



Facultad de Ciencias Económicas y Empresariales
UNIVERSIDAD AUTÓNOMA DE MADRID

Estimadores Puntuales





ESTIMADORES PUNTUALES : INSESGADEZ. EFICIENCIA



La variable aleatoria poblacional "renta de las familias" del municipio de Madrid se distribuye siguiendo un modelo $N(\mu, \sigma)$. Se extraen muestras aleatorias simples de tamaño 4. Como estimadores del parámetro μ , se proponen los siguientes:

$$\hat{\mu}_1 = \frac{x_1 + 2x_2 + 3x_3}{6} \quad \hat{\mu}_2 = \frac{x_3 - 4x_2}{-3} \quad \hat{\mu}_3 = \bar{x}$$

Se pide:

- Comprobar si los estimadores son insesgados
- ¿Cuál es el más eficiente?
- Si tuviera que escoger entre ellos, ¿cuál escogería?. Razone su respuesta a partir del Error Cuadrático Medio.

Solución:

a) Un estimador $\hat{\theta}$ es insesgado (o centrado) cuando se verifica $E(\hat{\theta}) = \theta$

$$E(\hat{\mu}_1) = E\left[\frac{x_1 + 2x_2 + 3x_3}{6}\right] = \frac{1}{6} E[x_1 + 2x_2 + 3x_3] = \frac{1}{6} [E(x_1) + 2E(x_2) + 3E(x_3)] = \frac{1}{6} [6\mu] = \mu$$

$$E(\hat{\mu}_2) = E\left[\frac{x_3 - 4x_2}{-3}\right] = -\frac{1}{3} E[x_3 - 4x_2] = -\frac{1}{3} [E(x_3) - 4E(x_2)] = -\frac{1}{3} [-3\mu] = \mu$$

$$E(\hat{\mu}_3) = E\left[\frac{x_1 + x_2 + x_3 + x_4}{4}\right] = \frac{1}{4} E[x_1 + x_2 + x_3 + x_4] = \frac{1}{4} [E(x_1) + E(x_2) + E(x_3) + E(x_4)] = \frac{1}{4} [4\mu] = \mu$$

Los tres estimadores son insesgados o centrados.

b) El estimador *más eficiente* es el que tenga menor varianza.

$$V[\hat{\mu}_1] = V\left[\frac{x_1 + 2x_2 + 3x_3}{6}\right] = \frac{1}{36} V[x_1 + 2x_2 + 3x_3] = \frac{1}{36} [V(x_1) + 4V(x_2) + 9V(x_3)] = \frac{1}{36} [14\sigma^2] = \frac{14}{36}\sigma^2 = 0,39\sigma^2$$

$$V[\hat{\mu}_2] = V\left[\frac{x_3 - 4x_2}{-3}\right] = \frac{1}{9} V[x_3 - 4x_2] = \frac{1}{9} [V(x_3) + 16V(x_2)] = \frac{1}{9} [17\sigma^2] = \frac{17}{9}\sigma^2 = 1,89\sigma^2$$



$$\begin{aligned}
 V[\hat{\mu}_3] &= V\left[\frac{x_1 + x_2 + x_3 + x_4}{4}\right] = \frac{1}{16} V[x_1 + x_2 + x_3 + x_4] = \\
 &= \frac{1}{16} [V(x_1) + V(x_2) + V(x_3) + V(x_4)] = \frac{1}{16} [4\sigma^2] = \frac{4}{16}\sigma^2 = 0,25\sigma^2
 \end{aligned}$$

El estimador $\hat{\mu}_3$ es el más eficiente.

$$c) \text{ECM}(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + [b(\hat{\theta})]^2 \quad \text{sesgo } b(\hat{\theta}) = [E(\hat{\theta}) - \theta]$$

Al ser los tres estimadores insesgados $b(\hat{\theta}) = 0 \rightarrow \text{ECM}(\hat{\theta}) = V(\hat{\theta})$

El error cuadrático medio coincide con la varianza, por tanto, se elige el estimador $\hat{\mu}_3$ por presentar menor varianza.



ESTIMADORES PUNTUALES: INSESGADEZ. EFICIENCIA



El peso en kilos de los jamones vendidos por una empresa sigue una distribución normal con varianza 4 y peso medio desconocido. Se conoce que el peso medio de los jamones vendidos es superior a 5 kg, y se toman m.a.s. de tamaño 4 para estimar θ .

¿Cuál de los dos estimadores sería el mejor respondiendo a la inexactitud y eficiencia?

$$\hat{\theta}_1 = \frac{x_1 + x_2 + x_3}{4} \quad \hat{\theta}_2 = \frac{x_1 + x_2}{2}$$

Solución:

La v.a X_i = 'peso en kg de los jamones' sigue una distribución normal de varianza $\sigma^2 = 4$

Para estudiar la inexactitud de los estimadores se calculan sus esperanzas:

$$E(\hat{\theta}_1) = E\left[\frac{x_1 + x_2 + x_3}{4}\right] = \frac{1}{4} [E(x_1) + E(x_2) + E(x_3)] = \frac{3}{4} \theta$$

El sesgo del estimador $\hat{\theta}_1$ será: $b(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta = \frac{3}{4} \theta - \theta = -\frac{1}{4} \theta$

$$E(\hat{\theta}_2) = E\left[\frac{x_1 + x_2}{2}\right] = \frac{1}{2} [E(x_1) + E(x_2)] = \frac{2}{2} \theta = \theta$$

El estimador $\hat{\theta}_2$ es inexacto, $b(\hat{\theta}_2) = 0$

Atendiendo al sesgo se elige $\hat{\theta}_2$

Para analizar la eficiencia relativa de los dos estimadores se calculan las respectivas varianzas

$$V(\hat{\theta}_1) = V\left[\frac{x_1 + x_2 + x_3}{4}\right] = \frac{1}{16} \underbrace{[V(x_1 + x_2 + x_3)]}_{\substack{\text{Observaciones} \\ \text{independientes}}} = \frac{1}{16} [V(x_1) + V(x_2) + V(x_3)]$$

$$= \frac{1}{16} \cdot 12 = \frac{12}{16} = \frac{3}{4}$$

$V(x_i) = 4$

$$V(\hat{\theta}_2) = V\left[\frac{x_1 + x_2}{2}\right] = \frac{1}{4} [V(x_1 + x_2)] = \frac{1}{4} [V(x_1) + V(x_2)] = \frac{1}{4} \cdot 8 = 2$$

Respecto a la varianza se elige el estimador $\hat{\theta}_1$ por ser el de menor varianza.

Aparecen decisiones contrapuestas, de modo que el estimador se elige en base al error cuadrático medio: $ECM = \text{Varianza} + (\text{sesgo})^2$

$$ECM(\hat{\theta}_1) = \frac{3}{4} + \left(-\frac{\theta}{4}\right)^2 = \frac{\theta^2 + 12}{16} \quad ECM(\hat{\theta}_2) = 2 + 0 = 2$$

Se analiza cuando es mayor el ECM en el primer estimador $\hat{\theta}_1$:



$$\text{ECM}(\hat{\theta}_1) > \text{ECM}(\hat{\theta}_2) \rightarrow \frac{\theta^2 + 12}{16} > 2 \rightarrow \theta^2 > 20 \Rightarrow |\theta| > \sqrt{20} \approx 4,47$$

Si θ es en valor absoluto mayor que 4,47, el error cuadrático medio de $\hat{\theta}_1$ es mayor, con lo que se elige el estimador $\hat{\theta}_2$.

Se conoce que el peso medio de los jamones es superior a 5 kg, no queda duda que el estimador a elegir (con menor error cuadrático medio) es $\hat{\theta}_2$



ESTIMADORES PUNTUALES: EFICIENCIA. CONSISTENCIA



La distribución de ingresos de cierta población es una variable aleatoria con media μ desconocida y varianza σ^2 también desconocida. Se desea estimar el ingreso medio de la población mediante una m.a.s. de tamaño n , respecto de la insesgadez y de la eficiencia.

¿Cuál de los dos estimadores se eligen?. ¿Son consistentes?

$$\hat{\mu}_1 = \frac{1}{n-1} \sum_{i=1}^n x_i \quad \hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^n x_i$$

Solución:

Sea la v.a x_i = 'ingresos de cierta población' sigue una distribución normal $N(\mu, \sigma)$. Para analizar el sesgo de los estimadores, se calcula la esperanza:

$$E(\hat{\mu}_1) = E\left[\frac{1}{n-1} \sum_{i=1}^n x_i\right] = \frac{1}{n-1} E\left[\sum_{i=1}^n x_i\right] = \frac{1}{n-1} \sum_{i=1}^n E(x_i) = \frac{1}{n-1} (n\mu) = \frac{n}{n-1} \mu$$

El sesgo del estimador $\hat{\mu}_1$ será: $b(\hat{\mu}_1) = E(\hat{\mu}_1) - \mu = \frac{n}{n-1} \mu - \mu = \frac{1}{n-1} \mu$

$$E(\hat{\mu}_2) = E\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} E\left[\sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{1}{n} (n\mu) = \mu$$

El estimador $\hat{\mu}_2$, que es la media muestral, es insesgado (centrado) y el estimador que se elige.

La eficiencia de los estimadores se analiza a través de su varianza:

$$V(\hat{\mu}_1) = V\left[\frac{1}{n-1} \sum_{i=1}^n x_i\right] = \frac{1}{(n-1)^2} V\left[\sum_{i=1}^n x_i\right] = \frac{1}{(n-1)^2} \sum_{i=1}^n V(x_i) = \frac{1}{(n-1)^2} (n\sigma^2) = \frac{n\sigma^2}{(n-1)^2}$$

$$V(\hat{\mu}_2) = V\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n^2} V\left[\sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n V(x_i) = \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n}$$

El estimador más eficiente será el de menor varianza. Comparando las varianzas de los estimadores:

$$V(\hat{\mu}_2) = \frac{\sigma^2}{n} < \frac{n\sigma^2}{(n-1)^2} = V(\hat{\mu}_1) \text{ puesto que } (n-1)^2 < n^2$$

El estimador $\hat{\mu}_2$, que es la media muestral, es el mejor tanto al sesgo como a la eficiencia.

Los dos estimadores son consistentes: $\lim_{n \rightarrow \infty} E(\hat{\mu}_1) = \mu$ y $\lim_{n \rightarrow \infty} V(\hat{\mu}_1) = \lim_{n \rightarrow \infty} \frac{n\sigma^2}{(n-1)^2} = 0$

$$\lim_{n \rightarrow \infty} E(\hat{\mu}_2) = \mu \text{ y } \lim_{n \rightarrow \infty} V(\hat{\mu}_2) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} = 0$$



ESTIMADORES PUNTUALES: INSESGADEZ. ERROR CUADRÁTICO MEDIO



Sea (x_1, x_2, \dots, x_n) una muestra aleatoria simple de una variable aleatoria X con $E(X) = \mu$ y $\text{Var}(X) = \sigma^2$. Calcular el error cuadrático medio para los siguientes estimadores de μ :

$$\hat{\mu}_1 = x_1 \quad \hat{\mu}_2 = \frac{3x_1 - 2x_2 + x_3}{6}$$

Solución:

a) Insesgadez:

$$E(\hat{\mu}_1) = E(x_1) = \mu \rightarrow b(\hat{\mu}_1) = 0$$

$$E(\hat{\mu}_2) = E\left[\frac{3x_1 - 2x_2 + x_3}{6}\right] = \frac{1}{6} [3E(x_1) - 2E(x_2) + E(x_3)] = \frac{1}{6} (3\mu - 2\mu + \mu) = \frac{2}{6}\mu = \frac{1}{3}\mu$$

$$\text{Sesgo: } b(\hat{\mu}_2) = E(\hat{\mu}_2) - \mu = \frac{1}{3}\mu - \mu = -\frac{2}{3}\mu$$

Respecto al sesgo es mejor el primer estimador $\hat{\mu}_1$ que es insesgado o centrado.

Varianza:

$$\text{Var}(\mu_1) = \sigma^2$$

$$\begin{aligned} \text{Var}(\mu_2) &= \text{Var}\left(\frac{3x_1 - 2x_2 + x_3}{6}\right) = \frac{1}{36} \text{Var}(3x_1 - 2x_2 + x_3) = \frac{1}{36} [9\text{Var}(x_1) + 4\text{Var}(x_2) + \text{Var}(x_3)] = \\ &= \frac{1}{36} [9\sigma^2 + 4\sigma^2 + \sigma^2] = \frac{14}{36}\sigma^2 = \frac{7}{18}\sigma^2 \end{aligned}$$

Respecto a la varianza es mejor el segundo estimador $\hat{\mu}_2$ por ser $\frac{7}{18}\sigma^2 < \sigma^2$

El mejor estimador será el que presente menor Error Cuadrático Medio:

$$\text{ECM}(\hat{\mu}_1) = V(\hat{\mu}_1) + [b(\hat{\mu}_1)]^2 = \sigma^2 + 0 = \sigma^2$$

$$\text{ECM}(\hat{\mu}_2) = V(\hat{\mu}_2) + [b(\hat{\mu}_2)]^2 = \frac{7}{18}\sigma^2 + \left(-\frac{2}{3}\mu\right)^2 = \frac{7}{18}\sigma^2 + \frac{4}{9}\mu^2$$

El primer estimador $\hat{\mu}_1$ será mejor si:

$$\sigma^2 < \frac{7}{18}\sigma^2 + \frac{4}{9}\mu^2 \rightarrow \frac{11}{18}\sigma^2 < \frac{4}{9}\mu^2 \rightarrow \sigma^2 < \frac{4 \times 18}{9 \times 11}\mu^2 \rightarrow \sigma^2 < \frac{8}{11}\mu^2$$



ESTIMADORES PUNTUALES: INSESGADEZ. ERROR CUÁDRATICO MEDIO



Para analizar la característica media μ de una población, se extraen m.a.s. de tamaño n , considerando los estimadores de la media:

$$\hat{\mu}_1 = \bar{x} \qquad \hat{\mu}_2 = \frac{1}{n+1} \sum_{i=1}^n x_i$$

- a) Estudiar la inexactitud, la eficiencia relativa y la consistencia de ambos estimadores
- b) Elegir uno de los dos en término del error cuadrático medio

Solución:

a) Inexactitud:

$$E(\hat{\mu}_1) = E(\bar{x}) = E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} E\left(\sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{1}{n} (n\mu) = \mu$$

Sesgo: $b(\hat{\mu}_1) = E(\hat{\mu}_1) - \mu = \mu - \mu = 0$

$$E(\hat{\mu}_2) = E\left(\frac{1}{n+1} \sum_{i=1}^n x_i\right) = \frac{1}{n+1} E\left(\sum_{i=1}^n x_i\right) = \frac{1}{n+1} \sum_{i=1}^n E(x_i) = \frac{1}{n+1} (n\mu) = \frac{n\mu}{n+1}$$

$$b(\hat{\mu}_2) = E(\hat{\mu}_2) - \mu = \frac{n\mu}{n+1} - \mu = \frac{n\mu - n\mu - \mu}{n+1} = \underbrace{-\frac{\mu}{n+1}}_{\substack{\text{sesgado} \\ \text{asintóticamente}}} \xrightarrow{0 \text{ cuando } n \rightarrow \infty}$$

Eficiencia: Sean dos estimadores inexactos $\hat{\theta}_1$ y $\hat{\theta}_2$ de un parámetro desconocido θ . Se dice que $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ si se verifica: $\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2)$.

La eficiencia relativa se mide por la ratio $\frac{\text{Var}(\hat{\theta}_1)}{\text{Var}(\hat{\theta}_2)}$

$$V(\hat{\mu}_1) = V(\bar{x}) = V\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(x_i) = \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n}$$

$$V(\hat{\mu}_2) = V\left(\frac{1}{n+1} \sum_{i=1}^n x_i\right) = \frac{1}{(n+1)^2} \sum_{i=1}^n V(x_i) = \frac{1}{(n+1)^2} (n\sigma^2) = \frac{n}{(n+1)^2} \sigma^2$$

Eficiencia relativa: $\frac{\text{Var}(\hat{\mu}_1)}{\text{Var}(\hat{\mu}_2)} = \frac{\sigma^2/n}{n\sigma^2/(n+1)^2} = \frac{(n+1)^2}{n^2} > 1 \rightarrow \text{Var}(\hat{\mu}_1) > \text{Var}(\hat{\mu}_2)$

El estimador $\hat{\mu}_2$ tiene menor varianza, por lo que es *más eficiente* que $\hat{\mu}_1$



Consistencia: Un estimador $\hat{\theta}$ es *consistente* cuando $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$ y $\lim_{n \rightarrow \infty} V(\hat{\theta}) = 0$

$$\hat{\mu}_1 \equiv \begin{cases} \lim_{n \rightarrow \infty} E(\hat{\mu}_1) = \lim_{n \rightarrow \infty} E(\bar{x}) = \mu \\ \lim_{n \rightarrow \infty} V(\hat{\mu}_1) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} = 0 \end{cases} \quad \text{es consistente}$$

$$\hat{\mu}_2 \equiv \begin{cases} \lim_{n \rightarrow \infty} E(\hat{\mu}_2) = \lim_{n \rightarrow \infty} \left(\mu - \frac{1}{n+1} \mu \right) = \mu \\ \lim_{n \rightarrow \infty} V(\hat{\mu}_2) = \lim_{n \rightarrow \infty} \left[\frac{n}{(n+1)^2} \sigma^2 \right] = 0 \end{cases} \quad \text{es consistente}$$

b) Se elige el estimador que presente menor error cuadrático medio (ECM):

$$ECM(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + [b(\hat{\theta})]^2 \quad \text{sesgo } b(\hat{\theta}) = [E(\hat{\theta}) - \theta]$$

$$ECM(\hat{\mu}_1) = V(\hat{\mu}_1) + [b(\hat{\mu}_1)]^2 = \frac{\sigma^2}{n} + 0 = \frac{\sigma^2}{n}$$

$$ECM(\hat{\mu}_2) = V(\hat{\mu}_2) + [b(\hat{\mu}_2)]^2 = \frac{n}{(n+1)^2} \sigma^2 + \left(\frac{1}{n+1} \mu \right)^2 = \frac{n\sigma^2 + \mu^2}{(n+1)^2}$$

$$\text{El estimador } \hat{\mu}_1 \text{ será elegido si } ECM(\hat{\mu}_1) \leq ECM(\hat{\mu}_2) \rightarrow \frac{\sigma^2}{n} \leq \frac{n\sigma^2 + \mu^2}{(n+1)^2}$$

$$\frac{n\sigma^2 + \mu^2}{(n+1)^2} = \frac{n\sigma^2}{(n+1)^2} + \frac{\mu^2}{(n+1)^2} \rightarrow \frac{\sigma^2}{n} \leq \frac{n\sigma^2}{(n+1)^2} + \frac{\mu^2}{(n+1)^2}$$

$$\frac{\sigma^2}{n} - \frac{n\sigma^2}{(n+1)^2} \leq \frac{\mu^2}{(n+1)^2} \rightarrow \frac{(n+1)^2 \sigma^2 - n^2 \sigma^2}{n(n+1)^2} \leq \frac{\mu^2}{(n+1)^2} \rightarrow$$

$$\rightarrow \frac{(2n+1)\sigma^2}{n(n+1)^2} \leq \frac{\mu^2}{(n+1)^2} \rightarrow \frac{(2n+1)}{n} \leq \frac{\mu^2}{\sigma^2}$$

$$\begin{cases} \text{Si } \frac{2n+1}{n} \leq \frac{\mu^2}{\sigma^2} \rightarrow \hat{\mu}_1 \text{ se elige antes que } \hat{\mu}_2 \\ \text{Si } \frac{2n+1}{n} \geq \frac{\mu^2}{\sigma^2} \rightarrow \hat{\mu}_2 \text{ se elige antes que } \hat{\mu}_1 \end{cases}$$



ESTIMADORES PUNTUALES: INSESGADEZ



La variable aleatoria X representa los gastos mensuales de una empresa, cuya función de densidad es $f(\theta, x) = \theta x^{\theta-1}$ con $\theta > 0$ y $0 < x < 1$. Se realiza una muestra aleatoria simple (m.a.s.) de tamaño 3, y se proponen tres estimadores:

$$\hat{\theta}_1 = \bar{x} \quad \hat{\theta}_2 = \frac{x_1^2 + 2x_2^2 + 3x_3^2}{6} \quad \hat{\theta}_3 = \frac{x_3 - 2x_1 + 4x_2}{6}$$

- a) Calcule los sesgos
- b) Si la muestra que se obtiene es (0,7; 0,1; 0,3), calcule las estimaciones puntuales
- c) ¿Cuáles son las funciones estimadas para las estimaciones anteriores?

Solución:

a) $\hat{\theta}_1 = \bar{x} \Rightarrow E(\hat{\theta}_1) = E\left[\frac{x_1 + x_2 + x_3}{3}\right] = \frac{1}{3} E[x_1 + x_2 + x_3] = \frac{1}{3} (3\mu) = \mu$

donde,

$$\mu = \int_{-\infty}^{\infty} x f(x, \theta) dx = \int_0^1 x f(x, \theta) dx = \int_0^1 x \theta x^{\theta-1} dx = \int_0^1 \theta x^{\theta} dx = \left[\frac{\theta x^{\theta+1}}{\theta+1} \right]_0^1 = \frac{\theta}{\theta+1}$$

Sesgo: $E(\hat{\theta}_1) = \frac{\theta}{\theta+1} \rightarrow b(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta = \frac{\theta}{\theta+1} - \theta = -\frac{\theta^2}{\theta+1}$

• $E(\hat{\theta}_2) = E\left[\frac{x_1^2 + 2x_2^2 + 3x_3^2}{6}\right] = \frac{1}{6} \left[\underbrace{E(x_1^2)}_{\alpha_2} + 2\underbrace{E(x_2^2)}_{\alpha_2} + 3\underbrace{E(x_3^2)}_{\alpha_2} \right] = \frac{1}{6} (6\alpha_2) = \alpha_2$

$$\alpha_2 = E(x^2) = \int_{-\infty}^{\infty} x^2 f(x, \theta) dx = \int_0^1 x^2 f(x, \theta) dx = \int_0^1 x^2 \theta x^{\theta-1} dx = \int_0^1 \theta x^{\theta+1} dx = \left[\frac{\theta x^{\theta+2}}{\theta+2} \right]_0^1 = \frac{\theta}{\theta+2}$$

$$E(\hat{\theta}_2) = E\left[\frac{x_1^2 + 2x_2^2 + 3x_3^2}{6}\right] = \frac{1}{6} \left[\underbrace{E(x_1^2)}_{\alpha_2} + 2\underbrace{E(x_2^2)}_{\alpha_2} + 3\underbrace{E(x_3^2)}_{\alpha_2} \right] = \alpha_2 = \frac{\theta}{\theta+2}$$

Sesgo: $b(\hat{\theta}_2) = E(\hat{\theta}_2) - \theta = \frac{\theta}{\theta+2} - \theta = -\frac{\theta^2 + \theta}{\theta+2}$

• $E(\hat{\theta}_3) = E\left[\frac{x_3 - 2x_1 + 4x_2}{6}\right] = \frac{1}{6} E[x_3 - 2x_1 + 4x_2] = \frac{1}{6} (3\mu) = \frac{1}{2} \mu$



$$\mu = \int_{-\infty}^{\infty} x f(x, \theta) dx = \int_0^1 x f(x, \theta) dx = \int_0^1 x \theta x^{\theta-1} dx = \int_0^1 \theta x^{\theta} dx = \left[\frac{\theta x^{\theta+1}}{\theta+1} \right]_0^1 = \frac{\theta}{\theta+1}$$

$$\text{Sesgo: } b(\hat{\theta}_3) = E(\hat{\theta}_3) - \theta = \frac{1}{2} \left(\frac{\theta}{\theta+1} \right) - \theta = -\frac{2\theta^2 + \theta}{2(\theta+1)}$$

b) Las estimaciones puntuales de la muestra (0,7 ; 0,1 ; 0,3):

$$\hat{\theta}_1 = \frac{0,7 + 0,1 + 0,3}{3} = 0,367$$

$$\hat{\theta}_2 = \frac{0,7^2 + 2 \cdot 0,1^2 + 3 \cdot 0,3^2}{6} = 0,13$$

$$\hat{\theta}_3 = \frac{0,3 - 2 \cdot 0,7 + 4 \cdot 0,1}{6} = -0,117 \rightarrow \text{No puede ser, puesto que } \hat{\theta} > 0$$

c) Las funciones estimadas para las estimaciones de la muestra:

$$\hat{\theta}_1 \rightarrow f(0,367, x) = 0,367 x^{0,367-1} = 0,367 x^{-0,633}$$

$$\hat{\theta}_2 \rightarrow f(0,13, x) = 0,13 x^{0,13-1} = 0,13 x^{-0,87}$$



ESTIMADORES PUNTUALES: EFICIENCIA



La distribución del peso de las manzanas de una determinada cosecha sigue una distribución normal, cuyo peso medio es desconocido y cuya desviación típica es de 7 gramos.

Se pide:

- a) Analizar cuál de los estimadores $\hat{\mu}_1, \hat{\mu}_2$ del peso medio es mejor respecto del sesgo y de la eficiencia, para una muestra aleatoria simple de tamaño cinco.
- b) Si $\hat{\mu}_1 = \frac{1}{5} \sum_{i=1}^5 x_i$ y $\hat{\mu}_2 = x_1 + 2x_2 + 3x_3 - 4x_4 - x_5$, obtener los pesos medios estimados a partir de la muestra (125, 135, 130, 137, 142).

Solución:

- a) El peso de las manzanas sigue una distribución $N(\mu, 7)$. Se calculan las esperanzas para analizar el sesgo de los estimadores.

$$E(\hat{\mu}_1) = E\left[\frac{1}{5} \sum_{i=1}^5 x_i\right] = \frac{1}{5} E\left[\sum_{i=1}^5 x_i\right] = \frac{1}{5} \sum_{i=1}^5 E[x_i] = \frac{1}{5}(5\mu) = \mu$$

$$E(\hat{\mu}_2) = E(x_1 + 2x_2 + 3x_3 - 4x_4 - x_5) = E(x_1) + 2E(x_2) + 3E(x_3) - 4E(x_4) - E(x_5) = \mu + 2\mu + 3\mu - 4\mu - \mu = \mu$$

Los estimadores $\hat{\mu}_1, \hat{\mu}_2$ son insesgados (centrados).

Para analizar la eficiencia de los estimadores se obtienen las varianzas:

$$V(\hat{\mu}_1) = V\left[\frac{1}{5} \sum_{i=1}^5 x_i\right] = \frac{1}{25} V\left[\sum_{i=1}^5 x_i\right] = \frac{1}{25} \sum_{i=1}^5 V[x_i] = \frac{1}{25}(5 \times 49) = \frac{49}{5}$$

$$V(\hat{\mu}_2) = V(x_1 + 2x_2 + 3x_3 - 4x_4 - x_5) = V(x_1) + 4V(x_2) + 9V(x_3) + 16V(x_4) + V(x_5) = (49) + 4(49) + 9(49) + 16(49) + (49) = 31(49) = 1519$$

Como los dos estimadores son insesgados y $V(\hat{\mu}_1) < V(\hat{\mu}_2)$ se elige como mejor el estimador $\hat{\mu}_1$, el peso medio de la muestra de las cinco manzanas.



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD. EFICIENCIA. CONSISTENCIA



Un atleta olímpico de salto de altura se enfrenta a un listón de 2,3 metros. Su entrenador desea estudiar el comportamiento del saltador. Sabe que el número de saltos fallidos por hora es una variable aleatoria distribuida como una Poisson de parámetro λ .

- a) Calcular el estimador máximo verosímil del parámetro λ .
- b) Analizar sus propiedades.

Solución:

a) **MÁXIMA VEROSIMILITUD:** El estimador de máxima verosimilitud (probabilidad conjunta) de θ es el formado por los valores $(\hat{\theta}_1, \dots, \hat{\theta}_n)$ que maximizan la *función de verosimilitud* de la muestra (x_1, \dots, x_n) obtenida:

$$L(\theta) = L(X; \theta) = L(x_1, \dots, x_n; \theta) = P(x_1, \theta) \dots P(x_n, \theta) \text{ caso discreto}$$

En muchas ocasiones, es más práctico encontrar el estimador de máxima verosimilitud al considerar la función *Soporte o Log-verosimilitud* $\ln L(X, \theta)$, en lugar de la función de verosimilitud $L(\theta)$, ya que es más fácil de manejar y presenta los mismos máximos y mínimos.

Se despeja $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_n)$ de la ecuación: $\frac{\partial \ln L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0$ y se obtiene el estimador de máxima verosimilitud E.M.V($\hat{\theta}$)

En este sentido, sea la v.a. $X =$ 'Número de saltos fallidos por hora'

En la distribución de Poisson: $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$ donde, $E(x) = \lambda$, $V(x) = \lambda$

La función de verosimilitud $L(X, \lambda)$ en una muestra aleatoria simple de tamaño n :

$$L(\lambda) = L(X, \lambda) = \prod_{i=1}^n P(x_i, \lambda) = \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \dots \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} = \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\lambda}$$

Función soporte o Log-verosimilitud:

$$\ln L(X, \lambda) = \ln \left[\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\lambda} \right] = \ln \left(\lambda^{\sum_{i=1}^n x_i} \right) - \ln \left(\prod_{i=1}^n x_i! \right) + \ln(e^{-n\lambda}) = \sum_{i=1}^n x_i \ln \lambda - \sum_{i=1}^n \ln(x_i!) - n\lambda$$

$$\ln L(X, \lambda) = \sum_{i=1}^n x_i \ln \lambda - \sum_{i=1}^n \ln(x_i!) - n\lambda$$



Para obtener el estimador de máxima verosimilitud EMV($\hat{\lambda}$), se deriva la expresión anterior respecto del parámetro para obtener, sustituyendo λ por $\hat{\lambda}$

$$\left. \frac{\partial \ln L(X, \lambda)}{\partial \lambda} \right|_{\lambda = \hat{\lambda}} = 0 \rightarrow \sum_{i=1}^n x_i \frac{1}{\lambda} - n \left|_{\lambda = \hat{\lambda}} = 0 \rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

El estimador de máxima verosimilitud (estimador que ofrece mayor credibilidad) viene dado por la media muestral EMV($\hat{\lambda}$) = \bar{x}

b) Propiedades de Insesgadez, Consistencia y Eficiencia

Insesgadez

El estimador $\hat{\lambda}$ es insesgado (centrado) sí $E(\hat{\lambda}) = \lambda$

$$E(\hat{\lambda}) = E\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} E\left(\sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{1}{n} (n\lambda) = \lambda$$

$$V(\hat{\lambda}) = V(\bar{x}) = V\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(x_i) = \frac{1}{n^2} (n\lambda) = \frac{\lambda}{n}$$

Consistencia

Cuando no es posible emplear estimadores de máxima verosimilitud, el requisito mínimo deseable para un estimador es que sea consistente.

El estimador $\hat{\lambda}$ es *consistente* cuando $\lim_{n \rightarrow \infty} E(\hat{\lambda}) = \lambda$ y $\lim_{n \rightarrow \infty} V(\hat{\lambda}) = 0$

$$\lim_{n \rightarrow \infty} E(\hat{\lambda}) = \lim_{n \rightarrow \infty} \lambda = \lambda \quad \text{y} \quad \lim_{n \rightarrow \infty} V(\hat{\lambda}) = \lim_{n \rightarrow \infty} \frac{\lambda}{n} = 0$$

El estimador $\hat{\lambda}$ es consistente

EFICIENCIA: Para que un estimador sea *eficiente* tiene que ser *centrado* y de *varianza mínima*.

La varianza mínima se analiza en virtud de la acotación de Cramer-Rao:

$$V(\hat{\lambda}) \geq \text{CCR} = \frac{[1 + b(\hat{\lambda})]^2}{E\left[\frac{\partial \ln P(x_1, x_2, \dots, x_n; \lambda)}{\partial \lambda}\right]^2}$$

La cantidad de información de Fisher $I(\lambda)$ puede simplificarse en una muestra aleatoria simple (m.a.s.), obteniendo la expresión:

$$I(\lambda) = E\left[\frac{\partial \ln P(x_1, x_2, \dots, x_n; \lambda)}{\partial \lambda}\right]^2 = n \cdot E\left[\frac{\partial \ln P(x; \lambda)}{\partial \lambda}\right]^2$$



$$V(\hat{\lambda}) \geq \text{CCR} = \frac{[1 + b(\hat{\lambda})]^2}{n \cdot E\left[\frac{\partial \ln P(x; \lambda)}{\partial \lambda}\right]^2} \xrightarrow{b(\hat{\lambda})=0} V(\hat{\lambda}) \geq \text{CCR} = \frac{1}{n \cdot E\left[\frac{\partial \ln P(x; \lambda)}{\partial \lambda}\right]^2}$$

El estimador es eficiente cuando $V(\hat{\lambda}) = \text{CCR}$

Siendo, $P(X=x) = \frac{\lambda^x}{x!} e^{-\lambda}$

$$\ln P(x, \lambda) = \ln \left[\frac{\lambda^x}{x!} e^{-\lambda} \right] = x \ln \lambda - \ln(x!) - \lambda$$

$$\frac{\partial \ln P(x, \lambda)}{\partial \lambda} = \frac{x}{\lambda} - 1 = \frac{x - \lambda}{\lambda}$$

$$E\left[\frac{\partial \ln P(x, \lambda)}{\partial \lambda}\right]^2 = E\left[\frac{x - \lambda}{\lambda}\right]^2 = \frac{1}{\lambda^2} E(x - \lambda)^2 = \frac{1}{\lambda^2} E(x - \bar{x})^2 = \frac{1}{\lambda^2} V(x) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

En consecuencia, $V(\hat{\lambda}) \geq \frac{1}{n \frac{1}{\lambda}} = \frac{\lambda}{n}$

El menor valor de la varianza del estimador será $\frac{\lambda}{n}$

Se sabe que $V(\hat{\lambda}) = V(\bar{x}) = \frac{\lambda}{n}$, lo que muestra $V(\hat{\lambda}) = \text{CCR}$, el estimador empleado es *eficiente*.



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD



Una muestra aleatoria (x_1, \dots, x_n) de la población tiene como función de densidad

$$f_{\theta}(x) = \begin{cases} \theta^2 x e^{-\theta x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

Hallar el estimador de máxima verosimilitud de θ

Solución

Función de verosimilitud $L(\theta)$

$$\begin{aligned} L(\theta) &= f_{\theta}(x_1) f_{\theta}(x_2) \dots f_{\theta}(x_n) = \left[\theta^2 x_1 e^{-\theta x_1} \right] \left[\theta^2 x_2 e^{-\theta x_2} \right] \dots \left[\theta^2 x_n e^{-\theta x_n} \right] = \\ &= \theta^{2n} (x_1 \dots x_n) e^{-(\theta x_1 + \theta x_2 + \dots + \theta x_n)} = \theta^{2n} (x_1 \dots x_n) e^{-\theta \sum_{i=1}^n x_i} \end{aligned}$$

Función soporte o Log-verosimilitud:

$$L(\theta) = \theta^{2n} (x_1 \dots x_n) e^{-\theta \sum_{i=1}^n x_i} \rightarrow \ln L(\theta) = \ln \left(\theta^{2n} (x_1 \dots x_n) e^{-\theta \sum_{i=1}^n x_i} \right)$$

$$\ln L(\theta) = (2n) \ln \theta + \ln \prod_{i=1}^n x_i - \theta \sum_{i=1}^n x_i \rightarrow \ln L(\theta) = (2n) \ln \theta + \sum_{i=1}^n \ln x_i - \theta \sum_{i=1}^n x_i$$

Para obtener el estimador de máxima verosimilitud de $\hat{\theta}$ se deriva la expresión anterior respecto al parámetros sustituyendo θ por $\hat{\theta}$, e igualando a cero:

$$\left(\frac{\partial \ln L(\theta)}{\partial \theta} \right)_{\theta = \hat{\theta}} = \left(\frac{2n}{\theta} - \sum_{i=1}^n x_i \right)_{\theta = \hat{\theta}} = 0 \rightarrow \hat{\theta} = \frac{2n}{\sum_{i=1}^n x_i}$$



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD



Sea la distribución $N(\mu, \sigma)$, con la media y varianza desconocidas.
 Calcular los estimadores máximo-verosímiles de μ y σ^2

Solución:

Función de densidad poblacional: $f(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Función de verosimilitud:

$$L(X; \mu, \sigma^2) = \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \right] \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \right] \cdots \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} \right] =$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}}$$

Función soporte o Log-verosimilitud:

$$\ln [L(X; \mu, \sigma^2)] = \ln \left[\frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}} \right] = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}$$

$$\ln [L(X; \mu, \sigma^2)] = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}$$

Para obtener los estimadores de máxima verosimilitud de $\hat{\mu}$ y $\hat{\sigma}^2$ se deriva la expresión anterior respecto de los parámetros, sustituyendo μ por $\hat{\mu}$ e σ^2 por $\hat{\sigma}^2$, e igualando a cero:

$$\left[\frac{\partial \ln L(X; \mu, \sigma^2)}{\partial \mu} \right]_{\mu=\hat{\mu}} = \frac{\sum_{i=1}^n (x_i - \hat{\mu})}{\sigma^2} = 0 \rightarrow \hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

$$\left[\frac{\partial \ln L(X; \mu, \sigma^2)}{\partial \sigma^2} \right]_{\sigma^2=\hat{\sigma}^2} = -\frac{n}{\hat{\sigma}^2} + \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{\hat{\sigma}^3} = 0 \rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n}$$



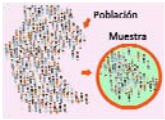
La condición de máximo se verifica, pues: $\left[\frac{\partial^2 \ln L(X; \mu)}{\partial \mu^2} \right]_{\mu = \hat{\mu}} = -\frac{n}{\sigma^2} < 0$

Los estimadores máximo-verosímiles de μ y σ^2 son la media y la varianza muestrales.

En conclusión, $\hat{\mu} = \bar{x}$ y $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \sigma_x^2$



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD



Sea (x_1, \dots, x_n) una muestra aleatoria de tamaño n , distribuida según $f(x, \theta)$ con θ desconocido, donde X representa el tiempo máximo necesario para determinar un proceso en segundos:

$$f(x, \theta) = \begin{cases} (\theta + 1)x^\theta & 0 \leq x \leq 1; \theta > -1 \\ 0 & \text{otro caso} \end{cases}$$

- a) Determinar el estimador máximo verosímil de θ
- b) Determinar la estimación máximo verosímil de θ en una muestra aleatoria simple constituida por los datos:

0,7 0,9 0,6 0,8 0,9 0,7 0,9 0,8

Estimar la probabilidad del tiempo máximo necesario para terminar un proceso, que no exceda de 0,25 segundos ni supere los 0,75 segundos.

- c) Determinar el estimador máximo verosímil de: (i) $\theta + 1$ (ii) $\frac{2\theta + 1}{\theta - 1}$

Solución

a) Función de verosimilitud: $L(\theta) = L(x, \theta) = \prod_{i=1}^n f(x_i, \theta)$

$$L(x_1, x_2, \dots, x_n; \theta) = [(\theta + 1)x_1^\theta] [(\theta + 1)x_2^\theta] \dots [(\theta + 1)x_n^\theta] = (\theta + 1)^n \prod_{i=1}^n x_i^\theta$$

Función soporte o Log-verosimilitud:

$$\ln L(x_1, x_2, \dots, x_n; \theta) = \ln \left((\theta + 1)^n \prod_{i=1}^n x_i^\theta \right) = \ln(\theta + 1)^n + \ln \left(\prod_{i=1}^n x_i^\theta \right) = n \ln(\theta + 1) + \theta \sum_{i=1}^n \ln x_i$$

Para obtener el estimador de máxima verosimilitud $EMV(\hat{\theta})$, se deriva la expresión anterior respecto del parámetro para obtener, sustituyendo θ por $\hat{\theta}$

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n; \theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} = \frac{n}{\hat{\theta} + 1} + \sum_{i=1}^n \ln x_i = 0 \rightarrow \frac{n}{\hat{\theta} + 1} = - \sum_{i=1}^n \ln x_i$$

$$\hat{\theta} = EMV(\theta) = - \frac{n}{\sum_{i=1}^n \ln x_i} - 1$$

- b) El estimador máximo verosímil de θ con los datos de la muestra:

$$\hat{\theta} = - \frac{8}{\ln 0,7 + \ln 0,9 + \ln 0,6 + \ln 0,8 + \ln 0,9 + \ln 0,7 + \ln 0,9 + \ln 0,8} - 1 = 3,027$$

Estimador máximo verosímil $\hat{\theta} = 3,027$



De otra parte,

$$\begin{aligned}
 P(0,25 \leq X \leq 0,75) &= \int_{0,25}^{0,75} f(X, \hat{\theta}) dx = \int_{0,25}^{0,75} (\hat{\theta} + 1) x^{\hat{\theta}} dx = (\hat{\theta} + 1) \frac{x^{\hat{\theta}+1}}{(\hat{\theta} + 1)} \Big|_{0,25}^{0,75} = \\
 &= x^{\hat{\theta}+1} \Big|_{0,25}^{0,75} = 0,75^{4,027} - 0,25^{4,047} = 0,31
 \end{aligned}$$

La probabilidad del tiempo máximo necesario para terminar el proceso (entre 0,25 y 0,75 segundos) es 0,31

c) Estimador máximo verosímil de $\theta + 1$ y $\frac{2\theta + 1}{\theta - 1}$

$$EMV[\theta + 1] = EMV[\theta] + 1 = \hat{\theta} + 1 = 3,027 + 1 = 4,027$$

$$EMV\left(\frac{2\theta + 1}{\theta - 1}\right) = \frac{2 EMV(\theta) + 1}{EMV(\theta) - 1} = \frac{2\hat{\theta} + 1}{\hat{\theta} - 1} = \frac{2 \times 3,027 + 1}{3,027 - 1} = 2,493$$



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD



Calcular el estimador máximo verosímil del parámetro 'a' de las funciones:

- a) $f(x, a) = a^2 e^{-a x}$ siendo $x \geq 0$ en muestras aleatorias simples de tamaño n
 b) $f(x, a) = a e^{-a x}$ para $x \geq 0$, $a > 0$ en muestras aleatorias simples de tamaño 2

Solución:

a) Función de verosimilitud:

$$L(X, a) = L(x_1, x_2, \dots, x_n; a) = (a^2 e^{-a x_1}) \cdot (a^2 e^{-a x_2}) \dots (a^2 e^{-a x_n}) = a^{2n} e^{-a \sum_{i=1}^n x_i}$$

$$\text{Función soporte o Log-verosimilitud: } \ln L(X, a) = \ln (a^{2n} e^{-a \sum_{i=1}^n x_i}) = 2n \ln a - a \sum_{i=1}^n x_i$$

Para obtener el estimador de máxima verosimilitud EMV(\hat{a}), se deriva la expresión anterior respecto del parámetro para obtener, sustituyendo a por \hat{a} e igualando a cero:

$$\left. \frac{\partial \ln L(X, a)}{\partial a} \right|_{a=\hat{a}} = \left. \frac{2n}{a} - \sum_{i=1}^n x_i \right|_{a=\hat{a}} = 0 \rightarrow \hat{a} = \frac{2n}{\sum_{i=1}^n x_i} = \frac{2}{\bar{x}}$$

b) Función de verosimilitud: $L(X, a) = L(x_1, x_2; a) = (a e^{-a x_1}) \cdot (a e^{-a x_2}) = a^2 e^{-a(x_1 + x_2)}$

$$\text{Función soporte o Log-verosimilitud: } \ln L(X, a) = \ln [a^2 e^{-a(x_1 + x_2)}] = 2 \ln a - a(x_1 + x_2)$$

Derivando respecto de a, igualando a cero y sustituyendo a por \hat{a} :

$$\left. \frac{\partial \ln L(X, a)}{\partial a} \right|_{a=\hat{a}} = \left. \frac{2}{a} - (x_1 + x_2) \right|_{a=\hat{a}} = 0 \rightarrow \hat{a} = \frac{2}{x_1 + x_2} = \frac{1}{\bar{x}}$$



ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD. EFICIENCIA



En un estacionamiento el número de veces que se abre la barrera en un intervalo de 10 minutos, para que pasen vehículos en un sector de seguridad, se considera una variable aleatoria con distribución de Poisson de parámetro λ desconocido.

a) En una muestra aleatoria de 8 intervalos de 10 minutos cada uno, elegidos de forma independiente, se registra para cada intervalo el valor que toma la variable en estudio.

3	5	8	7	4	5	6	2
---	---	---	---	---	---	---	---

Encontrar la estimación máximo verosímil de λ

b) Sea (x_1, \dots, x_n) una muestra aleatoria de tamaño n que sigue una distribución de Poisson. Si

$\hat{\lambda}_1 = \sum_{i=1}^n \frac{x_i}{n}$, $\hat{\lambda}_2 = \frac{x_1 + 3x_n}{4}$ son estimadores. Determinar el mejor estimador del parámetro λ

Solución

a) $X =$ "Número de veces que se abre la barrera en un intervalo de 10 minutos", $X \sim P(\lambda)$

$$P(X, \lambda) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & x = 0, 1, 2, \dots \\ 0 & \text{otro caso} \end{cases}$$

Función de verosimilitud: $L(\lambda) = L(X, \lambda) = \prod_{i=1}^n P(x_i, \lambda) = \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \dots \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} = \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\lambda}$

Función soporte o Log-verosimilitud:

$$\ln L(X, \lambda) = \ln \left(\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\lambda} \right) = \ln(\lambda^{\sum_{i=1}^n x_i}) - \ln \left(\prod_{i=1}^n x_i! \right) + \ln(e^{-n\lambda}) = \sum_{i=1}^n x_i \ln \lambda - \sum_{i=1}^n \ln(x_i!) - n\lambda$$

Para obtener el estimador de máxima verosimilitud $EMV(\hat{\lambda})$, se deriva la expresión anterior respecto del parámetro para obtener, sustituyendo λ por $\hat{\lambda}$

$$\frac{\partial \ln L(X, \lambda)}{\partial \lambda} \Big|_{\lambda=\hat{\lambda}} = 0 = \sum_{i=1}^n x_i \frac{1}{\lambda} - n \Big|_{\lambda=\hat{\lambda}} = 0 \rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

El estimador de máxima verosimilitud viene dado por la media muestral $EMV(\hat{\lambda}) = \bar{x}$



Utilizando la muestra aleatoria de ocho intervalos de 10 minutos, se obtiene el estimador máximo verosímil:

$$\hat{\lambda} = \frac{3+5+8+7+4+5+6+2}{8} = 5$$

En consecuencia, en una muestra aleatoria de ocho intervalos de 10 minutos cada uno, elegidos de forma independiente, la estimación máxima verosímil corresponde a que la barrera se abre 5 veces.

b) Se analizan si los estimadores son o no insesgados, esto es, si la esperanza del estimador coincide con el parámetro a estimar.

$$E(\hat{\lambda}_1) = E\left(\sum_{i=1}^n \frac{x_i}{n}\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{n\lambda}{n} = \lambda$$

$$E(\hat{\lambda}_2) = E\left(\frac{x_1 + 3x_n}{4}\right) = \frac{1}{4} [E(x_1) + 3E(x_2)] = \frac{\lambda + 3\lambda}{4} = \lambda$$

Ambos estimadores son insesgados.

Varianza de cada estimador:

$$V(\hat{\lambda}_1) = V\left(\sum_{i=1}^n \frac{x_i}{n}\right) = \frac{1}{n^2} \sum_{i=1}^n V(x_i) = \frac{n\lambda}{n^2} = \frac{\lambda}{n}$$

$$V(\hat{\lambda}_2) = V\left(\frac{x_1 + 3x_n}{4}\right) = \frac{1}{16} [V(x_1) + 9V(x_2)] = \frac{\lambda + 9\lambda}{16} = \frac{10\lambda}{16} = \frac{5\lambda}{8}$$

La efectividad de los estimadores depende del tamaño de la muestra:

- Si la muestra es igual a 1 ($n = 1$) $\rightarrow \hat{\lambda}_2$ es el estimador más eficiente
- Si la muestra es mayor que 1 ($n > 1$) $\rightarrow \hat{\lambda}_1$ es el estimador más eficiente



ESTIMADORES PUNTUALES: EFICIENCIA



En una distribución $N(\mu, \sigma)$ se estima la media poblacional μ mediante la media de una muestra aleatoria simple (x_1, x_2, \dots, x_n) por medio de la función muestral:

$$\hat{\mu} = \frac{\sum_{i=1}^n i x_i}{n}$$

Estúdiense la eficiencia del estimador.

Solución:

Un estimador es eficiente cuando $V(\hat{\theta}) = \text{CCR}$

Para hallar la cota de Cramer-Rao se necesita saber en primer lugar si el estimador es insesgado o no, se calcula para ello la esperanza matemática:

$$\begin{aligned} E(\hat{\mu}) &= E\left[\frac{\sum_{i=1}^n i x_i}{n}\right] = \frac{1}{n} \sum_{i=1}^n E(i x_i) = \frac{1}{n} \sum_{i=1}^n i E(x_i) = \frac{1}{n} [E(x_1) + 2E(x_2) + 3E(x_3) + \dots + nE(x_n)] = \\ &= \frac{1}{n} [\mu + 2\mu + 3\mu + \dots + n\mu] = \frac{\mu}{n} [1+2+3+ \dots + n] = \frac{\mu}{n} \times \frac{(n+1)n}{2} = \frac{(n+1)}{2} \mu \end{aligned}$$

$$E(\hat{\mu}) = \mu + b(\hat{\mu}) \rightarrow \text{sesgo: } b(\hat{\mu}) = E(\hat{\mu}) - \mu = \frac{n+1}{2} \mu - \mu = \frac{n-1}{2} \mu$$

Varianza del estimador:

$$\begin{aligned} V(\hat{\mu}) &= V\left[\frac{\sum_{i=1}^n i x_i}{n}\right] = \frac{1}{n^2} \sum_{i=1}^n V(i x_i) = \frac{1}{n^2} \sum_{i=1}^n i^2 V(x_i) = \frac{1}{n^2} [1^2 V(x_1) + 2^2 V(x_2) + 3^2 V(x_3) + \dots + n^2 V(x_n)] = \\ &= \frac{1}{n^2} [1^2 \sigma^2 + 2^2 \sigma^2 + 3^2 \sigma^2 + \dots + n^2 \sigma^2] = \frac{\sigma^2}{n^2} [1^2 + 2^2 + 3^2 + \dots + n^2] = \frac{\sigma^2}{n^2} \times \frac{(n+1)(2n+1)}{6} \end{aligned}$$

$$V(\hat{\mu}) = \frac{\sigma^2 (n+1)(2n+1)}{6n}$$

$$\text{Cota de Cramer-Rao: } \text{CCR} = \frac{[1+b(\hat{\mu})]^2}{n \cdot E\left[\frac{\partial \ln f(x, \mu)}{\partial \mu}\right]^2}, \quad [1+b(\hat{\mu})]^2 = \left[1 + \frac{n-1}{2} \mu\right]^2 = \frac{(n+1)^2}{4}$$



La Información de Fisher $I(\mu) = n I_0(\mu)$ de la muestra se obtiene a partir de la función de densidad poblacional:

$$I_0(\mu) = E \left[\frac{\partial \ln L(X, \mu)}{\partial \mu} \right]^2 = n \cdot E \left[\frac{\partial \ln L(x, \mu)}{\partial \mu} \right]^2$$

Función de densidad poblacional: $f(x, \mu) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Tomando logaritmos neperianos: $\ln f(x, \mu) = \ln \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) = -\ln \sigma - \frac{1}{2} \ln 2\pi - \frac{(x-\mu)^2}{2\sigma^2}$

Derivando respecto a μ : $\frac{\partial \ln f(x, \mu)}{\partial \mu} = -\frac{1}{2\sigma^2} [-2(x-\mu)] = \frac{x-\mu}{\sigma^2} \rightarrow \left(\frac{\partial \ln f(x, \mu)}{\partial \mu} \right)^2 = \frac{(x-\mu)^2}{\sigma^4}$

La esperanza matemática: $E \left[\frac{\partial \ln f(x, \mu)}{\partial \mu} \right]^2 = E \left[\frac{(x-\mu)^2}{\sigma^4} \right] = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}$

Atendiendo a la aditividad de esta medida: $I(\mu) = n I_0(\mu) = \frac{n}{\sigma^2}$

$$CCR = \frac{[1 + b(\hat{\mu})]^2}{n \cdot E \left[\frac{\partial \ln f(x, \mu)}{\partial \mu} \right]^2} = \frac{(n+1)^2 \sigma^2}{4n}$$

Cuando $n > 1$ la varianza del estimador es mayor que la cota de Cramer-Rao, con lo que el estimador no es eficiente:

$$V(\hat{\mu}) = \frac{\sigma^2 (n+1) (2n+1)}{6n} > \frac{(n+1)^2 \sigma^2}{4n} = CCR$$

Sí $n = 1$ $V(\hat{\mu}) = CCR = \frac{\sigma^2}{n}$ el estimador es suficiente.



ESTIMADORES PUNTUALES: EFICIENCIA



En una distribución $N(\mu, \sigma)$ se estima la media poblacional μ mediante la media de una muestra aleatoria simple (x_1, x_2, \dots, x_n) . El estimador es insesgado y su varianza $V(\bar{x}) = \frac{\sigma^2}{n}$.

Demostrar que la media muestral es un estimador eficiente.

Solución:

La varianza de un estimador verifica siempre la *Cota de Cramer-Rao*: $V(\hat{\theta}) \geq CCR$.

Un estimador es eficiente cuando $V(\hat{\theta}) = CCR$

Para obtener la cota de Cramer-Rao se parte de la función de densidad poblacional:

$$CCR = \frac{1}{n \cdot E \left[\left(\frac{\partial \ln f(x, \mu)}{\partial \mu} \right)^2 \right]}$$

Función de densidad poblacional: $f(x, \mu) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Tomando logaritmos neperianos, se tiene:

$$\ln f(x, \mu) = \ln \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) = -\ln \sigma - \frac{1}{2} \ln 2\pi - \frac{(x-\mu)^2}{2\sigma^2}$$

Derivando respecto a μ :

$$\frac{\partial \ln f(x, \mu)}{\partial \mu} = -\frac{1}{2\sigma^2} [-2(x-\mu)] = \frac{x-\mu}{\sigma^2} \rightarrow \left(\frac{\partial \ln f(x, \mu)}{\partial \mu} \right)^2 = \frac{(x-\mu)^2}{\sigma^4}$$

La esperanza matemática: $E \left[\left(\frac{\partial \ln f(x, \mu)}{\partial \mu} \right)^2 \right] = E \left[\frac{(x-\mu)^2}{\sigma^4} \right] = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}$

Sustituyendo el valor de la esperanza matemática en la expresión de la cota para estimadores insesgados:

$$CCR = \frac{1}{n \cdot E \left[\left(\frac{\partial \ln f(x, \mu)}{\partial \mu} \right)^2 \right]} = \frac{1}{n \cdot \frac{1}{\sigma^2}} = \frac{\sigma^2}{n}$$

La varianza del estimador coincide con la cota de Cramer-Rao, $V(\bar{x}) = CCR$, concluyendo que la media muestral es un estimador eficiente de la media poblacional en la distribución normal.



ESTIMADORES PUNTUALES: EFICIENCIA



A partir de una muestra aleatoria simple X de tamaño n , determinar la matriz de información de Fisher de una población con distribución $N(\mu, \sigma)$, respecto a los dos parámetros μ y σ .

Solución:

EFICIENCIA: Para que un estimador sea *eficiente* tiene que ser *centrado* y de *varianza mínima*. La varianza mínima se analiza en virtud de la acotación de Cramer-Rao (conocida también como desigualdad de la información).

En el caso continuo:

$$V(\hat{\theta}) \geq \text{CCR} = \frac{1}{I(\theta)} = \frac{1}{E\left[\frac{\partial \ln L(x; \theta)}{\partial \theta}\right]^2} = \frac{1}{n \cdot E\left[\frac{\partial \ln f(x; \theta)}{\partial \theta}\right]^2}$$

La información de Fisher $I(\theta)$ es una cota inferior para la varianza de un estimador insesgado del parámetro.

Un estimador es eficiente cuando $V(\hat{\theta}) = \text{CCR} = I(\theta)^{-1}$

Cuando se extiende la condición de regularidad a la segunda derivada en un parámetro bidimensional $\theta = (\mu, \sigma)$, se puede utilizar una forma alternativa de la información de Fisher para obtener una nueva desigualdad de Cramer-Rao:

$$V(\hat{\theta}) \geq \text{CCR} = \frac{1}{I(\theta)} = \frac{1}{-E\left[\frac{\partial^2 \ln L(x; \theta)}{\partial \theta^2}\right]} = \frac{1}{-n E\left[\frac{\partial^2 \ln f(x)}{\partial \theta^2}\right]}$$

Un estimador es eficiente cuando $V(\hat{\mu}, \hat{\sigma}) = \text{CCR} = I(\mu, \sigma)^{-1}$

La matriz de información de Fisher $I(\mu, \sigma)$, matriz de segundas derivadas (Jacobiano), viene dada por la expresión:

$$I(\mu, \sigma) = -E \begin{bmatrix} \frac{\partial^2}{\partial \mu \partial \mu} \ln L(\mu, \sigma) & \frac{\partial^2}{\partial \mu \partial \sigma} \ln L(\mu, \sigma) \\ \frac{\partial^2}{\partial \sigma \partial \mu} \ln L(\mu, \sigma) & \frac{\partial^2}{\partial \sigma \partial \sigma} \ln L(\mu, \sigma) \end{bmatrix}$$

Función de densidad poblacional: $f(x, \mu) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



Función de verosimilitud:

$$L(X; \mu, \sigma^2) = \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_1 - \mu)^2}{2\sigma^2}} \right] \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_2 - \mu)^2}{2\sigma^2}} \right] \cdots \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n - \mu)^2}{2\sigma^2}} \right] =$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}}$$

Función soporte o Log-verosimilitud:

$$\ln [L(X; \mu, \sigma^2)] = \ln \left[\frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}} \right] = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}$$

Para determinar la información de Fisher se deriva respecto de los parámetros μ y σ la expresión:

$$\ln [L(X; \mu, \sigma^2)] = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}$$

$$\frac{\partial \ln L(\mu, \sigma)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \quad \frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma^3}$$

Se calculan las segundas derivadas $\left(\mu = \bar{x}, \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)$ para encontrar la matriz de información de Fisher $I(\mu, \sigma)$

Al ser x_i un elemento de una muestra aleatoria simple de una población normal $N(\mu, \sigma)$, se tiene que $E(x_i - \mu) = 0$ y $E(x_i - \mu)^2 = \sigma^2$

$$\frac{\partial^2 \ln L(X; \mu, \sigma)}{\partial \mu \partial \mu} \Big|_{(\bar{x}, \hat{\sigma}^2)} = \frac{\partial}{\partial \mu} \left(\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \right) \Big|_{(\bar{x}, \hat{\sigma}^2)} = -\frac{1}{\sigma^2} \Big|_{(\bar{x}, \hat{\sigma}^2)} = -\frac{n}{\hat{\sigma}^2}$$

$$\frac{\partial^2 \ln L(X; \mu, \sigma)}{\partial \mu \partial \sigma} \Big|_{(\bar{x}, \hat{\sigma}^2)} = \frac{\partial^2 \ln L(X; \mu, \sigma)}{\partial \sigma \partial \mu} \Big|_{(\bar{x}, \hat{\sigma}^2)} = \frac{\partial}{\partial \sigma} \left(\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \right) \Big|_{(\bar{x}, \hat{\sigma}^2)} = -\frac{2 \sum_{i=1}^n (x_i - \mu)}{\sigma^3} \Big|_{(\bar{x}, \hat{\sigma}^2)} = 0$$



$$\begin{aligned} \left. \frac{\partial^2 \ln L(X; \mu, \sigma)}{\partial \sigma \partial \sigma} \right|_{(\bar{x}, \hat{\sigma}^2)} &= \left. \frac{\partial}{\partial \sigma} \left(-\frac{n}{\sigma} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma^3} \right) \right|_{(\bar{x}, \hat{\sigma}^2)} = \frac{n}{\sigma^2} + \sum_{i=1}^n (x_i - \mu)^2 \left(\frac{-3\sigma^2}{\sigma^6} \right) \Bigg|_{(\bar{x}, \hat{\sigma}^2)} = \\ &= \frac{n}{\sigma^2} - \frac{3 \sum_{i=1}^n (x_i - \mu)^2}{\sigma^4} \Bigg|_{(\bar{x}, \hat{\sigma}^2)} = \frac{n}{\hat{\sigma}^2} - \frac{3n\hat{\sigma}^2}{\hat{\sigma}^4} = -\frac{2n}{\hat{\sigma}^2} \end{aligned}$$

$$I(\mu, \sigma) = -E \begin{bmatrix} \frac{\partial^2}{\partial \mu \partial \mu} \ln L(\mu, \sigma) & \frac{\partial^2}{\partial \mu \partial \sigma} \ln L(\mu, \sigma) \\ \frac{\partial^2}{\partial \sigma \partial \mu} \ln L(\mu, \sigma) & \frac{\partial^2}{\partial \sigma \partial \sigma} \ln L(\mu, \sigma) \end{bmatrix} = - \begin{bmatrix} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & -\frac{2n}{\hat{\sigma}^2} \end{bmatrix} = \begin{bmatrix} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & \frac{2n}{\hat{\sigma}^2} \end{bmatrix}$$



ESTIMADORES PUNTUALES: EFICIENCIA



En la distribución $B(n, p)$ se considera como estimador del parámetro p el estadístico

$$\hat{p} = \frac{\bar{x}}{n}, \text{ siendo } \bar{x} \text{ la media muestral en muestras aleatorias simples de tamaño } n.$$

Hallar la eficiencia del estimador.

Solución

Insesgadez: $E(\hat{p}) = E\left(\frac{\bar{x}}{n}\right) = \frac{np}{n} = p$

El estimador es insesgado o centrado

Un estimador es eficiente cuando su varianza coincide con la Cota de Cramer-Rao: $V(\hat{p}) = CCR$

$$CCR = \frac{1}{n \cdot E\left[\frac{\partial \ln P(x; p)}{\partial \theta}\right]^2} \quad I(\theta) = n \cdot E\left[\frac{\partial \ln P(x; p)}{\partial \theta}\right]^2 \quad \text{Información de Fisher}$$

Se considera la función de cuantía de la distribución binomial: $P(x, p) = \binom{n}{x} p^x (1-p)^{n-x}$

Función soporte o Log-verosimilitud:

$$\ln P(x, p) = \ln \binom{n}{x} + x \ln p + (n-x) \ln(1-p)$$

$$\frac{\partial \ln P(x, p)}{\partial p} = \frac{x}{p} - \frac{(n-x)}{1-p} = \frac{(1-p)x - p(n-x)}{p(1-p)} = \frac{x - np}{p(1-p)}$$

$$E\left[\frac{\partial \ln P(x, p)}{\partial p}\right]^2 = E\left[\frac{x - np}{p(1-p)}\right]^2 = \frac{E(x - np)^2}{p^2(1-p)^2}$$

En una distribución binomial: $E(X) = np, V(X) = E(X - np)^2 = np(1-p)$

con lo que, $E\left[\frac{\partial \ln P(x, p)}{\partial p}\right]^2 = E\left[\frac{x - np}{p(1-p)}\right]^2 = \frac{E(x - np)^2}{p^2(1-p)^2} = \frac{np(1-p)}{p^2(1-p)^2} = \frac{n}{p(1-p)}$

$$CCR = \frac{1}{n \cdot E\left[\frac{\partial \ln P(x; p)}{\partial \theta}\right]^2} = \frac{1}{n \times \frac{n}{p(1-p)}} = \frac{p(1-p)}{n^2}$$

De otra parte, la varianza del estimador:

$$V(\hat{p}) = V\left(\frac{\bar{x}}{n}\right) = \frac{1}{n^2} V(\bar{x}) = \frac{1}{n^2} \times \frac{np(1-p)}{n} = \frac{p(1-p)}{n^2}$$

La varianza del estimador coincide con la cota de Cramer-Rao, con lo que el estadístico $\hat{p} = \frac{\bar{x}}{n}$ es eficiente.



Sea una población definida por: $P(X = -1) = \frac{1-\theta}{2}$, $P(X = 0) = \frac{\theta+\lambda}{2}$,
 $P(X = 1) = \frac{1-\lambda}{2}$, donde $0 < \theta < 1$, $0 < \lambda < 1$.

Estimar los parámetros θ y λ por el método de los momentos, estudiando si son insesgados.

Solución

MÉTODO DE LOS MOMENTOS: Procedimiento consistente en igualar momentos poblacionales respecto al origen α_r a los correspondientes momentos muestrales respecto al origen a_r , formando así tantas ecuaciones como parámetros poblacionales se pretenden estimar:

$$\alpha_1 = E(X) = \mu \rightarrow \hat{\alpha}_1 = a_1 = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

$$\alpha_2 = E(X^2) \rightarrow \hat{\alpha}_2 = a_2 = \frac{\sum_{i=1}^n x_i^2}{n}$$

.....

$$\alpha_r = E(X^r) \rightarrow \hat{\alpha}_r = a_r = \frac{\sum_{i=1}^n x_i^r}{n}$$

Como hay que estimar dos parámetros θ y λ hay que calcular los dos primeros momentos.

momentos poblacionales

$$\alpha_1 = \mu = E(X) = \sum_i x_i P(X=x_i) = (-1)\left(\frac{1-\theta}{2}\right) + (0)\left(\frac{\theta+\lambda}{2}\right) + (1)\left(\frac{1-\lambda}{2}\right) = \frac{\theta-\lambda}{2}$$

$$\alpha_2 = E(X^2) = \sum_i x_i^2 P(X=x_i) = (-1)^2\left(\frac{1-\theta}{2}\right) + (0)^2\left(\frac{\theta+\lambda}{2}\right) + (1)^2\left(\frac{1-\lambda}{2}\right) = \frac{2-\theta-\lambda}{2}$$

momentos muestrales

$$a_1 = \bar{x} = \frac{\sum_i x_i}{n} \quad a_2 = \frac{\sum_i x_i^2}{n}$$

$$\begin{cases} \alpha_1 = a_1 \rightarrow \frac{\theta-\lambda}{2} = \bar{x} \rightarrow \theta - \lambda = 2\bar{x} \\ \alpha_2 = a_2 \rightarrow \frac{2-\theta-\lambda}{2} = a_2 \rightarrow -\theta - \lambda = 2a_2 - 2 \end{cases} \rightarrow \begin{cases} \theta - \lambda = 2\bar{x} \\ -\theta - \lambda = 2a_2 - 2 \end{cases} \rightarrow \begin{cases} \hat{\lambda} = 1 - a_2 - \bar{x} \\ \hat{\theta} = 1 - a_2 + \bar{x} \end{cases}$$



Insesgadez

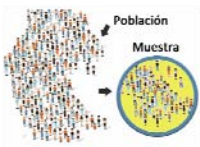
$$E(\hat{\theta}) = E(1 - a_2 + \bar{x}) = 1 - E(a_2) + E(\bar{x}) = 1 - \alpha_2 + \mu = 1 - \left(\frac{2 - \theta - \lambda}{2}\right) + \left(\frac{\theta - \lambda}{2}\right) = \theta$$

$$E(\hat{\lambda}) = E(1 - a_2 - \bar{x}) = 1 - E(a_2) - E(\bar{x}) = 1 - \alpha_2 - \mu = 1 - \left(\frac{2 - \theta - \lambda}{2}\right) - \left(\frac{\theta - \lambda}{2}\right) = \lambda$$

Los estimadores $\hat{\theta}$ y $\hat{\lambda}$ son insesgados.



ESTIMADORES PUNTUALES: SUFICIENCIA. VEROSIMILITUD. MÉTODO DE LOS MOMENTOS



Una muestra aleatoria (x_1, \dots, x_n) de la población tiene como función de densidad

$$f_{\theta}(x) = \begin{cases} \theta x^{\theta-1} & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases} \quad \theta > 0$$

- a) Hallar un estadístico suficiente
- b) Estimador de máxima verosimilitud de θ
- c) Estimador de θ por el método de los momentos

Solución

a) Un estimador $\hat{\theta}$ es suficiente cuando no da lugar a una pérdida de información. Es decir, cuando la información basada en $\hat{\theta}$ es tan buena como la que hiciera uso de toda la muestra.

TEOREMA DE FACTORIZACIÓN: Identifica estadísticos suficientes.

Establece que dada una muestra aleatoria (x_1, \dots, x_n) de una población X con función de densidad f_{θ} (caso continuo), un estadístico $\hat{\theta}$ es suficiente para θ sí y sólo sí:

$$f_{\theta}(x_1, \dots, x_n) = g[\hat{\theta}(x_1, \dots, x_n), \theta] \cdot h(x_1, \dots, x_n)$$

Para encontrar un estadístico suficiente $\hat{\theta}$ hay que factorizar la función de verosimilitud de la forma:

$$L(\theta) = g(\hat{\theta}, \theta) \cdot h(x_1, \dots, x_n)$$

Función de verosimilitud:

$$L(\theta) = f_{\theta}(x_1) \cdot f_{\theta}(x_2) \dots f_{\theta}(x_n) = (\theta x_1^{\theta-1}) (\theta x_2^{\theta-1}) \dots (\theta x_n^{\theta-1}) = \theta^n (x_1 \dots x_n)^{\theta-1}$$

La función de verosimilitud se ha factorizado, $\hat{\theta} = (x_1, \dots, x_n)$ es un estadístico suficiente.

b) Función de verosimilitud: $L(\theta) = \theta^n (x_1, \dots, x_n)^{\theta-1}$

Función soporte o Log-verosimilitud:

$$\ln L(\theta) = \ln \left[\theta^n (x_1, \dots, x_n)^{\theta-1} \right] = \ln \theta^n + \ln \prod_{i=1}^n x_i^{\theta-1} = \ln \theta^n + \sum_{i=1}^n \ln(x_i^{\theta-1})$$

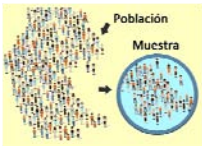
$$\ln L(\theta) = n \ln \theta + (\theta-1) \sum_{i=1}^n \ln(x_i) \rightarrow \frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \ln(x_i) = 0 \rightarrow \hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln(x_i)}$$

c) Se plantea la ecuación $E(X) = \bar{x}$

$$\bar{x} = E(X) = \int_0^1 x f_{\theta}(x) dx = \int_0^1 x \theta x^{\theta-1} dx = \int_0^1 \theta x^{\theta} dx = \theta \left[\frac{x^{\theta+1}}{\theta+1} \right]_0^1 = \frac{\theta}{\theta+1} \rightarrow \hat{\theta} = \frac{\bar{x}}{1-\bar{x}}$$



ESTIMADORES PUNTUALES: MÉTODO DE LOS MOMENTOS. INSESGADEZ



Una muestra aleatoria (x_1, \dots, x_n) de la población tiene como función de densidad

$$f_{\theta}(x) = \begin{cases} e^{-x+\theta} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

- a) Hallar un estimador por el método de los momentos de θ
- b) Estudiar si el estimador encontrado en el apartado anterior es insesgado para estimar θ

Solución

a) Se plantea la ecuación: $E(X) = \bar{x}$

$$\bar{x} = E[X] = \int_{\theta}^{\infty} x f_{\theta}(x) dx = \int_{\theta}^{\infty} x e^{-x+\theta} dx \stackrel{\text{integración por partes}}{=} 1 + \theta \rightarrow \hat{\theta} = \bar{x} - 1$$

$$\int_{\theta}^{\infty} x e^{-x+\theta} dx = \int_{\theta}^{\infty} \underbrace{x}_{u} \underbrace{e^{-x+\theta}}_{dv} dx = \underbrace{x}_{u} \underbrace{(-e^{-x+\theta})}_{v} - \int \underbrace{-e^{-x+\theta}}_{v} \underbrace{dx}_{du} = -x e^{-x+\theta} - e^{-x+\theta} = -(1+x) e^{-x+\theta} = -e^{-\theta} \left(\frac{1+x}{e^x} \right)$$

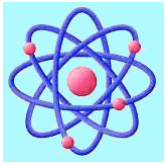
$$\int_{\theta}^{\infty} x e^{-x+\theta} dx = -e^{-\theta} \left(\frac{1+x}{e^x} \right)_{\theta}^{\infty} = 1 + \theta$$

b) Un estimador es insesgado o centrado cuando su valor probable coincide con el valor del parámetro a estimar. Es decir, $E(\hat{\theta}) = \theta$

$$E(\hat{\theta}) = E(\bar{x} - 1) = E(\bar{x}) - 1 = (\theta + 1) - 1 = \theta$$



ESTIMADORES PUNTUALES: MÉTODO DE LOS MOMENTOS. CONSISTENCIA



El coseno X del ángulo con el que se emiten los electrones en un proceso radioactivo es una variable aleatoria con función de densidad:

$$f_{\theta}(x) = \begin{cases} (1+\theta x)/2 & -1 \leq x \leq 1, \quad -1 \leq \theta \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

Se considera una muestra aleatoria (x_1, \dots, x_n) de la variable aleatoria

- a) Obtener el estimador $\hat{\theta}$ por el método de los momentos
- b) Calcular la varianza de este estimador y demostrar que es consistente

Solución

a) Se plantea la ecuación: $E(X) = \bar{x}$

$$\bar{x} = E(X) = \int_{-1}^1 x \left(\frac{1+\theta x}{2} \right) dx = \left[\frac{x^2}{2} + \frac{\theta x^3}{6} \right]_{-1}^1 = \frac{\theta}{3} \rightarrow \hat{\theta} = 3\bar{x}$$

b) $V(\hat{\theta}) = V(3\bar{x}) = 9V(\bar{x}) = 9 \frac{V(X)}{n} = \frac{9}{n} V(X)$

$$V(X) = E(X^2) - [E(X)]^2 = \int_{-1}^1 x^2 \left(\frac{1+\theta x}{2} \right) dx - \left[\frac{\theta}{3} \right]^2 = \left[\frac{x^3}{6} + \frac{\theta x^4}{8} \right]_{-1}^1 - \left[\frac{\theta}{3} \right]^2 = \frac{3-\theta^2}{9}$$

de donde, $V(\hat{\theta}) = \frac{9}{n} V(X) = \frac{9}{n} \left[\frac{3-\theta^2}{9} \right] = \frac{3-\theta^2}{n}$

Consistencia o robustez del estimador $\hat{\theta}$:

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \lim_{n \rightarrow \infty} E(3\bar{x}) = \lim_{n \rightarrow \infty} 3E(\bar{x}) = 3E(X) = 3 \frac{\theta}{3} = \theta$$

$$\lim_{n \rightarrow \infty} V(\hat{\theta}) = \lim_{n \rightarrow \infty} V(3\bar{x}) = \lim_{n \rightarrow \infty} \frac{3-\theta^2}{n} = 0$$

Queda probado que el estimador $\hat{\theta}$ es consistente





Instrumentos Estadísticos Avanzados
Facultad Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández