

**INTRODUCCIÓN A LA ESTADÍSTICA. (ADE). FEBRERO 2008. Segunda semana**  
(Código de la asignatura 202. Código de la carrera 42)

**PREGUNTAS TIPO TEST:**

1.- Sea la recta de regresión  $y = -80.2 + 1.01x$  de dicha recta conocemos además que  $r = 1$ . Con estos datos podemos afirmar que

- a) la correlación lineal es perfecta y negativa ☐ **b) la varianza residual es 0**  
c) la covarianza es 0 ☐ d) Ninguna de las respuestas es correcta

2.- A un grupo de estudiantes universitarios se le ha preguntado si va a participar en el viaje de "paso de Ecuador". Con las respuestas obtenidas ¿qué medidas podríamos calcular?

- a) Media aritmética ponderada ☐ **b) Moda** c) Mediana d) Ninguna de las respuestas es correcta

3.- De una distribución  $X_i$  se sabe que  $M_e = 8$  ( $M_e$  = mediana). Si multiplicamos cada uno de los valores de dicha distribución por una constante  $a$  ( $a > 0$ ) ¿cuál será el valor de la mediana en la nueva distribución que llamaremos  $Y_i$ ?

- a)  $M_e(Y) = M_e(X)$  ☐ **b)  $M_e(Y) = a M_e(X)$**  c)  $a M_e(Y) = M_e(X)$  d) Ninguna de las respuestas es correcta.

4.- De una muestra de 100 personas se observó que durante el año 2006, 30 adquirieron un coche, 75 compraron al menos un traje y 15 adquirieron un televisor.

Convenimos que al suceso "comprar un coche" lo denominamos A, "comprar un traje" B y "comprar un televisor" C. ¿Cuál es la probabilidad de que una persona compre un coche y por lo menos un traje?

- a)  $P(A \cup B)$  b)  $P(A \cap (B \cup C))$  c)  $P(A \cap C)$  ☐ **d) Ninguna de las respuestas es correcta**

5.- ¿Cuál de las siguientes opciones es correcta?

- a) Los valores observados de cualquier serie temporal son siempre el resultado de la adición de los componentes tendencia, variaciones cíclicas, variaciones estacionales y variaciones accidentales.  
b) Los valores observados de cualquier serie temporal son siempre el resultado de multiplicar la componente tendencial, cíclica, estacional y accidental  
c) Una serie temporal esta formada siempre por la tendencia, variaciones cíclicas, variaciones estacionales y variaciones estacionales.  
☐ **d) Ninguna de las respuestas es correcta**

6.- La varianza es igual a cero ( $S^2 = 0$ ) cuando

- a) la concentración es máxima  
☐ **b) la variable solo toma un valor**  
c) cuando  $\sum (X_i - \bar{X}) = 0$   $i = 1, \dots, r$   
d) Ninguna de las respuestas es correcta

7.- Sean dos distribuciones de las que conocemos la media y la varianza

Distribución 1:  $\bar{X} = 13$   $S^2 = 4$

Distribución 2:  $\bar{X} = 25$   $S^2 = 9$

¿Cuál de las dos medias es más representativa?

- a) la media aritmética de la distribución 1 es más representativa pues su varianza es menor.  
☐ **b) la media aritmética de la distribución 2 es más representativa.**  
c) depende de las unidades de medida de la distribución  
d) Ninguna de las respuestas es correcta

8.- Se dispone de la siguiente información sobre salarios mensuales de los trabajadores de una empresa ¿cuál de las respuestas propuestas es correcta?

Salarios mensuales	[800-2000]	(2000-2500]	(2500-2800]	(2800-3500]
Nº trabajadores	50	42	8	140

- a) el 87.62% de los trabajadores recibe el 81.65% de los salarios  
b) el 34.74% de los trabajadores recibe el 47.62% de los salarios  
c) La distribución de los salarios está muy concentrada y el índice correspondiente es próximo a 0.  
☐ **d) Ninguna de las respuestas es correcta**

9.- Un ahorrador compró un paquete de acciones el 1.01.03 cuyo nominal es de 20000 € el índice de cotización el 1.01.04 fue 82% el 1.01.05 70% el 1.01.06 cotizan al 110% y el 1.01.07 al 120% ¿cuál será la valoración de su cartera el 1.01.07 a precios corrientes?

- a) 24 000      b) 13 440      c) 15 153.6      d) Ninguna de las respuestas es correcta

10.- Un ahorrador compró un paquete de acciones el 1.01.03 cuyo nominal es de 20000 € el índice de cotización el 1.01.04 fue 82% el 1.01.05 70% el 1.01.06 cotizan al 110% y el 1.01.07 al 120% ¿cuál será la valoración de su cartera, a precios corrientes, durante estos años?

- a) 16 400    11 480    12 628    15 153      b) 16 400    14 000    22 000    24 000  
c) 24 690    34 843.2    31 675.6    26 396.4      d) Ninguna de las respuestas es correcta

11.- Dada la siguiente distribución de frecuencias con los datos agrupados en intervalos

$(L_{i-1}, L_i]$	$n_i$
20-30	14
30-40	4
40-50	9
50-60	3

la frecuencia relativa del segundo intervalo es:

- a) 0.599    b) 18      c) 0.133      d) Ninguna de las respuestas es correcta

12.- Se tiene la siguiente información sobre las variables  $X_1$   $X_2$   $Y_i$

$$[x'x]^{-1} = \begin{bmatrix} 0.56 & -0.68 & 0.43 \\ -0.68 & 1.43 & -1.06 \\ 0.43 & -1.06 & 0.81 \end{bmatrix} \quad x'y = \begin{bmatrix} 15 \\ 43 \\ 48 \end{bmatrix} \quad y'y = 55 \quad N = 5 \quad \bar{Y} = 2$$

¿cuál de las respuestas corresponde a la bondad del ajuste?

- a) 0.990      b) 0.985      c) 0.992      d) Ninguna de las respuestas es correcta

13.- Se tiene la siguiente información sobre las variables  $X_1$   $X_2$   $Y_i$

$$[x'x]^{-1} = \begin{bmatrix} 0.56 & -0.68 & 0.43 \\ -0.68 & 1.43 & -1.06 \\ 0.43 & -1.06 & 0.81 \end{bmatrix} \quad x'y = \begin{bmatrix} 15 \\ 43 \\ 48 \end{bmatrix} \quad y'y = 55 \quad N = 5 \quad \bar{Y} = 2$$

¿cuál de las respuestas corresponde a la varianza explicada por la regresión?

- a) 54.5      b) 10.9      c) 8.2      d) Ninguna de las respuestas es correcta

14.- Dada la siguiente distribución de salarios mensuales de una empresa ¿cuál de las respuestas propuestas corresponde al intervalo modal y a la moda?

Salarios mensuales	[800-2000]	(2000-2500]	(2500-2800]	(2800-3500]
Nº trabajadores	50	42	8	140

- a) [800-2000]  $M_o = 1834.48$       b) (2000-2500]  $M_o = 2195.65$   
c) [800-2000]  $M_o = 1268$       d) Ninguna de las respuestas es correcta

15.- ¿Cuál de estas afirmaciones es falsa?

- a) La varianza es la suma de las desviaciones de la variable respecto a la media aritmética  
b) La varianza es la media aritmética del cuadrado de la suma de las desviaciones de los valores de la variable con respecto a su media aritmética.  
c) La varianza nunca puede ser negativa  
d) Ninguna de las respuestas es falsa

### ALGUNAS ACLARACIONES

1.- Si  $r = 1 \rightarrow r^2 = 1$  es decir, toda la varianza está explicada por la regresión luego la varianza residual es cero.

4.- La respuesta correcta sería  $P(A \cap B)$

7.- Los respectivos coeficientes de variación son:  $CV_1 = \frac{2}{13} \cong 0,15$ ;  $CV_2 = \frac{3}{25} = 0,12$

**8.-** Para ver si es cierta la opción a, hallaremos la imagen de 87,62 en la curva de Lorentz. Para poder hacer los cálculos supondremos que los salarios se distribuyen uniformemente dentro de cada intervalo entre los trabajadores correspondientes, de forma que el volumen de salario percibido por todos los trabajadores de cada clase se obtiene multiplicando el número de trabajadres por la marca de clase del intervalo respectivo. Tenemos pues la tabla:

Marcas de clase $x_i$	$n_i$	$N_i$	$p_i$	$x_i \cdot n_i$	$u_i$	$q_i$
1400	50	50	20,83%	70000	70000	11,17%
2250	42	92	38,33%	94500	164500	26,25%
2650	8	100	41,67%	21200	185700	29,63%
3150	140	240		441000	626700	
			100,83%			67,05%

En la curva de Lorentz representada a la derecha las coordenadas de los puntos A, B y C son:

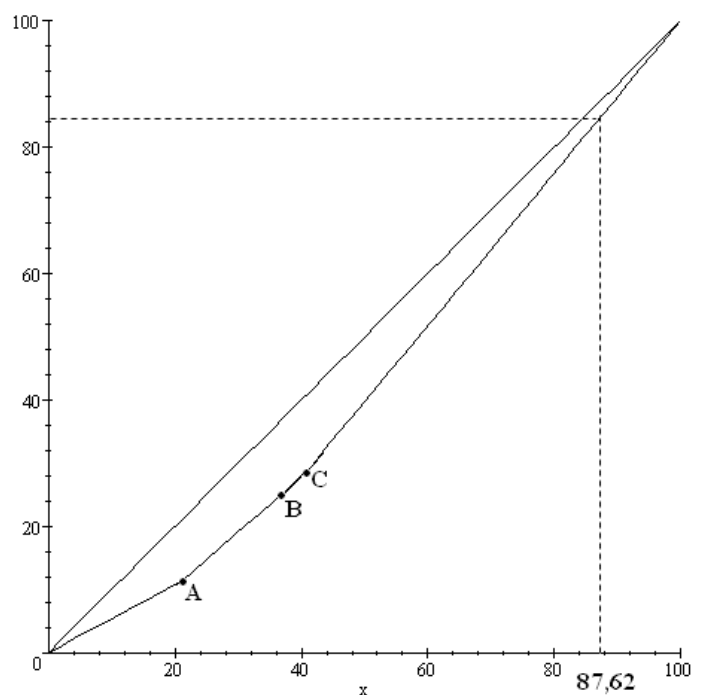
A(20,83; 11,17), B(38,33; 26,25) y C(41,67; 29,63).

Para hallar el porcentaje de salario que correspondería al 87,62% de trabajadores efectuaremos una interpolación lineal entre los puntos C y el punto de coordenadas (100, 100). Así pues, dicho porcentaje sería:

$$29,63 + \frac{100 - 29,63}{100 - 41,67}(87,62 - 41,67) = 85,06$$

que no coincide con el que se proporciona en el apartado a.

Desde luego tampoco es cierto el apartado b, pues los porcentajes de trabajadores no pueden ser inferiores a los de salarios.



Por otra parte, el índice de Gini sería  $\frac{100,83 - 67,05}{100,83} \cong 0,33$ , que no puede considerarse muy próximo a cero (en su escala de variación que es de 0 a 1).

**12.-** Véase la respuesta siguiente.

**13.-** La matriz  $b = [x'x]^{-1}x'y = \begin{pmatrix} 0,56 & -0,68 & 0,43 \\ -0,68 & 1,43 & -1,06 \\ 0,43 & -1,06 & 0,81 \end{pmatrix} \begin{pmatrix} 15 \\ 43 \\ 48 \end{pmatrix} = \begin{pmatrix} -0,2 \\ 0,41 \\ -0,25 \end{pmatrix}$ , de donde la

$$\text{varianza explicada por la regresión } S_{y,12}^2 = \frac{1}{N} [b'x'y - N\bar{Y}^2] = \frac{1}{5} \left[ (-0,2 \quad 0,41 \quad -0,25) \begin{pmatrix} 15 \\ 43 \\ 48 \end{pmatrix} - 20 \right] =$$

$= \frac{1}{5} (2,63 - 20) = -3,474$ . Este resultado es imposible porque la varianza debe ser positiva. La información sobre las variables  $X_1$ ,  $X_2$  e  $Y$  será, por lo tanto, errónea.

**14.-** La clase modal es la que tiene mayor densidad de frecuencia que se obtiene dividiendo la frecuencia por la amplitud de la clase, de forma que resultan las densidades:

Clases	Amplitud	Frecuencia	Densidad
[800, 2000]	1200	50	0,0417
(2000, 2500]	500	42	0,084
(2500, 2800]	300	8	0,027
(2800, 3500]	700	140	0,2

Es decir, la clase modal es (2800, 3500].

**15.-** La varianza de una variable  $X$  que toma los valores  $x_i$ , con frecuencias respectivas  $n_i$ ,  $i = 1, 2, \dots, n$ , es  $\frac{1}{N} \sum_{i=1}^n (x_i - \bar{X})^2 n_i$ . Es decir, es la media aritmética de los cuadrados de las

desviaciones de los valores de la variable con respecto a su media aritmética. Por tanto:

- No es la suma de las desviaciones de la variable respecto a la media aritmética.
  - No es la media aritmética del cuadrado de la suma de las desviaciones de los valores de la variable con respecto a su media aritmética.
  - Puesto que las afirmaciones a) y b) son falsas, entonces el apartado d) también es falso.
- El único apartado cierto es el c).