" en miles de euros de seis países de la U.E. son las siguientes:**

	ALEMANIA	BÉLGICA	DINAMARCA	ESPAÑA	FRANCIA	ITALIA
CONSUMO (y)	5,7	5,0	5,1	2,7	4,6	3,1
RENTA (x)	11,1	8,5	11,3	4,5	9,9	6,5

a) Calcula la recta de regresión del consumo de energía (y) sobre la renta (x).

b) Indica el coeficiente de correlación entre el consumo y la renta.

c) ¿Qué predicción podemos hacer sobre el consumo de energía "per cápita" de Grecia si su renta es de 4,4 miles de euros?

a) $y = 0{,}8 + 0{,}4x$

b) $r = 0{,}93$

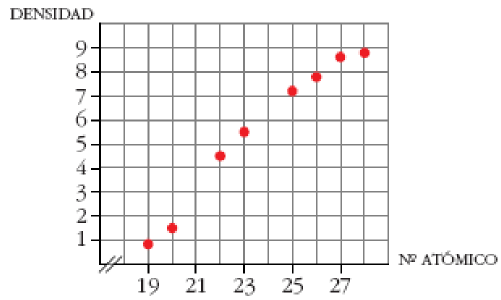
c) $\hat{y}(4{,}4) = 2{,}56$ kWh

16 **La siguiente tabla relaciona el número atómico de varios metales de la misma fila en el sistema periódico (periodo 4), con su densidad:**

ELEMENTO	K	Ca	Ti	V	Mn	Fe	Co	Ni
NÚMERO ATÓMICO	19	20	22	23	25	26	27	28
DENSIDAD (g/cm^3)	0,86	1,54	4,5	5,6	7,11	7,88	8,7	8,8

Representa los puntos, calcula el coeficiente de correlación y halla la ecuación de la recta de regresión. A partir de ella, estima la densidad del cromo (Cr), cuyo número atómico es 24.

Haz otro tanto con la del escandio (Sc), de número atómico 21.



$r = 0{,}98$ $\hat{y} = -16{,}5 + 0{,}93x$ $\hat{y}(24) = 5{,}86$ $\hat{y}(21) = 3{,}06$

Las densidades del Cr y del Sc son, aproximadamente, 5,86 y 3,01. (Los valores reales de estas densidades son 7,1 y 2,9.)

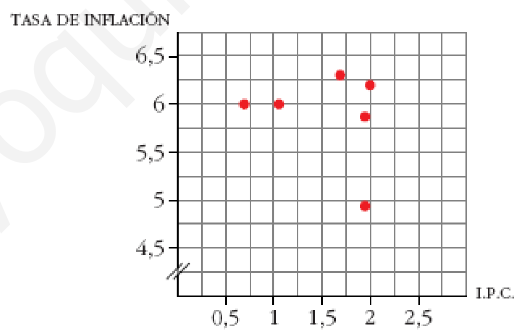
La evolución del *IPC* (índice de precios al consumo) y de la tasa de inflación en 1987 fue:

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO
IPC	0,7	1,1	1,7	2	1,9	1,9
TASA DE INFLACIÓN	6	6	6,3	6,2	5,8	4,9

a) Representa la nube de puntos.

b) Calcula el coeficiente de correlación entre el *IPC* y la tasa de inflación.

c) ¿Se puede estimar la tasa de inflación a partir del *IPC*?



$r = -0{,}24$. La nube de puntos es muy dispersa. No se puede estimar de forma fiable la tasa de inflación a partir del IPC (pues $|r|$ es muy bajo).