

MATEMÁTICAS APLICADAS A LAS CC. SOCIALES I

Examen estadística descriptiva y estadística bidimensional

Nombre y apellidos _____ Fecha: 4 de mayo de 2010

1.- (2 puntos) Una empresa de publicidad hace una encuesta entre los lectores de una revista para saber su edad aproximada y estudiar si deben anunciarse en la revista.
Las respuestas obtenidas se reflejan en la siguiente tabla:

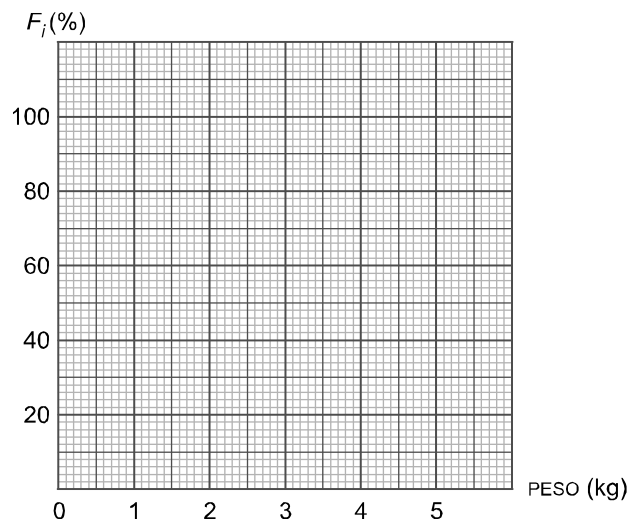
Edad	[10,13)	[13,16)	[16,19)	[19,22)	[22,25)	[25,28)
N.º de lectores	110	248	115	20	4	3

- Calcula la media y la desviación típica. Utiliza marcas de clase para la edad.
- ¿Es un grupo homogéneo o disperso en cuanto a la edad? Si la empresa anuncia productos para adolescentes, ¿les conviene poner un anuncio en esa revista?
- En otra revista *B*, la media de edad de los lectores es de 37,53 años con una desviación típica de 5,21. Calcula el coeficiente de variación en ambos casos y compara la dispersión en los dos grupos.

2.- (2 puntos) En un hospital materno-infantil se está haciendo un estudio sobre el peso alcanzado por 200 recién nacidos. La siguiente tabla refleja los resultados obtenidos:

Peso (kg)	[1,5,2)	[2,2'5)	[2'5, 3)	[3, 3'5)	[3'5, 4)	[4, 4'5)
Nº de bebés	4	16	54	65	35	26

Dibuja el polígono de frecuencias acumuladas. Halla Me y p_{80} .



3.- (1 punto) Comprueba que el producto de los coeficientes de regresión es igual al cuadrado del coeficiente de correlación.

4.- (2 puntos) En una determinada especie de tortugas se han medido la longitud y la anchura de una serie de seis ejemplares, obteniendo los siguientes resultados:

Longitud (cm)	31	40	45	33	35	38
Anchura (cm)	17	21	25	16	18	19

Calcula la covarianza y el coeficiente de correlación.
¿Cómo es la relación entre las dos variables?

5.- (2 puntos) Las notas obtenidas en Matemáticas en la primera y en la segunda evaluación por un grupo de seis alumnos de 1.º de Bachillerato han sido las siguientes:

1.ª evaluación	4,5	8	6,5	4	9	8
2.º evaluación	3,5	7	6	4	8,5	7,5

- Halla la recta de regresión que me permita intuir la nota de la segunda evaluación conociendo la nota de la primera.
- Calcula $y(6)$. ¿Es fiable esta estimación? (sabemos que $r=0,98$)

6.- (1 punto) Sabemos que dos rectas de regresión de una distribución bidimensional son:

$$y=8,7-0,76x$$

$$y=11,36-1,3x$$

Calcula la media de x y la de y .

Ejercicio nº 1.-

1.- Una empresa de publicidad hace una encuesta entre los lectores de una revista para saber su edad aproximada y estudiar si deben anunciarse en la revista.

Las respuestas obtenidas se reflejan en la siguiente tabla:

Edad	[10,13)	[13,16)	[16,19)	[19,22)	[22,25)	[25,28)
N.º de lectores	110	248	115	20	4	3

- Calcula la media y la desviación típica. Utiliza marcas de clase para la edad.
- ¿Es un grupo homogéneo o disperso en cuanto a la edad? Si la empresa anuncia productos para adolescentes, ¿les conviene poner un anuncio en esa revista?
- En otra revista *B*, la media de edad de los lectores es de 37,53 años con una desviación típica de 5,21. Calcula el coeficiente de variación en ambos casos y compara la dispersión en los dos grupos.

Solución:

a) Hallamos la marca de clase, x_i , de cada intervalo y hacemos la tabla:

Intervaleo	x_i	f_i	$f_i x_i$	$f_i x_i^2$
[10,13)	11,5	110	1265	14547,5
[13,16)	14,5	248	3596	52142
[16,19)	17,5	115	2012,5	35218,75
[19,22)	20,5	20	410	8405
[22,25)	23,5	4	94	2209
[25,28)	26,5	3	79,5	2106,75
		500	7457	11429

Media:

$$\bar{x} = \frac{\sum f_i x_i}{n} = \frac{7457}{500} = 14,914$$

Desviación típica:

$$\sigma = \sqrt{\frac{\sum f_i x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{114629}{500} - 14,914^2} = \sqrt{6,83} = 2,61$$

b) Es un grupo bastante homogéneo; la mayoría de los lectores son adolescentes, por tanto, sí les conviene poner el anuncio en la revista.

$$c) C.V._A = \frac{\sigma}{\bar{x}} = \frac{2,61}{14,914} = 0,175; \quad C.V._B = \frac{\sigma}{\bar{x}} = \frac{5,21}{37,53} = 0,139$$

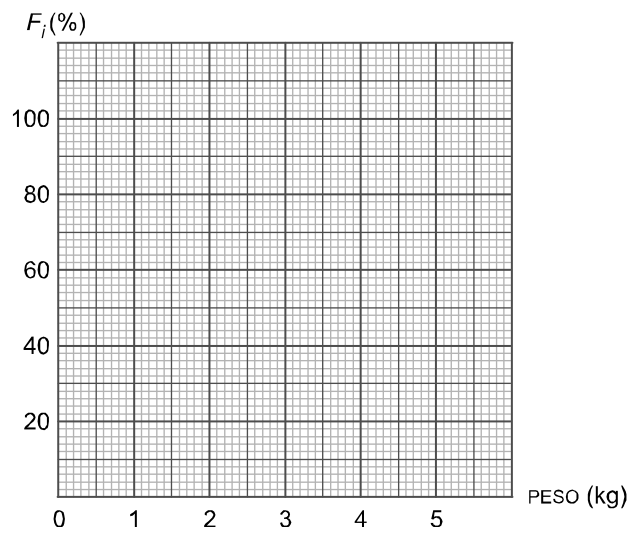
La dispersión es un poco mayor en el primer caso.

Ejercicio nº 2.-

En un hospital materno-infantil se está haciendo un estudio sobre el peso alcanzado por 200 recién nacidos. La siguiente tabla refleja los resultados obtenidos:

Peso(kg)	[1,5; 2)	[2; 2,5)	[2,5; 3)	[3; 3,5)	[3,5; 4)	[4; 4,5)
N.º de bebés	4	16	54	65	35	26

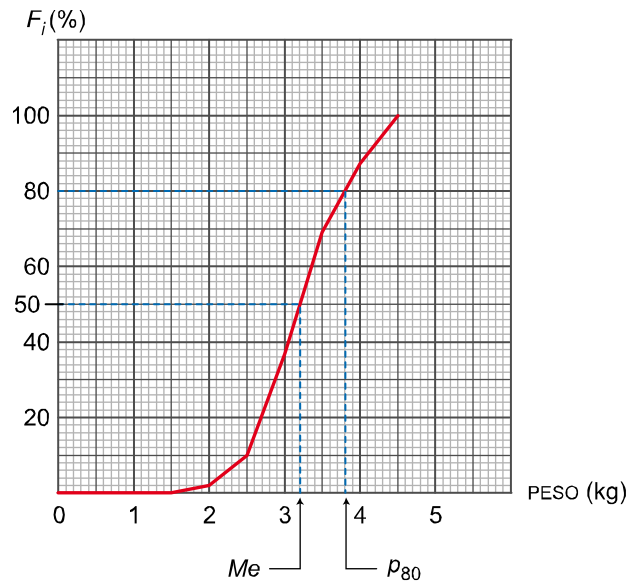
Halla gráficamente y numéricamente Me y p_{80} .
Puedes dibujar aquí el polígono de frecuencias acumuladas.



Solución:

Construimos el polígono de frecuencias acumuladas:

Extremo	F_i	%
1,5	0	0
2	4	2
2,5	20	10
3	74	37
3,5	139	69,5
4	174	87
4,5	200	100

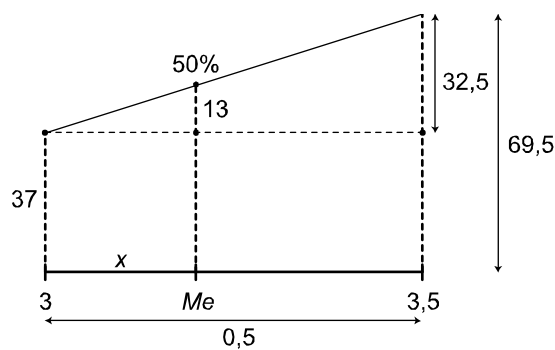


Gráficamente se observa que:

$$Me \approx 3,2 \quad p_{80} \approx 3,8$$

Obtengamos los valores exactos, razonando sobre el polígono de frecuencias:

• Me :

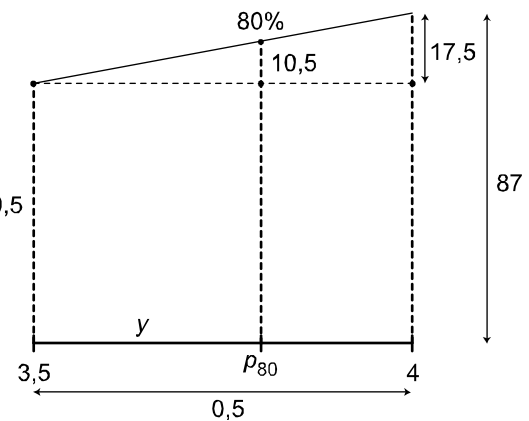


$$\frac{32,5}{0,5} = \frac{3}{x}$$

$$x = 0,2$$

$$Me = 3 + 0,2 = 3,2$$

• p_{80} :



$$\frac{17,5}{0,5} = \frac{10,5}{y}$$

$$y = 0,3$$

$$p_{80} = 3,5 + 0,3 = 3,8$$

Los valores exactos son $Me = 3,2$ y $p_{80} = 3,8$

Ejercicio nº 3.-

El coeficiente de regresión es la pendiente de la recta de regresión:

$$m_{xy} m_{yx} \frac{\delta xy}{(\delta x)^2} \cdot \frac{\delta xy}{(\delta y)^2} = \frac{(\delta xy)^2}{(\delta x \delta y)^2} = \left(\frac{\delta xy}{\delta x \delta y} \right)^2 = r^2$$

Ejercicio nº 4.-

En una determinada especie se han medido la longitud y la anchura de una serie de seis ejemplares, obteniendo los siguientes resultados:

Longitud (cm)	31	40	45	33	35	38
Anchura (cm)	17	21	25	16	18	19

Calcula la covarianza y el coeficiente de correlación.
¿Cómo es la relación entre las dos variables?

Solución:

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	31	17	961	289	527
	40	21	1600	441	840
	45	25	2025	625	1125
	33	16	1089	256	528
	35	18	1225	324	630
	38	19	1444	361	722
Total	222	116	8344	2296	4372

Medias:

$$\bar{x} = \frac{222}{6} = 37$$

$$\bar{y} = \frac{116}{6} = 19,33$$

Desviaciones típicas:

$$\sigma_x = \sqrt{\frac{8344}{6} - 37^2} = \sqrt{21,67} = 4,65$$

$$\sigma_y = \sqrt{\frac{2296}{6} - 19,33^2} = \sqrt{9,02} = 3$$

• Covarianza:

$$\sigma_{xy} = \frac{4372}{6} - 37 \cdot 19,33 = 13,46$$

• Coeficiente de correlación:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{13,46}{4,65 \cdot 3} = 0,96$$

• Hay una relación positiva y fuerte entre las dos variables.

Ejercicio nº 5.-

Las notas obtenidas en Matemáticas en la primera y en la segunda evaluación por un grupo de seis alumnos de 1.º de Bachillerato han sido las siguientes:

1.ª evaluación	4,5	8	6,5	4	9	8
2.ª evaluación	3,5	7	6	4	8,5	7,5

- a) Halla la recta de regresión de Y sobre X .
b) Calcula $\hat{y}(6)$. ¿Es fiable esta estimación? (Sabemos que $r=0,98$).

Solución:

a)

	x_i	y_i	x_i^2	$x_i y_i$
	4,5	3,5	20,25	15,75
	8	7	64	56
	6,5	6	42,25	39
	4	4	16	16
	9	8,5	81	76,5
	8	7,5	64	60
Total	40	36,5	287,5	263,25

Medias:

$$\bar{x} = \frac{40}{6} = 6,67$$

$$\bar{y} = \frac{36,5}{6} = 6,08$$

Varianza de X :

$$\sigma_x^2 = \frac{287,5}{6} - 6,67^2 = 3,43$$

• Covarianza:

$$\sigma_{xy} = \frac{263,25}{6} - 6,67 \cdot 6,08 = 3,32$$

• Coeficiente de regresión de Y sobre X :

$$m_{yx} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{3,32}{3,43} = 0,97$$

- Ecuación de la recta de regresión de Y sobre X :

$$y = 6,08 + 0,97(x - 6,67) \rightarrow y = 0,97x - 0,39$$

b) $\hat{y}(6) = 5,43$. Podemos considerar fiable la estimación, ya que $r = 0,98$ es alto y $x = 6$ está dentro del intervalo de valores que estamos considerando.

Ejercicio nº 6.-

Como el centro de gravedad (\bar{x}, \bar{y}) es el punto de intersección de las dos rectas de regresión, basta con resolver el sistema de ecuaciones:

$$\begin{aligned} y &= 8,7 - 0,76x \\ y &= 11,36 - 1,33x \end{aligned}$$

$$8,7 - 0,76x = 11,36 - 1,33x$$

$$\mathbf{X = 4.92}$$

$$\mathbf{Y = 4.96}$$