OPOSICIONES AL CDEE. CONVOCATORIA 2017. EJERCICIO 1

ESTADÍSTICA TEÓRICA BÁSICA

1. Una variable aleatoria X toma los valores -2, -1, 0, 1, 2. Se sabe que cada valor tiene la misma probabilidad que su opuesto, que P(X=0)=0,2 y que la probabilidad del valor 1 es el doble que la probabilidad del valor 2. Calcule la función de cuantía de X.

2. La longitud de una pieza se distribuye según una uniforme en el intervalo (1,3). Si sólo son válidas las piezas con longitud comprendida entre 1,2 y 2,8 se pide:

a) Probabilidad de que una pieza sea útil
b) Si las piezas se empaquetan en lotes de 5 unidades y se acepta el lote si contiene menos de 2 piezas defectuosas ¿Cuál es la probabilidad de que un cierto lote sea rechazado?

3. ¿Qué condiciones deben cumplirse para poder aproximar una distribución Binomial de parámetros n y p $\in (0,1)$ por una distribución Normal? Indique los parámetros de la Normal aproximada en relación con los parámetros de la Binomial.

4. Sean X e Y dos variables aleatorias con distribución conjunta normal bidimensional. Sabiendo que E(X)= E(Y), $E(X^2)=E(Y^2)$ y Cov(X,Y) = 2, razone si las afirmaciones siguientes son verdaderas o falsas:

a) X e Y son independientes.
b) Var(X) < 2

5. Dadas las variables aleatorias X_1 y X_2 idénticamente distribuidas, se construyen $Z=X_1 + X_2$ y $W=X_1 - X_2$. Demuestre que estas nuevas variables son incorrelacionadas.

6. ¿Cuál es la diferencia entre estimador y estimación? Ponga un ejemplo.

7. Sea (X_1, X_2) una muestra aleatoria simple de tamaño 2 de una variable aleatoria X cuya función de densidad viene dada por la expresión:

$$f(x) = ae^{-ax} \quad con \quad x \geq 0, \ a > 0$$

Calcule el estimador de máxima verosimilitud del parámetro a.

8. Sea $(X_1, X_2, \ldots, X_n)$ una muestra aleatoria simple de una característica X que se distribuye según una Normal de media μ y desviación típica σ_o conocida. Hallar el intervalo de confianza de longitud mínima para la media μ y nivel $1-\alpha$ con $\alpha \in (0,1)$ basado en el pivote $T(X_1, \ldots, X_n) = \frac{\bar{X}-\mu}{\sigma_o}\sqrt{n}$ con $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

9. Una fábrica elabora dos artículos A y B cuya demanda aleatoria sigue distribuciones normales independientes con medias μ_A y μ_B desconocidas y desviaciones típicas $\sigma_A = 100$ y $\sigma_B = 50$. Observados 100 puntos de venta, la demanda media de dichos artículos ha resultado de 200 y 150 unidades respectivamente. Obtenga una estimación insesgada para la diferencia de las demandas medias y calcule su error de muestreo. Razone las respuestas.

10. Sea X una variable aleatoria con media $\mu < 0$, se considera una muestra aleatoria simple para estimar μ y se construye el estimador $T(X_1, \ldots, X_n) = \bar{X}^2$ con $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

¿Es T insesgado para μ ? Razone la respuesta.

11. En una población se está realizando un estudio sobre determinada característica X con función de densidad:

$$f(x) = \theta x^{\theta-1} \qquad 0 < x < 1 \ \ y \ \ \theta > 0$$

Se toma una única observación y se pide contrastar $H_0 \equiv \theta \leq 1$ frente $H_1 \equiv \theta > 1$, haciendo uso de la región crítica $C = \{x/x \leq 1/2\}$. Halle la función de potencia y el tamaño del test.

12. Sea (X, Y) una variable aleatoria bidimensional con función de densidad:

$$f(x,y) = \begin{cases} k & 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad x + y \leq 1 \\ 0 & resto \end{cases}$$

Estudie la independencia de X e Y.

13. Al tirar 120 veces un dado se han obtenido los resultados siguientes:

Puntos	1	2	3	4	5	6
Frecuencia	20	14	23	12	26	25

Contraste la hipótesis de que el dado está equilibrado al nivel α = 0,05
(Nota: para un valor α = 0,05 el valor crítico necesario de la distribución del estadístico es 11,1)

14. Se dispone de una lista de 14 viviendas y, de ellas, las cuatro primeras son de protección oficial. El total de personas que habita cada una es: 2; 4; 3; 5; 2; 1; 6; 4; 3; 5; 2; 3; 4; 1. Calcule la varianza del estimador de la proporción de personas que habitan en viviendas protegidas, si se utiliza una muestra aleatoria simple sin reemplazamiento de 15 personas.

15. Se desea conocer el porcentaje de viviendas que tienen acceso a internet en una población formada por 1.000 viviendas, distribuidas en 20 conglomerados de 50 viviendas cada uno. Se realiza un muestreo bietápico con probabilidades iguales en ambas etapas, con reemplazamiento en la 1ª etapa y sin reemplazamiento en la 2ª etapa, seleccionando 4 conglomerados e investigando 10 viviendas en cada uno. Los resultados obtenidos son los siguientes:

Conglomerado	Total viviendas con acceso a internet
1	4
2	6
3	2
4	4

Estime el porcentaje de viviendas con acceso a internet.

ESTADÍSTICA DESCRIPTIVA Y PROCESO ESTADÍSTICO

1. Una institución pretende mejorar sus instalaciones y optimizar la superficie destinada al aparcamiento de sus empleados. Antes de acometer las obras, con el fin de conocer las necesidades del personal, pregunta a sus trabajadores cuál es el medio de transporte utilizado para llegar al trabajo. Defina la variable estadística adecuada para el objetivo del estudio, clasifíquela y proponga el gráfico más utilizado para su representación.

2. Los resultados de cumplimentación de cuestionarios de cinco unidades de recogida de un instituto de estadística se detallan a continuación:

	Cuestionarios cumplimentados	Horas dedicadas
Equipo 1	210	7
Equipo 2	240	6
Equipo 3	200	5
Equipo 4	315	9
Equipo 5	324	9

1) Calcule la productividad media de cada equipo, en cuestionarios por hora.

2) ¿Cuál cree que es la medida más adecuada para calcular la productividad media de los cinco equipos? Calcule ese promedio.

3. Sea X la variable que mide el número de niños que requieren atención psicopedagógica en un aula. Tras analizar las 47 aulas de un colegio, se obtiene la siguiente tabla de frecuencias absolutas y acumuladas:

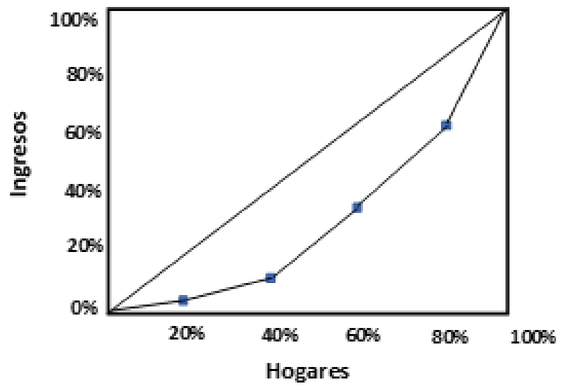
Nº niños que reciben atención psicopedagógica	Nº aulas (n_i)	N_i
1	7	
2		19
3		27
4		

Calcule el mayor número de niños por aula que reciben atención psicopedagógica en la mitad de las aulas con menor número de incidencias de este tipo.

4. Ventajas e inconvenientes de la utilización del recorrido intercuartílico como medida de dispersión para analizar el gasto turístico por comunidad autónoma realizado por los extranjeros que visitan España.

5. La media aritmética de una variable es 8 y su varianza es 4. ¿Qué transformación lineal debemos realizar para que la nueva variable tenga por media aritmética 42 y desviación típica 10?

6. Identifique razonadamente cuál de las siguientes afirmaciones se corresponde con la curva de Lorenz que se muestra a continuación:

<table>
<tr>
<td>
</td>
<td>a) El 40% de los hogares con menos ingresos gana menos del 20% del total de ingresos.
b) El 20% de los hogares con menos ingresos gana el 20% del total de ingresos.
c) Todos los hogares tienen más o menos los mismos ingresos.
d) Cuanto más se aleja la curva de la diagonal, mayor es la igualdad de ingresos.
e) Las respuestas a y d son correctas.</td>
</tr>
</table>

7. Dada una tabla de contingencia (X, Y) defina el concepto de independencia estadística entre X e Y. Si las variables son estadísticamente independientes ¿qué valor tiene su covarianza?

8. Si el coeficiente de correlación de Pearson es cero, entonces las variables objeto de análisis no están relacionadas. ¿Es correcta esta afirmación? Justifique la respuesta con una demostración o un contraejemplo, según el caso.

9. Defina el concepto de correlación parcial entre dos variables en el contexto de regresión lineal múltiple.

10. De la distribución de una variable bidimensional (X, Y) se conocen las siguientes características:

$$n = 200;\ \bar{y} = 75;\ \bar{x} = 24;\ \sum (x_i - \bar{x})^2 = 5.000;\ \sum (y_i - \bar{y})^2 = 20.000;\ r_{xy} = 0{,}9$$

Calcule la varianza de los valores de estimación de la variable Y obtenida por regresión lineal a partir de la variable X.

11. Dados los siguientes pares de rectas de regresión, justifique razonadamente si cada par puede pertenecer o no a la misma nube de puntos:

a) X/Y: X = 0,192Y - 0,76 Y/X: Y = 5,15X + 4,65

b) X/Y: X = -0,39Y + 12,09 Y/X: Y = 40,96 - 3,2X

c) X/Y: X = 1,33Y-1,882 Y/X: Y = 1,69 - 0,7X

12. Los coeficientes de variación trimestral de las ventas de un producto son: 0,9; 1,1; 1,0; 1,0. Si la tendencia crece a un ritmo del 1% trimestral, ordene los trimestres de menor a mayor número de ventas.

13. Defina el concepto de "base" en un número índice. ¿Qué base se toma como referencia para medir actualmente la evolución de los precios en España? ¿Cada cuántos años se cambia?

14. ¿Qué mide el Índice de Producción Industrial elaborado por el INE? Indique el tipo de índice utilizado y su periodicidad

15. Un estadístico analiza un fichero con datos sobre una muestra de empresas. Para ello comprueba que las variables se adecúan a los formatos estandarizados, identifica los outliers, detecta los casos con falta de respuesta y asigna valores según un algoritmo. ¿En qué fase y subproceso(s) del estándar GSBPM se clasifican estas tareas?