Oposición al Cuerpo Superior de Estadísticos de Estado

Primer Ejercicio

Convocatoria de la oferta pública de empleo de 2024

Resolución de 22 de diciembre de 2024, de la Subsecretaría, por la que se convocan procesos selectivos para ingreso, por el sistema de acceso libre y promoción interna, en el Cuerpo Superior de Estadísticos del Estado. (BOE 31 de Diciembre de 2024)

Acceso Promoción Interna

Producción Estadística Oficial: Principios Básicos del Ciclo de Producción de Operaciones Estadísticas

Pregunta 1. Defina el estimador de Horvitz-Thompson y demuestre formalmente que es insesgado. Además, derive su expresión específica en el caso de muestreo aleatorio simple sin reposición.

Pregunta 2. Se quiere estimar el total poblacional de una característica Y , para ello, consideramos dos estrategias: la estrategia e_1 compuesta por el diseño muestral de Bernoulli y el estimador de Horvitz-Thompson; y la estrategia e_2 , que utiliza el muestreo aleatorio simple sin reposición y el estimador de Horvitz-Thompson. Calcule el efecto de diseño de la estrategia e_1 comparada con la e_2 ¿Cuál de los dos diseños es más eficiente?

Pregunta 3. Describa de manera detallada el mecanismo de selección sistemática de muestra con probabilidades proporcionales al tamaño.

Pregunta 4. Determine las razones que pueden llevar a aplicar un muestreo por conglomerados en lugar de métodos de muestreo directo de elementos.

Pregunta 5. Exponga y formalice la clasificación de mecanismos de falta de respuesta según la categorización MCAR, MAR y NMAR.

Pregunta 6. Explique en qué consiste la fase 4 (Recoger/Obtener) del *Generic Statistical Business Process Model* (GSBPM) y qué subprocesos la conforman.

Pregunta 7. Describa los estándares SDMX y DDI, explicando sus propósitos y cómo se relacionan con el GSIM en la gestión de datos y metadatos estadísticos.

Pregunta 8. Enumere y describa tres indicadores relacionados con las dimensiones de la calidad en los productos estadísticos, especificando a qué dimensión pertenece cada uno.

Pregunta 9. Las unidades de una población de tamaño N están agrupadas de tal manera que la media poblacional en cada grupo, $\bar{Y}_r$, coincide. Si se desea estimar el total poblacional $X = X_1 + X_2 + \dots + X_N$ estando esta característica altamente correlada con la anterior, razone cuál de los siguientes dos planteamientos es mejor: un muestreo estratificado considerando cada grupo un estrato o un muestreo por conglomerados considerando cada grupo un conglomerado.

Pregunta 10. Compare los métodos de recogida de datos (autocumplimentación, entrevista telefónica y entrevista personal) en términos de coste, tiempo de recogida y tasa de respuesta.

Inferencia y Modelización Estadísticas

Pregunta 11. La gerente de un restaurante que sólo da servicio mediante reservas sabe, por experiencia, que el 20% de las personas que reservan una mesa no asistirán. Si el restaurante acepta 25 reservas pero sólo dispone de 20 mesas, ¿cómo calcularía la probabilidad de que a todas las personas que asistan al restaurante se les asigne una mesa? [No hace falta que efectúe el cálculo de la probabilidad. Basta con que dé la expresión que le llevaría a computar la misma].

Pregunta 12. Sea $X = \{x_1, x_2, \dots, x_n\}$ una muestra aleatoria simple de tamaño n . Se propone el siguiente estimador para la media μ :

$$\hat{\theta} = \frac{2x_1 + x_n}{3}$$

donde x_1, x_n son el primer y el último valor de la muestra. Estudie si se trata un estimador consistente e insesgado.

Pregunta 13. Se considera una población representada por una variable X , de suerte que la distribución de probabilidad de la población viene definida por la función de densidad:

$$f(x, \theta) = \begin{cases} \frac{2(\theta-x)}{\theta^2} & \text{si } 0 \leq x \leq \theta \\ 0 & \text{en otro caso} \end{cases}$$

Determinar el estimador, por el método de los momentos, del parámetro poblacional θ .

Pregunta 14. En una población $N(\mu, \sigma = 2)$ consideramos una hipótesis nula $H_0 : \mu = 1$ y una alternativa $H_1 : \mu \neq 1$. Se toma una muestra aleatoria de tamaño uno, y se considera como región crítica el intervalo $[2, \infty)$. Define con estos datos a qué nos referimos por error tipo I y por error tipo II, y calcula la probabilidad asociada al error tipo I.

Pregunta 15. Si en la población $N(\mu, \sigma = 2)$ consideramos las hipótesis $H_0 : [\mu = 6]$ y $H_1 : [\mu = 4]$, fijamos un nivel de significación de 0.05 y tenemos una muestra aleatoria simple de tamaño 4, entonces:

- establezca el contraste de hipótesis adecuado
- obtenga la región crítica
- calcule la potencia del contraste.

Pregunta 16. En relación con la Ley de las Esperanzas Iteradas, enuncie el contenido de dicha Ley y aplíquela, en su caso, para obtener el salario medio de las mujeres de una empresa de la que tenemos los siguientes datos:

- el salario medio de las mujeres españolas es de 40 euros/hora, y el de las mujeres con nacionalidad distinta a la española es de 30 euros/hora;
- la empresa está formada por 1000 trabajadores y trabajadoras, de los que 600 son mujeres, y 420 de ellas tienen nacionalidad española.

Pregunta 17. En el marco de los modelos lineales, explique cómo afecta a las denominadas «suma cuadrática de los residuos» y «suma cuadrática de la regresión» la inclusión de nuevas variables en el modelo.

Pregunta 18. En el marco del modelo lineal ordinario, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, con un error $\mathbb{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, indique las dimensiones de las matrices implicadas en la expresión anterior y compruebe si el estimador mínimo cuadrático del vector $\boldsymbol{\beta}$ es insesgado.

Almacenamiento y Modelos de Datos

Pregunta 19. ¿Qué es el código ASCII y cómo se diferencia del Unicode en la representación de caracteres?

Pregunta 20. ¿Cuáles son las principales diferencias entre los modelos de datos relacional, en red y jerárquico?