



**INTRODUCCIÓN A LA ESTADÍSTICA. (ADE). FEBRERO 2004 Examen tipo C**  
(Código de la asignatura 202. Código de la carrera 42)

**PREGUNTAS TIPO TEST**

1.- En una distribución simétrica se verifica que:

- a) La media coincide con la moda en todos los casos; b) El rango depende del número de observaciones; c) La mediana coincide con la moda en todos los casos; d) Ninguna de las respuestas es correcta.

**Respuesta.- d) Ninguna de las respuestas es correcta.**

2.- Dada la  $P(A) = 0,4$ , la  $P(B) = 0,25$  y además la probabilidad de la intersección es igual a 0,2, ¿cual es la probabilidad de  $P(A/B)$ ?:

- a) 0,8; b) 0,5; c) 0,32; d) Ninguna de las respuestas es correcta.

**Respuesta.- a) 0,8**

3.- Cual de las siguientes afirmaciones es verdadera:

- a) La media es un estadístico que no utiliza toda la información muestral; b) La mediana no se ve afectada por los valores extremos; c) La media no se ve afectada por los valores extremos; d) Ninguna de las respuestas es correcta.

**Respuesta.- b) La mediana no se ve afectada por los valores extremos**

4.- Si el coeficiente de curtosis de Fisher es mayor que 0:

- a) La distribución es platicúrtica; b) La distribución es mesocúrtica; c) La distribución es leptocúrtica; d) Ninguna de las respuestas es correcta.

**Respuesta.- c) La distribución es leptocúrtica**

5.- Cual de las siguientes variables es de tipo discreto:

- a) Tiempos de espera del Ave; b) Distancia entre las capitales de provincia; c) El numero de viviendas existentes en Madrid; d) Ninguna de las respuestas es correcta.

**Respuesta.- c) El numero de viviendas existentes en Madrid**

6.- Cuando el valor obtenido para la varianza residual en una regresión lineal es 0:

- a) Todos los valores teóricos coinciden con los observados. b) La correlación es perfecta; c) La respuestas a) y b) son correctas; d) Ninguna de las respuestas es correcta

**Respuesta.- c) Las respuestas a) y b) son correctas**

7.- Cuando el coeficiente de correlación lineal entre dos variables es 0 ( $r = 0$ ):

- a) La varianza residual y la varianza explicada por la regresión coinciden; b) La varianza explicada por la regresión es  $> 0$ ; c) La varianza residual y la varianza de la variable dependiente coinciden; d) Ninguna de las respuestas es correcta.

**Respuesta.- c) La varianza residual y la varianza de la variable dependiente coinciden**

8.- El conjunto de los posibles resultados que se pueden obtener cuando tiene lugar un experimento aleatorio se denomina:

- a) Fenómeno aleatorio; b) Experimento aleatorio; c) Espacio muestral; d) Ninguna de las respuestas es correcta.

**Respuesta.- c) Espacio muestral**

9.- Se ha realizado un cambio de escala en las variables  $x$  e  $y$  obteniendo las nuevas variables  $z = kx$  y  $c = ky$ . ¿Cuál sería el valor del coeficiente de correlación lineal de las



nuevas variables ( $r_{zc}$ )?: **a)**  $r_{zc} = \frac{k^2 S_{xy}}{k S_x k S_y}$ ; **b)**  $r_{zc} = \frac{k S_{xy}}{k S_x k S_y}$ ; **c)**  $r_{zc} = \frac{k^2 S_{xy}}{k^2 S_x k^2 S_y}$ ; **d)** Ninguna de las respuestas es correcta.

**Respuesta.- a)**  $r_x = \frac{k^2 S_{xy}}{k S_x k S_y}$

**10.-** ¿Qué consecuencias tiene para el ajuste mínimo cuadrático de un modelo lineal simple el hecho de que todos los datos de la variable y sean iguales?:

**a)** El origen de la ecuación es igual a 0; **b)** El origen de la ecuación es igual a 1; **c)** La pendiente de la ecuación es igual a 1; **d)** La pendiente de la ecuación es igual a 0.

**Respuesta.-d) La pendiente de la ecuación es igual a 0**

## EJERCICIOS PRACTICOS

**1.-** De la siguiente distribución bidimensional ( $x_i, y_i$ ), de frecuencias unitarias, tenemos los siguientes datos: los valores de x, (2, 3, 5, 7, 8), la recta de regresión de y sobre x ( $x + 2y - 1 = 0$ ) y el coeficiente de correlación lineal ( $r = 0,75$ ). Calcular: **a)** la media y la varianza de x, **b)** El coeficiente de regresión b de la recta de regresión de y sobre x. Comente su significado, **c)** La covarianza y **d)** La media y la varianza de y.

**Solución.-**

(Nota: hay una errata en el enunciado ya que la recta de regresión de Y/X:  $y = -\frac{1}{2}x + \frac{1}{2}$  tiene pendiente negativa y, por tanto, el coeficiente de correlación tendría que ser negativo. Para hacer el problema supondremos pues que  $r = -0,75$ )

**a)** media de la x =  $a_{10} = \frac{2+3+5+7+8}{5} = 5$ ; por otra parte  $a_{20} = \frac{4+9+25+49+64}{5} = 30,2$ , de donde varianza de la x =  $m_{20} = 30,2 - 25 = 5,2$

**b)** La recta de regresión es  $y = -\frac{1}{2}x + \frac{1}{2}$ . El coeficiente de regresión es la pendiente  $-\frac{1}{2}$ . Al ser negativo, la recta de regresión es decreciente y por tanto, al aumentar los valores de x, disminuirán los valores de y.

**c)** Puesto que  $b = \frac{m_{11}}{m_{20}}$ , la covarianza  $m_{11} = b \cdot m_{20} = -\frac{5,2}{2} = -2,6$ .

**d)** El término independiente de la recta de regresión es  $a_{01} + a_{10} \cdot b = \frac{1}{2}$ ; así pues: media de la y =  $a_{01} = -5 \cdot \left(-\frac{1}{2}\right) + \frac{1}{2} = 3$ ; para calcular la varianza de y  $m_{02}$ , puesto que  $r^2 = \frac{m_{11}^2}{m_{20} m_{02}} \rightarrow m_{02} = \frac{m_{11}^2}{r^2 m_{20}} = \frac{(-2,6)^2}{(-0,75)^2 \cdot 5,2} \cong 2,31$

**2.-** Dados los siguientes datos relativos al gasto en publicidad de una empresa (x) y su volumen de ventas (y):



$\sum_{i=1}^5 x_{1i} = 15$ ;  $\sum_{i=1}^5 x_{2i} = 14$ ;  $\sum_{i=1}^5 x_{1i}^2 = 51$ ;  $\sum_{i=1}^5 x_{2i}^2 = 54$ ;  $\sum_{i=1}^5 x_{1i}x_{2i} = 49$ ;  $\sum_{i=1}^5 y_i = 9$ ;  $\sum_{i=1}^5 x_{1i}y_i = 27$ ;  $\sum_{i=1}^5 x_{2i}y_i = 27$ . Obtenga el vector de coeficientes de regresión parcial e interprete los resultados.

**Solución.-**

De la matriz  $x^t x = \begin{pmatrix} 5 & 15 & 14 \\ 15 & 51 & 49 \\ 14 & 49 & 54 \end{pmatrix}$  obtenemos  $(x^t x)^{-1} = \frac{1}{199} \begin{pmatrix} 353 & -124 & 21 \\ -124 & 74 & -35 \\ 21 & -35 & 30 \end{pmatrix}$ ; además  $x^t y = \begin{pmatrix} 9 \\ 27 \\ 27 \end{pmatrix}$ , de donde  $b = (x^t x)^{-1} x^t y = \frac{9}{199} \begin{pmatrix} 44 \\ -7 \\ 6 \end{pmatrix} \cong \begin{pmatrix} 1,99 \\ -0,32 \\ 0,27 \end{pmatrix}$ . El plano de regresión sería

$y = 1,99 - 0,32x_1 + 0,27x_2$ , de forma que, a una variación de una unidad de  $x_1$ , permaneciendo constante  $x_2$ , corresponde una variación de  $-0,32$  unidades de  $y$ , y a una variación de una unidad de  $x_2$ , permaneciendo constante  $x_1$ , corresponde una variación de  $0,27$  unidades de  $y$ .

**3.-** Dada la siguiente distribución:

|       |      |       |       |       |
|-------|------|-------|-------|-------|
| $x_i$ | 1-10 | 11-20 | 21-30 | 31-40 |
| $n_i$ | 100  | 234   | 150   | 44    |

Calcular: a) La media, mediana y moda de la distribución.

**Solución.-**

Para que no existan discontinuidades en el recorrido de la variable, debemos precisar los extremos de los intervalos que, a falta de más datos, podremos tomar  $[1, 11[$ ,  $[11, 21[$ ,  $[21, 31[$  y  $[31, 41[$ . Así pues:

| Clases     | $x_i$ | $n_i$ | $x_i n_i$ | $N_i$ |
|------------|-------|-------|-----------|-------|
| $[1, 11[$  | 6     | 100   | 600       | 100   |
| $[11, 21[$ | 16    | 234   | 3744      | 334   |
| $[21, 31[$ | 26    | 150   | 3900      | 484   |
| $[31, 41[$ | 36    | 44    | 1584      | 528   |
|            |       | 528   | 9828      |       |

De donde: media =  $\frac{9828}{528} \cong 18,61$ ; mediana =  $L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} c_i = 11 + \frac{264 - 100}{234} 10 \cong 18$ ; moda =  $L_{i-1} + \frac{n_{i+1}}{n_{i-1} + n_{i+1}} c_i = 11 + \frac{150}{100 + 150} 10 = 17$