

GUIA #1 DE EJERCICIOS DE ESTADISTICA

MA-34B Prof. R. Gouet, 20/08/07

1. Se ha podido determinar que la duración de una componente electrónica es una v.a. con densidad de probabilidad $p_\theta(t) = \frac{1}{\theta}e^{-t/\theta}, t \geq 0, \theta > 0$, donde $\theta > 0$ es un parámetro desconocido que se quiere estimar, a partir de una muestra aleatoria simple (MAS) de tamaño n , dada por $T = (T_1, T_2, \dots, T_n)$.
 - (i) Calcule $E_\theta(T_1)$ y $V_\theta(T_1)$.
 - (ii) Muestre que el estimador máximo verosímil (EMV) de θ está dado por $\hat{\theta} = \bar{T} = \sum_{i=1}^n T_i$.
 - (iii) Muestre que $S(T) = \sum_{i=1}^n T_i$ es estadístico suficiente para θ .
 - (iv) Calcule la cota de Cramér-Rao para la varianza de los estimadores insesgados de θ y muestre que $\hat{\theta}$ es eficiente.

2. El número de fallas de una máquina, durante el intervalo de tiempo (expresado en horas) $[0, t]$ se puede modelar como una v.a. con función de probabilidad de Poisson, dada por

$$p_\theta(k) = \frac{e^{-\theta t}(\theta t)^k}{k!},$$

donde $k = 0, 1, 2, \dots, t > 0$ y $\theta > 0$ es un parámetro desconocido que se requiere estimar, a partir de una MAS $K = (K_1, K_2, \dots, K_n)$.

- (i) Encuentre el EMV de θ y calcule su varianza.
 - (ii) Calcule la cota de Cramér-Rao y decida si el EMV calculado en (i) es eficiente.
 - (iii) Se han efectuado 8 observaciones independientes de la máquina, en períodos de 24 horas, obteniéndose los resultados siguientes, respecto del número de fallas: 1,3,0,5,2,3,2,4. Calcule una estimación (numérica) para θ , a partir del EMV obtenido en (i).
3. Considere el experimento que consiste en hacer lanzamientos independientes de una moneda, con probabilidad de cara θ desconocida, hasta que hayan salido n caras. De este experimento se han registrado los largos $X_1, X_2, \dots, X_n$, eventualmente nulos, de las rachas de sellos entre las sucesivas caras.
 - (i) Muestre que las v.a. $X_1, X_2, \dots, X_n$ son independientes, idénticamente distribuidas con ley geométrica, dada por

$$p_\theta(k) = (1 - \theta)^k \theta, k \in \{0, 1, \dots\}, \theta \in [0, 1] = \Theta.$$

- (ii) Calcule el EMV de θ

4. Considere una MAS $X = (X_1, X_2, \dots, X_n)$ de una ley normal $N(\mu, \sigma_0^2)$, donde $\mu \in \mathbb{R}$ un parámetro desconocido pero $\sigma_0 > 0$ es conocido.
- (i) Muestre que $T(X) = \sum_{i=1}^n X_i$ es un estadístico suficiente para el parámetro μ .
 - (ii) Suponiendo que T es completo, deduzca que $\bar{X}$ es el estimador insesgado de varianza uniformemente mínima (EIVUM) para μ .
5. Sea $X = (X_1, X_2, \dots, X_n)$ una MAS de una ley normal $N(\mu, \sigma^2)$, donde $\mu \in \mathbb{R}$ y $\sigma > 0$ son parámetros desconocidos.
- (i) Muestre que $T(X) = (\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i)$ es un estadístico suficiente para el parámetro $\theta = (\mu, \sigma)$.
 - (ii) Suponiendo que T es completo, deduzca que $\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ es el EIVUM para σ^2 .
6. Considere una MAS $X = (X_1, X_2, \dots, X_n)$ de una ley normal $N(\mu, \sigma_0^2)$, donde $\mu \in \mathbb{R}$ es un parámetro desconocido pero $\sigma_0 > 0$ es conocido.
- (i) Muestre que $T(X) = \sum_{i=1}^n X_i$ es un estadístico suficiente para el parámetro μ .
 - (ii) Compruebe que $E_\mu(X_1^3) = \mu^3 + 3\sigma_0^2\mu$ (Indicación: Expresé X_1 como $\sigma_0 Y_1 + \mu$, con Y_1 normal $N(0, 1)$ y aproveche la simetría de la densidad de Y_1).
 - (iii) Muestre que $\tilde{g}_1(X) = \frac{1}{n} \sum_{i=1}^n X_i^3$ es estimador insesgado para $g(\mu) = \mu^3 + 3\sigma_0^2\mu$ y fuertemente consistente (converge casi seguramente a $g(\mu)$ cuando $n \rightarrow \infty$).
 - (iv) Suponiendo que T es completo, deduzca que

$$\hat{g}(X) = (\bar{X})^3 + 3\sigma_0^2\bar{X}(1 - 1/n)$$

es el EIVUM para $g(\mu)$. (Nota: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$).

7. Sean $\tilde{g}_1(X), \tilde{g}_2(X)$ estimadores insesgados de $g(\theta)$, donde $X = (X_1, X_2, \dots, X_n)$ proviene del modelo paramétrico $\mathcal{P} = \{P_\theta | \theta \in \Theta\}$.
- (i) Muestre que $\hat{g}_\alpha(X) = \alpha\tilde{g}_1(X) + (1 - \alpha)\tilde{g}_2(X)$ es un estimador insesgado para $g(\theta)$, cualquiera sea $\alpha \in [0, 1]$.
 - (ii) Sean σ_1^2, σ_2^2 las varianzas de $\tilde{g}_1(X), \tilde{g}_2(X)$ respectivamente y σ_{12} la covarianza entre $\tilde{g}_1(X)$ y $\tilde{g}_2(X)$. Suponiendo que σ_1^2, σ_2^2 y σ_{12} no dependen de θ , determine α óptimo de manera que $\hat{g}_\alpha(X)$ tenga mínima varianza.
8. Una caja contiene 5 nueces, θ de las cuales están frescas y las restantes rancias, donde θ es un entero (parámetro) desconocido con valores entre 0 y 5. Una persona extrae (al azar) dos nueces distintas de la caja, las prueba y constata que X están frescas, donde X es una va con valores en $\mathcal{X} = \{0, 1, 2\}$.
- (i) Plantee el modelo paramétrico $\{P_\theta | \theta \in \Theta\}$ para la situación descrita.
 - (ii) Construya el EMV de θ y averigüe si es insesgado. Notar que el EMV puede no ser único.

9. En la situación del ejercicio anterior (de las 5 nueces), la persona prueba al azar una nuez tras otra, hasta que le sale una rancia o se terminan las nueces. La observación X es el número de nueces probadas
- (i) Plantee el modelo paramétrico $\{P_\theta | \theta \in \Theta\}$ para la situación descrita.
 - (ii) Construya el EMV de θ y averigüe si es insesgado.
10. Dadas observaciones iid $X = (X_1, X_2, \dots, X_n)$ de un modelo de Poisson de parámetro λ , interesa estimar el tercer momento factorial $g(\lambda) = \lambda^3$.
- (i) Use el teorema de factorización para verificar que el estadístico $T(X) = \sum_{i=1}^n X_i$ es suficiente para λ .
 - (ii) Compruebe que $\tilde{g}(X) = X_1(X_1 - 1)(X_1 - 2)$ es estimador insesgado de $g(\lambda)$ y muestre que la ley condicional de X_1 dado $T(X) = m$ es una Binomial de parámetros $(m, 1/n)$.
 - (iii) Use el Teorema de Rao-Blackwell para obtener un estimador insesgado $\hat{g}(T(X))$, función del estadístico suficiente T , de menor varianza que $\tilde{g}$.
 - (iv) Apóyese en la Ley de Grandes Números para mostrar que $\hat{g}$ es consistente (converge casi seguramente a $g(\lambda)$ cuando $n \rightarrow \infty$).
- Nota: X_1 tiene ley de Poisson con parámetro λ si $P(X_1 = k) = e^{-\lambda} \lambda^k / k!$. La suma de va iid Poisson λ es Poisson $n\lambda$. Y tiene ley Binomial de parámetros (m, p) si $P(Y = k) = \binom{m}{k} p^k (1-p)^{m-k}$.
11. Sean $X_1, X_2, \dots, X_n$ puntos escogidos al azar (uniformes) e independientemente en el intervalo $[\theta_1, \theta_2]$, donde $\theta_1 < \theta_2$ son parámetros (reales) desconocidos.
- (i) Mostrar que $T(X) = (Y_n, Z_n)$ es estadístico suficiente para $\theta = (\theta_1, \theta_2)$, donde $Y_n = \min\{X_1, \dots, X_n\}$, $Z_n = \max\{X_1, \dots, X_n\}$.
 - (ii) Compruebe que el EMV de $\theta = (\theta_1, \theta_2)$ está dado por $\hat{\theta} = (Y_n, Z_n)$.
 - (iii) Considere el estimador de $g(\theta) = (\theta_1 + \theta_2)/2$, dado por $(Y_n + Z_n)/2$. Calcule su esperanza y compruebe si es insesgado.
- Considere las siguientes indicaciones para resolver (iii): 1) Si X es uniforme en $[0, \theta]$ entonces $\theta - X$ también es uniforme en $[0, \theta]$ (compruébelo). 2) Concluya de lo anterior que si las X_i son iid uniformes $[0, \theta]$ entonces las va $\max\{X_1, \dots, X_n\}$ y $\max\{\theta - X_1, \dots, \theta - X_n\} = \theta - \min\{X_1, \dots, X_n\}$ tienen igual distribución. 3) Deduzca el valor de la esperanza de $\min\{X_1, \dots, X_n\} + \max\{X_1, \dots, X_n\}$. 4) Compruebe que si X es uniforme en $[\theta_1, \theta_2]$ entonces $X - \theta_1$ es uniforme en $[0, \theta_2 - \theta_1]$. Ahora está en condiciones de calcular la esperanza de $(Y_n + Z_n)/2$ en (iii).
12. Se desea determinar lo más precisamente posible el área de un rectángulo de lados a, b desconocidos. Para ello se realizan n mediciones (con error) de cada lado: $X_1, \dots, X_n$ para el lado a e $Y_1, \dots, Y_n$ para el lado b . Se sabe que las observaciones están sujetas

a errores que pueden describirse como variables aleatorias normales independientes $N(0, \sigma^2)$, donde $\sigma > 0$ es conocido. Es decir, $X_i = a + \epsilon_i$, $Y_i = b + \delta_i$, $i = 1, \dots, n$, donde los ϵ_i, δ_i son va iid $N(0, \sigma^2)$.

- (i) Escriba la densidad conjunta de las observaciones $p(x, y, \theta)$, donde $\theta = (a, b)$.
- (ii) Muestre que $T(X, Y) = T(X_1, \dots, X_n, Y_1, \dots, Y_n) = (\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i)$ es estadístico suficiente para θ .
- (iii) Muestre que $\tilde{g}(X, Y) = X_1 Y_1$ es estimador insesgado del área $g(\theta) = ab$.
- (iv) Suponiendo que T es suficiente completo, muestre que $\bar{X} \cdot \bar{Y}$ es el EIVUM de $g(\theta)$.
- (v) Compare el EIVUM con el estimador razonable $\frac{1}{n} \sum_{i=1}^n X_i Y_i$. Para ello calcule las respectivas varianzas y compruebe que la varianza del EIVUM es menor.
- (vi) Suponga que el lado a se mide n veces y el lado b m veces ($m > n$). Estaría usted de acuerdo en ignorar las observaciones en "exceso" del lado b y aplicar el EIVUM encontrado arriba? Puede proponer usted algún estimador del área que use toda la información disponible de manera óptima?