Estadística inferencial

Oromion

19 de agosto del 2019

Contents

La *estadística inferencial*

Part I.

Teoría

Definición 1 (Estimador)**.** *Un estimador es una regla, a menudo expresada como una fórmula, que indica cómo calcular el valor de una estimación con base en las mediciones contenidas en una muestra.*

Ejemplo 1 (Media muestral)**.**

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

Definición 2 (Estimador insesgado e insesgado)**.** *Si $\hat{\theta}$ es un estimador puntual de un parámetro θ, entonces $\hat{\theta}$ es un* estimador insesgado *si* $\mathbb{E}\left(\hat{\theta}\right) = \theta$. *Si* $\mathbb{E}\left(\hat{\theta}\right) \neq \theta$, *se dice que $\hat{\theta}$ está* sesgado*.*

Definición 3 (Sesgo)**.** *El* sesgo *de un estimador puntual $\hat{\theta}$ está dado por* $B\left(\hat{\theta}\right) = \mathbb{E}\left(\hat{\theta}\right) - \theta$.

Definición 4 (Error cuadrático medio)**.** *El* error cuadrático medio *de un estimador puntual $\hat{\theta}$ es*

$$\mathrm{MSE}\left(\hat{\theta}\right) = \mathbb{E}\left[\left(\hat{\theta} - \theta\right)^2\right].$$

Proposición 1.

$$\mathrm{MSE}\left(\hat{\theta}\right) = V\left(\hat{\theta}\right) + \left[B\left(\hat{\theta}\right)\right]^2.$$

1. Ejercicios propuestos del capítulo *Estimación*

1. Usando la identidad

$$\left(\hat{\theta}-\theta\right)=\left[\hat{\theta}-\mathbb{E}\left(\hat{\theta}\right)-\theta\right]=\left[\hat{\theta}-\mathbb{E}\left(\hat{\theta}\right)\right]+B\left(\hat{\theta}\right)$$

   demuestre que

$$\mathrm{MSE}\left(\hat{\theta}\right)=\mathbb{E}\left[\left(\hat{\theta}-\theta\right)^{2}2\right]=V\left(\hat{\theta}\right)+B\left(\hat{\theta}\right)^{2}.$$

2. a) Si $\hat{\theta}$ es un estimador insesgado para θ, ¿cuál es $B\left(\hat{\theta}\right)$?

   b) Si $B\left(\hat{\theta}\right)=5$, ¿cuál es $\mathbb{E}\left(\hat{\theta}\right)$?

3. Suponga que $\hat{\theta}$ es un estimador para un parámetro θ y $\mathbb{E}\left(\hat{\theta}\right)=a\theta+b$ para algunas constantes diferentes de cero a y b.

   a) En términos de a, b y θ, ¿cuál es $B\left(\hat{\theta}\right)$?

   b) Encuentre una función de $\hat{\theta}$, por ejemplo $\hat{\theta}^*$, que es un estimador insesgado para θ.

4. Del ejercicio (1)

   a) Si $\hat{\theta}$ es un estimador insesgado para θ, ¿cómo se compara $\mathrm{MSE}\left(\hat{\theta}\right)$ con $V\left(\hat{\theta}\right)$?

   b) Si $\hat{\theta}$ es un estimador insesgado para θ, ¿cómo se compara $\mathrm{MSE}\left(\hat{\theta}\right)$ con $V\left(\hat{\theta}\right)$?

5. Del ejercicio (1), considere el estimador insesgado $\hat{\theta}^*$ que usted propuso en el tercer ejercicio.

   a) Exprese $\mathrm{MSE}\left(\hat{\theta}^*\right)$ como función de $V\left(\hat{\theta}\right)$.

   b) Dé un ejemplo de un valor para a para el cual $\mathrm{MSE}\left(\hat{\theta}^*\right)<\mathrm{MSE}\left(\hat{\theta}\right)$.

   c) Dé un ejemplo de un valor para a para el cual $\mathrm{MSE}\left(\hat{\theta}^*\right)>\mathrm{MSE}\left(\hat{\theta}\right)$.

6. Suponga que $\mathbb{E}\left(\hat{\theta}_1\right) = \mathbb{E}\left(\hat{\theta}_2\right) = \theta$, $V\left(\hat{\theta}_1\right) = \sigma_1^2$ y $V\left(\hat{\theta}_2\right) = \sigma_2^2$. Considere el estimador $\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\,\hat{\theta}_2$.
   a) Demuestre que $\hat{\theta}_3$ es un estimador insesgado para θ.
   b) Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son independientes, ¿cómo debe escogerse la constante a para minimizar la varianza de $\hat{\theta}_3$?

7. Considere la situación descrita en el ejercicio (6). ¿Cómo debe elegirse la constante a para minimizar la varianza de $\hat{\theta}_3$, si $\hat{\theta}_1$ y θ_2 no son independientes pero son tales que $\operatorname{Cov}\left(\hat{\theta}_1, \hat{\theta}_2\right) = c \neq 0$?

8. Suponga que Y_1, Y_2, Y_3 denotan una muestra aleatoria de una distribución exponencial con función de densidad

$$f(y) = \begin{cases} \left(\frac{1}{\theta} e^{-y/\theta}\right), & y > 0 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

   Considere los siguientes cinco estimadores de θ:

$$\hat{\theta}_1, \quad \hat{\theta}_2, \quad \hat{\theta}_3 = \frac{Y_1 + 2Y_2}{3}, \quad \theta_4 = \min(Y_1, Y_2, Y_3), \quad \theta_5 = \overline{Y}.$$

   a) ¿Cuáles de estos estimadores son insesgados?
   b) Entre los estimadores insesgados, ¿cuál tiene la varianza más pequeña?

9. Suponga que $Y_1, Y_2, \ldots, Y_n$ constituyen una muestra aleatoria de una población con función de densidad de probabilidad

$$f(y) = \begin{cases} \left(\frac{1}{\theta+1} e^{-y/(\theta+1)}\right), & y > 0, \theta > -1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

   Sugiera un estadístico apropiado para usarlo como estimador insesgado para θ.

10. El número de descomposturas por semana para un tipo de minicomputadora es una variable aleatoria Y con una distribución de Poisson y media λ. Existe una muestra aleatoria $Y_1, Y_2, \ldots, Y_n$ de observaciones del número semanal de descomposturas.
   a) Sugiera un estimador insesgado para λ.
   b) El costo semanal de reparar estas descomposturas es $C = 3Y + Y^2$. Demuestre que $E(C) = 4\lambda + \lambda^2$.
   c) Encuentre una función de $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño n de una población con media 3. Suponga que $\hat{\theta}_2$ es un estimador insesgado para el tercer momento central de la distribución subyacente.

11. La lectura en un voltímetro conectado a un circuito de prueba está distribuida uniformemente en el intervalo $(\theta, \theta + 1)$, donde θ es el valor desconocido del voltaje real del circuito. Suponga que $Y_1, Y_2, \ldots, Y_n$ denota una muestra aleatoria de esas lecturas.
    a) Demuestre que $\overline{Y}$ es un estimador sesgado de θ y calcule el sesgo.
    b) Encuentre una función de $\overline{Y}$ que sea un estimador insesgado de θ.
    c) Encuentre MSE $\left(\overline{Y}\right)$ cuando $\overline{Y}$ se use como estimador de θ.
12. Hemos visto que si Y tiene una distribución binomial con parámetros n y p, entonces Y/n es un estimador insesgado de p. Para calcular la varianza de Y, por lo general usamos $n\,(Y/n)\,(1 - Y/n)$.
    a) Demuestre que el estimador sugerido es un estimador sesgado de $V\,(Y)$.
    b) Modifique ligeramente $n\,(Y/n)\,(1 - Y/n)$ para formar un estimador insesgado de $V\,(Y)$.
13. Sea $Y_1, Y_2 \ldots, Y_n$ una muestra aleatoria de tamaño n de una población cuya densidad está dada por

$$f\,(y) = \begin{cases} \alpha y^{\alpha-1}/\theta^{\alpha}, & 0 \leq y \leq \theta, \\ 0, & \text{en cualquier otro punto,} \end{cases}$$

    donde $\alpha > 0$ es un valor fijo conocido, pero θ no se conoce. Considere el estimador $\hat{\theta} = \max\,(Y_1, Y_2, \ldots, Y_n)$.
    a) Demuestre que $\hat{\theta}$ es un estimador sesgado de $\hat{\theta}$.
    b) Determine un múltiplo de $\hat{\theta}$ que constituya un estimador de $\hat{\theta}$.
    c) Deduzca MSE $\left(\hat{\theta}\right)$.
14. Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño n de una población cuya densidad está dada por

$$f\,(y) = \begin{cases} 3\beta^3 y^{-4}, & \beta \leq y, \\ 0, & \text{en cualquier otro punto,} \end{cases}$$

    donde $\beta > 0$ es desconocido. Considere el estimador $\hat{\beta} = \min\,(Y_1, Y_2, \ldots, Y_n)$.
    a) Deduzca el sesgo del estimador $\hat{\beta}$.
    b) Deduzca MSE $\left(\hat{\beta}\right)$.
15. Suponga que $Y_1, Y_2, \ldots, Y_n$ constituyen una muestra aleatoria de una distribución normal con parámetros μ y σ^2.
    a) Demuestre que $S = \sqrt{S^2}$ es un estimador sesgado de σ.
    b) Ajuste S para formar un estimador insesgado de σ.

c) Encuentre un estimador insesgado de $\mu - z_\alpha \sigma$, el punto que corta un área de cola inferior de α bajo esta curva normal.

16. Si Y tiene otra distribución binomial con parámetros n y p. entonces $\hat{p}_1 = Y/n$ es un estimador insesgado de p. Otro estimador de p es $\hat{p}_2 = (Y+1)/(n+2)$.

 a) Deduzca el sesgo de $\hat{p}_2$.

 b) Deduzca $\mathrm{MSE}(\hat{p}_1)$ y $\mathrm{MSE}(\hat{p}_2)$.

 c) ¿Para qué valores de p es $\mathrm{MSE}(\hat{p}_1) < \mathrm{MSE}(\hat{p}_2)$?

17. Si $Y_1, Y_2, \ldots, Y_n$ representan una muestra aleatoria de tamaño n de una población con una distribución uniforme en el intervalo $(0, \theta)$, considere a $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ el estadístico de orden más bajo. Deduzca $\mathbb{E}\left(Y_{(1)}\right)$. Encuentre un múltiplo de $Y_{(1)}$ que sea un estimador insesgado para θ.

18. Suponga que $Y_1, Y_2, \ldots, Y_n$ denotan una muestra aleatoria de tamaño n de una población con una distribución exponencial cuya densidad está dada por

$$f(y) = \begin{cases} (1/\theta)\, e^{-y/\theta}, & y > 0, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Si $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ denota el estadístico de orden más bajo, demuestre que $\hat{\theta} = nY_{(1)}$ es un estimador insesgado para θ y encuentre $\mathrm{MSE}\left(\hat{\theta}\right)$.

19. Suponga que Y_1, Y_2, Y_3, Y_4 denotan una muestra aleatoria de tamaño 4 de una población con una distribución exponencial cuya densidad está dada por

$$f(y) = \begin{cases} (1/\theta)\, e^{-y/\theta}, & y > 0, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

 a) Sea $X = \sqrt{Y_1 Y_2}$. Encuentre un múltiplo de X que sea un estimador insesgado para θ.

 b) Sea $W = \sqrt{Y_1, Y_2, Y_3, Y_4}$. Encuentre un múltiplo de W que sea un estimador insesgado para θ^2.

Part II.

Práctica

2. Marco de datos en R