ACTIVIDAD 3: INTRODUCCIÓN A LA INFERENCIA
INFERENCIA ESTADÍSTICA

Problema 1 La Agencia para la Protección Ambiental está interesada en el problema de establecer normas para la cantidad permisible de ciertos productos químicos tóxicos en lagos y ríos de agua dulce. Una medida común de la toxicidad de cualquier contaminante es la concentración del contaminante que matará la mitad de especímenes de prueba en un intervalo de tiempo dado (normalmente 96 horas para peces). Esta medida de toxicidad se denomina CL50 del contaminante en cuestión. En muchos estudios se ha observado que el logaritmo neperiano del CL50 de un contaminante tiene una distribución normal. Estudios sobre la toxicidad del cobre para una especie de peces A revelan una varianza de $ln(CL50)$ de alrededor de 0.4 mg/l. Si se efectúan $n = 10$ mediciones de la toxicidad del cobre, encuentre la probabilidad de que la media de la muestra de $ln(CL50)$ difiera de la verdadera media de $ln(CL50)$ en no más de 0.5.

Problema 2 (Continuación del Problema 1). Se efectúan mediciones de toxicidad del cobre sobre otra especie de peces B, para la que otros estudios indican que la varianza de $ln(CL50)$ es de 0.8. Si se asume que las medias poblacionales de $ln(CL50)$ son iguales para ambas especies, encuentre la probabilidad de que con muestras aleatorias de 10 mediciones para cada especie, la media muestral para la especie A exceda a la media muestral de la especie B por lo menos en una unidad.

Problema 3 Sea $(X_1, \ldots, X_n)$ una m.a.s. de una $N(\mu, \sigma^2)$, con μ y σ^2 desconocidas. A la hora de estimar σ^2, se toman la varianza muestral S^2 y la cuasivarianza muestral S'^2.

(a) Calcula la esperanza de los dos estimadores. ¿Cuál es más cercana a σ^2?

(b) Calcula la varianza de ambos estimadores. ¿Cuál es menor?

Nota. Utiliza el Teorema de Fisher.

Problema 4 Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una m.a.s. de un vector aleatorio bidimensional (X, Y) con distribución normal bivariante, con vector de medias cero y matriz de covarianzas

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}.$$

Calcular la esperanza de la covarianza muestral

$$S_{XY} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}).$$

Problema 5 Los amperímetros producidos por una compañía en particular se venden en el mercado con la especificación de que la varianza de las lecturas no es mayor que 0.04

amp. Se utilizó uno de estos amperímetros para efectuar 10 lecturas independientes en un circuito de prueba con corriente constante. Si la varianza de estas mediciones es de 0.08, y es razonable suponer que las lecturas tienen distribución normal, ¿indican los resultados que el amperímetro que se utilizó no cumple la especificación, o han sido simplemente mediciones desafortunadas? Para contestar, calcule la probabilidad de que la varianza muestral exceda 0.08, si la varianza real es de 0.04.

Problema 6 Sea X una v.a. $U(0, \theta)$, y sea $(X_1, \ldots, X_n)$ una m.a.s. de X. Para estimar θ, se considera el estimador $\hat{\theta} = X_{(n)}$.

(a) Calcula la función de distribución de $\hat{\theta}$.

(b) Calcula la función de densidad de $\hat{\theta}$.

(c) Calcula la media y la varianza de $\hat{\theta}$.

Problema 7 Un antropólogo desea estimar la estatura promedio de los hombres de cierta raza. Si se supone que la desviación estándar de la población es de 2.5 pulgadas y se selecciona al azar a 100 hombres de la raza en cuestión:

(a) Encuentre la probabilidad de que la diferencia entre la estatura media de la muestra y la estatura media verdadera de la población no exceda de 0.5 pulgadas.

(b) Si el antropólogo desea que la diferencia entre la media muestral y la media poblacional sea menor de 0.4 pulgadas con probabilidad 0.95, ¿cuántos hombres tendría que seleccionar?

Problema 8 Se colocan 25 tubos de luz infrarroja en un invernadero, de tal menera que si falla un foco, automáticamente se enciende otro. Los focos tienen una vida media de 50 horas y una desviación estándar de 4 horas, siendo la vida de distintos focos independiente. Si no se inspecciona el invernadero durante 1300 horas despúes de encender el sistema de focos, ¿cuál es la probabilidad de que haya luz al final del periodo de tiempo de 1300 horas?

Problema 9 Sean $(X_1, \ldots, X_{n_1})$ e $(Y_1, \ldots, Y_{n_2})$ dos m.a.s. independientes de dos poblaciones con medias μ_1 y μ_2 y con varianzas σ_1^2 y σ_2^2 respectivamente. Si n_1 y n_2 son grandes,

(a) ¿Qué distribución aproximada tendrá $\bar{X} - \bar{Y}$?

(b) Si $\sigma_1^2 = 2$, $\sigma_2^2 = 2{,}5$ y $n_1 = n_2$, determinar el tamaño de ambas muestras para que $\bar{X} - \bar{Y}$ se aleje como máximo en 0.5 unidades de $\mu_1 - \mu_2$ con probabilidad 0.95.

Problema 10 Se procede a detener el funcionamiento de una máquina para repararla si en una muestra aleatoria de 100 artículos de la producción diaria de la máquina se encuentran por lo menos 15 artículos defectuosos. Si realmente la máquina produce sólo un 10 % de defectuosos, encuentre la probabilidad de que se pare la máquina un día dado.

Problema 11 (a) Sea $(X_1, \ldots, X_n)$ una m.a.s. de una v.a. con media μ_1 y varianza σ_1^2, y sea $(Y_1, \ldots, Y_n)$ otra m.a.s., independiente de la anterior, de una v.a. con media μ_2 y varianza σ_2^2. Calcular la distribución asintótica ($n \to \infty$), justificando el resultado, de la v.a.

$$U_n = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}.$$

(b) El flujo de agua a través de los suelos depende, entre otras cosas, de la porosidad (proporción del volumen de huecos) del suelo. Para la comparación de dos tipos de suelos arenosos, A y B, se obtienen 50 mediciones de la porosidad de cada tipo de suelo. Por estudios anteriores, se sabe que las varianzas de la porosidades en los tipos de suelo son $\sigma_A^2 = 0{,}01$ y $\sigma_B^2 = 0{,}02$ respectivamente. Calcular la probabilidad de que el valor de la diferencia entre las porosidades medias de las muestras diste como máximo 0.05 unidades de la diferencia entre las porosidades medias reales, $\mu_A - \mu_B$.

Problema 12 En un control sobre la abundancia relativa de una especie de pez en dos lagos distintos A y B, se ponen 50 trampas en cada lago, y en cada trampa se apunta si se recogen o no peces de la especie en cuestión. La experiencia previa ha mostrado que esta especie aparecerá en las trampas del lago A aproximadamente en un 10 % de la veces y en las del lago B en un 20 % de las veces. Aproximar la probabilidad de que la diferencia de proporciones de aparición de la especie en cada lago difiera como máximo en 0.1 de la diferencia de proporciones reales.

Problema 13 Sean $X_i \sim N(i\theta, 1)$, $i = 1, \ldots, n$, variables aleatorias independientes. Obtener la distribución del estimador de θ,

$$\hat{\theta} = \frac{\sum_{i=1}^n iX_i}{\sum_{i=1}^n i^2}.$$

Problema 14 Ciertos componentes electrónicos tienen tiempos de vida que se pueden modelizar mediante una distribución Weibull de parámetro β desconocido. La función de distribución de esta variable aleatoria es

$$F(x; \beta) = 1 - \exp(-x^{3/2}/\beta), \quad x > 0.$$

Si se toma una m.a.s. $(X_1, \ldots, X_n)$ de tiempos de vida de n de estas componentes, calcula la distribución asintótica (para n grande) de $\hat{\beta} = \sum_{i=1}^n X_i^{3/2}/n$.

Problema 15 (Exámen Junio 2006) (5 ptos.) A la hora de estimar los ingresos medios de los hogares, se toma una muestra aleatoria de n hogares. Se sabe que los ingresos totales de un hogar i, Y_i, están relacionados con el número de personas activas en dicho hogar, x_i. Concretamente, se sabe que $Y_i \sim N(\beta x_i, \sigma^2)$, $i = 1, \ldots, n$, donde x_i son constantes conocidas (no aleatorias), y β y σ^2 son parámetros desconocidos. Si se toman los siguientes estimadores

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{\beta}x_i)^2,$$

calcular su esperanza y su varianza, suponiendo que los ingresos de hogares distintos son independientes.

Nota. Se verifica $n\hat{\sigma}^2/\sigma^2 \sim \mathcal{X}_{n-1}^2$.