

EXAMEN ESTADÍSTICA 1

1.- Con la variable edad, en años, de una muestra de 100 personas se forma la siguiente tabla de frecuencias:

Edad en años	Frecuencia acumulada
[10,30)	10
[30,50)	30
[50,70)	60
[70,90)	84
[90,110)	100

- Completa la tabla de frecuencias. (0,75 puntos)
- Usando la tabla, calcula la media y la desviación típica. (1 punto)
- Calcula exactamente la moda. (0,75 puntos)
- Calcula exactamente la mediana. (0,75 puntos)
- ¿A partir de qué valor podemos encontrar el 25% de las personas más jóvenes? (1 punto)

2.- Hemos calculado la covarianza de una cierta distribución y ha resultado negativa. Justifica por qué podemos afirmar que tanto el coeficiente de correlación como las pendientes de las dos rectas de regresión son también números negativos. (1 punto)

3.- Sean x = “gastos en publicidad de un producto” (en miles de euros) e y = “ventas conseguidas de ese producto” (en miles de euros).

x	1	2	3	4	5	6
y	10	17	30	28	39	47

- Calcula el coeficiente de correlación. (1 punto)
- Comenta la influencia del gasto publicitario en las ventas. (1 punto)
- Halla la ecuación de la recta de regresión de las ventas respecto del gasto publicitario. (1 punto)
- Halla las ventas esperadas para un gasto en publicidad de 3200 euros. (1 punto)
- Dibuja la nube de puntos y la recta de regresión hallada en el apartado c). (0,75 puntos)

SOLUCIONES

1.-a)

Edad en años	x_i	f_i	F_i	h_i	H_i	$x_i \cdot f_i$	$(x_i - \bar{x})$	$f_i(x_i - \bar{x})^2$
[10,30)	20	10	10	0'1	0'1	200	-43'20	18662'4
[30,50)	40	20	30	0'2	0'3	800	-23'20	10764'8
[50,70)	60	30	60	0'3	0'6	1800	-3'20	307'2
[70,90)	80	24	84	0'24	0'84	1920	16'80	6773'76
[90,110)	100	16	100	0'16	1	1600	36'80	21667'84

$$b) \bar{x} = \frac{\sum x_i f_i}{n} = \frac{6320}{100} = 63'20 \text{ años} \quad \sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{n}} = \sqrt{5817'6} = 24'12 \text{ años}$$

$$c) \text{Intervalo modal [50,70) (} f_i \text{ mayor) } Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot c = 62'5$$

d) Intervalo mediano: [50,70) ya que $F_i = 60 > 50$

$$M = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot c = 50 + \frac{50 - 30}{30} \cdot 20 = 63'33$$

e) Nos piden $P_{25} = Q_1$; está en el intervalo [30,50) ya que $n/4 = 25$

$$Q_3 = L_i + \frac{\frac{n}{4} - F_{i-1}}{f_i} \cdot c = 30 + \frac{25 - 10}{20} \cdot 20 = 45 \text{ luego el 25\% está por debajo de 45 años}$$

2.- Si la covarianza es negativa ($\sigma_{xy} < 0$) entonces como el coeficiente de correlación es

$\frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$ y las desviaciones típicas son siempre positivas, tendremos que el coeficiente es también negativo.

Por otra parte, las pendientes de las rectas de regresión son: $\frac{\sigma_{xy}}{\sigma_x^2}$ y $\frac{\sigma_{xy}}{\sigma_y^2}$ que, por la misma razón son también negativas.

3.- a)

x	y	x_i^2	y_i^2	$x_i \cdot y_i$
1	10	1	100	10
2	17	4	289	34
3	30	9	900	90
4	28	16	784	112
5	39	25	1521	195
6	47	36	2209	282
21	171	91	5803	723

$$\text{Medias: } \bar{x} = \frac{21}{6} = 3'5 \text{ miles de euros (3500 €)} \quad \bar{y} = \frac{171}{6} = 28'5 \text{ (28500 €)}$$

$$\text{Desviaciones típicas: } \sigma_x = \sqrt{\frac{91}{6} - 3'5^2} = 1,71 \quad \sigma_y = \sqrt{\frac{5803}{6} - 28'5^2} = 12'45$$

$$\text{Covarianza: } \sigma_{xy} = \frac{723}{6} - 3'5 \cdot 28'5 = 20'75$$

$$\text{Coeficiente de correlación: } r = \frac{20'75}{1'71 \cdot 12'45} = 0,97$$

b) Es una correlación muy fuerte (próxima a 1) y positiva. Lo que significa que a más gastos en publicidad mas ventas, con una incidencia muy clara.

c) tenemos que hallar la recta de y sobre x:

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x}); \quad y - 28'5 = \frac{20'75}{1'71^2} (x - 3'5); \quad y - 28'5 = 7'1(x - 3'5)$$

$$y - 28'5 = 7'1x - 24'85, \text{ luego la recta pedida es: } y = 7'1x + 3'65$$

d) Para un gasto de 3200 euros ($x = 3'2$) en publicidad, las ventas esperadas serán:

$$y = 7'1 \cdot 3'2 + 3'65 = 26'37 \text{ miles de euros, es decir } \mathbf{26370 \text{ euros.}}$$

e)

