

ESTADÍSTICA – LIC. DOCUMENTACIÓN – Febrero 2007

Examen Tipo B

PROBLEMA 1: (4 puntos)

En una biblioteca se han contabilizado el número de libros consultados por día a lo largo de un año (365 días), información que se recoge en la siguiente tabla:

Nº de libros consultados	Nº de días	Fr. Absolutas acumuladas	Fr. Rrelativas	Fr. Relativas acumuladas
(0,30]			0,082	
(30,50]		71		
(50,90]	98			
(90,120]				
(100,180]	116			

1. Completar la tabla anterior. (0.75 puntos)
2. ¿Cuántos días se consultan más de 90 libros? ¿Y cuántos a lo sumo 50? (0.5 puntos)
3. Obtener las medidas de tendencia central. Interpretarlas. (1.5 puntos)
4. ¿Es una distribución simétrica? Interpretar. (0,5 puntos)
5. Supongamos que de otra biblioteca se conoce el número medio de libros consultados por día 63, y la desviación típica 12. ¿En cuál de las dos bibliotecas se presenta una mayor variabilidad? (0.75 puntos)

PROBLEMA 2: (6 puntos)

La siguiente tabla muestra para una colección de libros, el número de páginas (X), y el precio en euros de cada libro (Y):

X	Y
282	12
295	12.5
317	12.3
328	13
425	13.8

1. ¿Qué tipo de relación podría plantearse entre estas dos variables? Justificar la respuesta (no sirve explicarlo interpretando únicamente la tabla: utilizar métodos numéricos y gráficos) (2 puntos)
2. Obtener la recta de regresión. Interpretar (2 puntos)
3. Qué precio se espera que tenga un libro de 325 páginas? (1 punto)
4. Valorar el ajuste. (1 punto)

ESTADÍSTICA – LIC. DOCUMENTACIÓN – Febrero 2007

Examen Tipo A

PROBLEMA 1: (4 puntos)

En una biblioteca se han contabilizado el número de libros consultados por día a lo largo de un año (365 días), información que se recoge en la siguiente tabla:

Nº de libros consultados	Nº de días	Fr. Absolutas acumuladas	Fr. Rrelativas	Fr. Relativas acumuladas
(0,20]	26	26		
(20,50]		71		
(50,70]			0,279	
(70,100]				
(100,150]	112			

6. Completar la tabla anterior. (0.75 puntos)
7. ¿Cuántos días se consultan más de 70 libros? ¿Y cuántos a lo sumo 100? (0.5 puntos)
8. Obtener las medidas de tendencia central. Interpretarlas. (1.5 puntos)
9. ¿Es una distribución simétrica? Interpretar. (0,5 puntos)
10. Supongamos que de otra biblioteca se conoce el número medio de libros consultados por día 53, y la desviación típica 12. ¿En cuál de las dos bibliotecas se presenta una mayor variabilidad? (0.75 puntos)

PROBLEMA 2: (6 puntos)

La siguiente tabla muestra para una colección de libros, el número de páginas (X), y el precio en euros de cada libro (Y):

X	Y
282	12.2
293	12.5
312	12.3
325	13
405	13.5

5. ¿Qué tipo de relación podría plantearse entre estas dos variables? Justificar la respuesta (no sirve explicarlo interpretando únicamente la tabla: utilizar métodos numéricos y gráficos) (2 puntos)
6. Obtener la recta de regresión. Interpretar (2 puntos)
7. Qué precio se espera que tenga un libro de 350 páginas? (1 punto)
8. Valorar el ajuste. (1 punto)

Estadística – Lic. Documentación – Febrero 2007
Solución Examen Tipo B

PROBLEMA 1

1)

Nº de libros consultados	Nº de días	Fr. Abs. Acumuladas	Fr. Relativas	Fr. Relativas Acumuladas
(0,30]	30	30	0,082	0,082
(30,50]	41	71	0,11252055	0,19452055
(50,90]	98	169	0,26849315	0,4630137
(90,120]	80	249	0,21917808	0,68219178
(120,180]	116	365	0,31780822	1
	365		1	

Nota: Los datos en negrita eran los que daban en el enunciado.

2) Más de 90 libros: 80 + 116 = 196 días.

A lo sumo 50 libros: 30 + 41 = 71 días.

3)

Nº de libros consultados	Marca de Clase (x)	Nº de días (n)	nx	nx ²
(0,30]	15	30	450	6750
(30,50]	40	41	1640	65600
(50,90]	70	98	6860	480200
(90,120]	105	80	8400	882000
(120,180]	150	116	17400	2610000
Sumas		365	34750	4044550

Media: $\bar{x} = \frac{\sum n_i x_i}{N} = \frac{34750}{365} = 95,2054 \approx 95 \text{ días.}$

Mediana: Primero localizamos el intervalo de la mediana que es aquel en el cual la frecuencia absoluta acumulada supera por primera vez la mitad de los datos, en este problema el (90,240] y aplicamos la fórmula de la mediana:

$$Me = a_{i-1} + \frac{\left(\frac{N}{2} - N_{i-1}\right)}{n_i} (a_i - a_{i-1}) = 90 + \frac{\frac{365}{2} - 169}{80} (120 - 90) = 95,0625 \approx 95 \text{ días.}$$

Moda: Primero localizamos el intervalo modal, que es el que tiene mayor frecuencia absoluta, en este problema el (120,180]:

$$Mo = a_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} (a_i - a_{i-1}) =$$

$$= 120 + \frac{116 - 80}{(116 - 80) + (116 - 0)} (180 - 120) = 134,21 \square 134$$

- 4) Podemos decir que lo es ya que la media y la mediana coinciden.
 5) Cuanto mayor sea el coeficiente de variación de Pearson mayor será la variabilidad de los datos. Para calcularlo necesitamos conocer la desviación típica de los datos (s_x). Así pues empezaremos calculando la varianza de los datos:

$$s_x^2 = \frac{\sum n_i x_i^2}{N} - \bar{x}^2 = \frac{4044550}{365} - 95^2 = 2055,9589$$

Ahora podemos halla el coeficiente de variación de Pearson:

$$C.V. = \frac{s_x}{\bar{x}} = \frac{\sqrt{2055,9589}}{95} = 0.4772914$$

Para la segunda biblioteca tenemos:

$$C.V. = \frac{s_x}{\bar{x}} = \frac{12}{63} = 0.19047619$$

Por tanto la primera biblioteca presenta una mayor variabilidad.

PROBLEMA 2

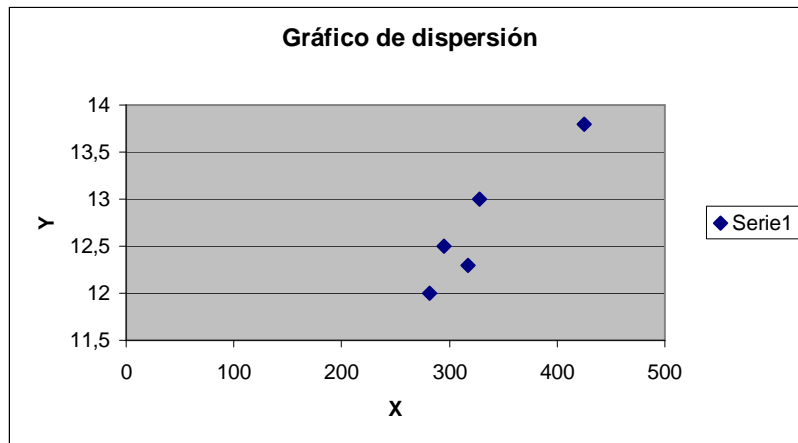
X	Y	X ²	Y ²	XY	
282	12	79524	144	3384	
295	12,5	87025	156,25	3687,5	
317	12,3	100489	151,29	3899,1	
328	13	107584	169	4264	
425	13,8	180625	190,44	5865	
Sumas	1647	63,6	555247	810,98	21099,6

- 1) Para justificar numéricamente la relación entre las variables hallaremos la covarianza:

$$s_{xy} = \frac{\sum x_i y_i}{N} - \bar{x} \bar{y} = \frac{21099,6}{5} - \frac{1647}{5} \cdot \frac{63,6}{5} = 29,952$$

Como la covarianza es distinta de cero podemos plantear una relación lineal.

Para justificar gráficamente el tipo de relación usamos un gráfico de dispersión:



2) La recta de regresión de y sobre x es de la forma:

$$y = a + bx$$

$$b = \frac{s_{xy}}{s_x^2} = \frac{29,952}{2545,04} = 0,01176877$$

$$a = \bar{y} - b\bar{x} = \frac{63,6}{5} - 0,01176877 \cdot \frac{1647}{5} = 8,84336592$$

$$s_x^2 = \frac{\sum x_i^2}{N} - \bar{x}^2 = \frac{555247}{5} - 329,4^2 = 2545,04$$

$$\text{esto es: } y = 8,84336592 + 0,01176877x.$$

3) El precio esperado será: $y = 8,84336592 + 0,01176877 \cdot 325 = 12,668 \approx 12,67\text{€}$

4) Para valorar el ajuste necesitamos el coeficiente de correlación lineal. Cuanto más próximo sea a 1, en valor absoluto, más fuerte será el ajuste:

$$r = \frac{s_{xy}}{s_x s_y} = \frac{29,952}{\sqrt{2545,04} \cdot \sqrt{0,3976}} = 0,9415759$$

$$s_y^2 = \frac{\sum y_i^2}{N} - \bar{y}^2 = \frac{810,98}{5} - 12,72^2 = 0,3976$$

Por tanto se trata de un ajuste fuerte y positivo.

Estadística – Lic. Documentación – Febrero 2007
Solución Examen Tipo A
PROBLEMA 1

1)

Nº de libros consultados	Nº de días	Fr. Abs. Acumuladas	Fr. Relativas	Fr. Relativas Acumuladas
(0,20]	26	26	0,071	0,071
(20,50]	45	71	0,124	0,195
(50,70]	102	173	0,279	0,474
(70,100]	80	253	0,219	0,693
(100,150]	112	365	0,307	1
	365		1	

Nota: Los datos en negrita eran los que daban en el enunciado.

2) Más de 70 libros: 80 + 112 = 192 días.

A lo sumo 100 libros: 26 + 45 + 102 + 80 = 253 días.

3)

Nº de libros consultados	Marca de Clase (x)	Nº de días (n)	nx	nx^2
(0,20]	10	26	260	2600
(20,50]	35	45	1575	55125
(50,70]	60	102	6120	367200
(70,100]	85	80	6800	578000
(100,150]	125	112	14000	1750000
Sumas		365	28755	2752925

Media: $\bar{x} = \frac{\sum n_i x_i}{N} = \frac{28755}{365} = 78,78 \approx 79$ días .

Mediana: Primero localizamos el intervalo de la mediana que es aquel en el cual la frecuencia absoluta acumulada supera por primera vez la mitad de los datos, en este problema el (70,100] y aplicamos la fórmula de la mediana:

$$Me = a_{i-1} + \frac{\left(\frac{N}{2} - N_{i-1}\right)}{n_i} (a_i - a_{i-1}) = 90 + \frac{\frac{365}{2} - 173}{80} (100 - 70) = 73,56 \approx 74 \text{ días.}$$

Moda: Primero localizamos el intervalo modal, que es el que tiene mayor frecuencia absoluta, en este problema el (70,100]:

$$Mo = a_{i-1} + \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} (a_i - a_{i-1}) =$$

$$= 100 + \frac{112 - 80}{(112 - 80) + (112 - 0)} (150 - 100) = 111,11 \square 111$$

- 4) Podemos decir que es asimétrica a la izquierda ya que la mediana es mas pequeña que la media.
- 5) Cuanto mayor sea el coeficiente de variación de Pearson mayor será la variabilidad de los datos. Para calcularlo necesitamos conocer la desviación típica de los datos (s_x). Así pues empezaremos calculando la varianza de los datos:

$$s_x^2 = \frac{\sum_i n_i x_i^2}{N} - \bar{x}^2 = \frac{2752925}{365} - 79^2 = 1301,26027$$

Ahora podemos halla el coeficiente de variación de Pearson:

$$C.V. = \frac{s_x}{\bar{x}} = \frac{\sqrt{1335,84237}}{79} = 0,37971564$$

Para la segunda biblioteca tenemos:

$$C.V. = \frac{s_x}{\bar{x}} = \frac{12}{53} = 0.226415094$$

Por tanto la primera biblioteca presenta una mayor variabilidad.

PROBLEMA 2

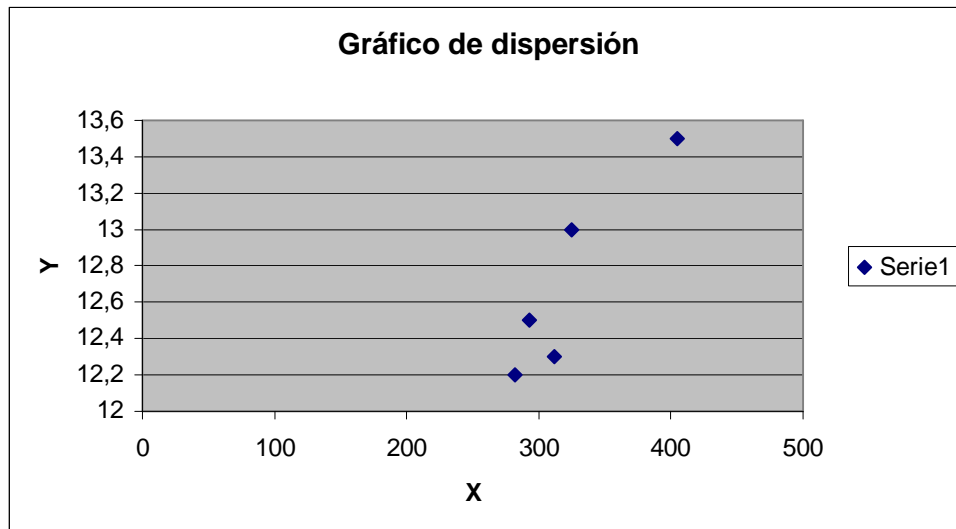
X	Y	X^2	Y^2	XY
282	12,2	79524	148,84	3440,4
293	12,5	85849	156,25	3662,5
312	12,3	97344	151,29	3837,6
325	13	105625	169	4225
405	13,5	164025	182,25	5467,5
Sumas	1617	532367	807,63	20633

- 1) Para justificar numéricamente la relación entre las variables hallaremos la covarianza:

$$s_{xy} = \frac{\sum x_i y_i}{N} - \bar{x} \bar{y} = \frac{20633}{5} - \frac{1617}{5} \cdot \frac{63,5}{5} = 19,42$$

Como la covarianza es distinta de cero podemos plantear una relación lineal.

Para justificar gráficamente el tipo de relación usamos un gráfico de dispersión:



2) La recta de regresión de y sobre x es de la forma:

$$y = a + bx$$

$$b = \frac{s_{xy}}{s_x^2} = \frac{19,42}{1885,84} = 0,0102978$$

$$a = \bar{y} - b\bar{x} = \frac{63,5}{5} - 0,0102978 \cdot \frac{1617}{5} = 9,36969202$$

$$s_x^2 = \frac{\sum x_i^2}{N} - \bar{x}^2 = \frac{532367}{5} - 323,4^2 = 1885,84$$

esto es: $y = 9,36939202 + 0,0102978x$.

3) El precio esperado será:

$$y = 9,36939202 + 0,0102978 \cdot 350 = 12,9739214 \approx 12,97\text{€}$$

4) Para valorar el ajuste necesitamos el coeficiente de correlación lineal. Cuanto más próximo sea a 1, en valor absoluto, más fuerte será el ajuste:

$$r = \frac{s_{xy}}{s_x s_y} = \frac{19,42}{\sqrt{1885,84} \cdot \sqrt{0,236}} = 0,92053605$$

$$s_y^2 = \frac{\sum y_i^2}{N} - \bar{y}^2 = \frac{807,63}{5} - 12,7^2 = 0,236$$

Por tanto se trata de un ajuste fuerte y positivo.