

OPOSICIONES AL CDEE. CONVOCATORIA 2019. EJERCICIO 1

ESTADÍSTICA TEÓRICA

1. Se consideran dos sucesos, A y B, asociados a un experimento aleatorio con $P(A)=1/4$, $P(B/A)=1/2$ y $P(A/B)=1/4$. Determine, justificando la respuesta, si los sucesos A y B son incompatibles.

2. Dada la función
$$f(x) = \begin{cases} x^a & \text{si } 0 \leq x \leq 1 \\ 0 & \text{resto} \end{cases}$$

Determine el valor de a para que $f(x)$ sea función de densidad.

3. Sea X una variable aleatoria que sigue una distribución t de Student con n grados de libertad. Determine, razonadamente, la distribución de la variable aleatoria X elevada al cuadrado.

4. Sea $(X_1, X_2, X_3, X_4, X_5)$ una muestra aleatoria simple de una variable aleatoria X con $E(X)=\mu - 5$ y $V(X)=\sigma^2$. Se proponen los siguientes estimadores para estimar el parámetro μ :

$$\hat{\mu}_1 = \sum_{i=1}^5 X_i \quad \hat{\mu}_2 = 8X_2 - 3X_5$$

Calcule el error cuadrático medio de cada uno de los estimadores. ¿Cuál es mejor y por qué?

5. La probabilidad de que una impresora tenga un fallo de funcionamiento cada día es p . Sabiendo que en 30 días la impresora ha fallado 5 días, estime, por el método de los momentos, el valor de la probabilidad p .

6. Una variable discreta X puede tomar los valores $\{1, 2, 3, 4, 5, 6\}$. Se plantea el contraste:

$$H_0: P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

$$H_1: P(1) = P(6) = \frac{2}{15}; P(2) = P(5) = \frac{1}{6}; P(3) = P(4) = \frac{1}{5}$$

Se dispone de una única observación x y se toma como región crítica $C=\{x / x=3 \text{ ó } 4\}$. Calcule las probabilidades de ambos tipos de error.

7. Calcule el tamaño de muestra necesario para estimar una proporción P en una población de 1000 individuos con muestreo aleatorio simple con reemplazamiento y un error relativo de muestreo del 5%. (Nota: en estudios anteriores esta proporción es el 50% de la población).

8. Un punto aleatorio X sigue una distribución uniforme en el intervalo $[0,1]$ y otro punto aleatorio Y tiene distribución uniforme en el intervalo $[2,3]$. Se define la variable aleatoria $Z=Y^2 - X^2$. Calcule la esperanza matemática de la variable Z .

9. Una muestra de una población normal de media desconocida μ y desviación típica $\sigma = 2$ ha dado los siguientes resultados: $n=100$ y $\sum x_i = 700$. Calcule de forma razonada un intervalo de confianza al 95% para la media poblacional μ . (Nota: para un valor $\alpha = 0.025$ el valor crítico necesario de la distribución del estadístico es 1,96)

10. Sea (X, Y) una variable bidimensional discreta con una función de probabilidad definida por la tabla siguiente:

X	Y		
	-2	0	3
-1	2/6	0	1/6
1	1/6	1/6	1/6

¿Son X e Y variables aleatorias independientes?

ESTADÍSTICA DESCRIPTIVA Y PROCESO ESTADÍSTICO

1. La empresa HOGARMAD pretende realizar un estudio estadístico sobre las viviendas de las distintas promociones de obra nueva disponibles en las afueras de Madrid. Para analizar el tamaño de la vivienda se encuentra con el problema de que cada promotora emplea una variable diferente, incluso de tipología. Defina la variable "tamaño de la vivienda" mediante una escala cualitativa ordinal y proponga razonadamente las medidas de síntesis que resuman esta distribución.

2. Los gastos medios mensuales en salarios de los trabajadores de una empresa ascienden a 128.000 euros. Sabiendo que el gasto medio para los trabajadores mujeres es de 120.000 euros mensuales y el de los hombres es 140.000 euros, determine el porcentaje de hombres y mujeres empleados en esta empresa.

3. A partir de la siguiente información relativa a la situación de una muestra de 100 hospitales, ¿cuál será el número mínimo de camas que deberá tener el hospital para poder beneficiarse de una ayuda especial, destinada únicamente al 30% de los hospitales con mayor capacidad?

Camas	0 - 100	100 - 200	200 - 300	300 - 400	400 - 500
Nº hospitales	11	25	32	10	22

4. Una empresa ha lanzado al mercado un nuevo producto. En la tabla adjunta se muestran los resultados de un estudio publicitario en el que se pide a diez personas potenciales clientes que valoren su grado de acuerdo (1=nada de acuerdo, 2=en desacuerdo, 3=de acuerdo, 4=totalmente de acuerdo) con las siguientes afirmaciones relacionadas con el producto:
 X_1 = "Compraría el producto"
 X_2 = "El producto tiene una buena relación calidad-precio"
 X_3 = "El producto tiene una buena presentación"
 X_4 = "La campaña de publicidad me ha hecho comprar el producto"

	X_1	X_2	X_3	X_4
1	2	2	1	1
3	4	3	4	4
4	3	1	4	4
3	4	3	4	4
2	3	2	2	2
3	4	3	1	1
4	4	4	4	4
1	2	2	1	1
3	4	2	3	3
4	2	1	2	2

Calcule, utilizando la distribución de frecuencias asociada, el porcentaje de potenciales clientes que, aun no estando de acuerdo con la presentación del producto ni con la campaña de publicidad, estarían dispuestos a comprar el producto.

5. Justifique razonadamente la veracidad o falsedad de la siguiente afirmación: "El ajuste lineal es el método de ajuste más utilizado porque es el que minimiza el error cometido".

6. La media aritmética y la varianza de una variable X son respectivamente 21 y 9. Sea la variable $Y = 4X - 8$. Compare la dispersión relativa respecto a la media de las dos variables.

7. Las rectas de regresión de las variables X e Y son:

$$\text{De } Y \text{ sobre } X: y = 1,7 + 0,8x$$

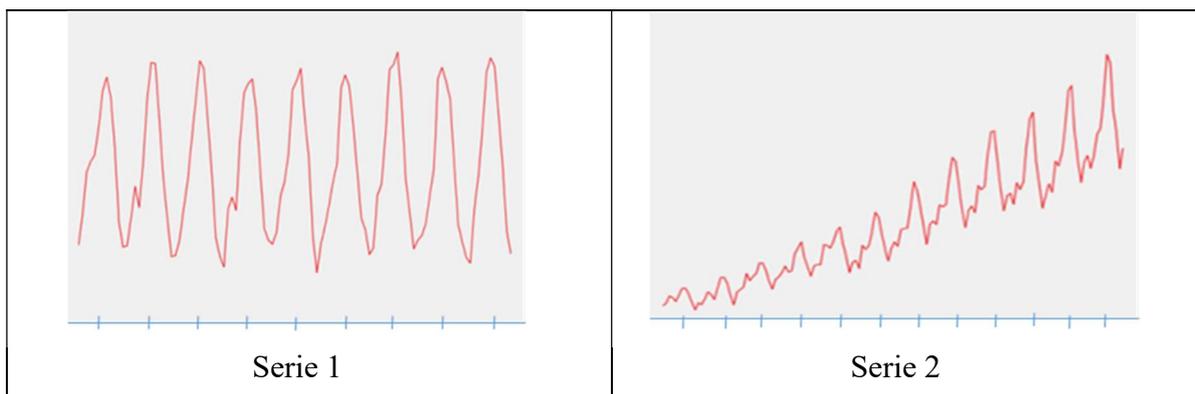
$$\text{De } X \text{ sobre } Y: y = 0,125 + 1,25x$$

Calcule el coeficiente de correlación y las medias aritméticas de cada variable.

8. Calcule el índice de concentración de Gini de la siguiente distribución salarial de la plantilla de una empresa:

- El 10 % de los empleados con menor sueldo supone el 2% del total de salarios pagados
- El 10% de empleados con mayor sueldo representa el 50% del total
- El 40% de empleados con sueldo inmediatamente superior al mínimo representa el 18% del total
- El 40% de empleados con sueldo inmediatamente inferior al máximo representa el 30% del total

9. Indique el tipo de tendencia, la componente estacional y el esquema de agregación que mejor se adapta a unas series temporales de datos trimestrales cuya representación gráfica es la que aparece en las figuras siguientes (los puntos del eje horizontal representan años):



10. En el análisis de los datos recibidos de una encuesta dirigida a empresas, se detecta una falta de respuesta significativa en la variable *Ingresos anuales*. Se pretende minimizar este problema utilizando una fuente externa, mediante un modelo que chequea condiciones de similitud basadas en variables comunes. Especifique la fase/subprocesos del GSBPM donde describirían estas actividades.