Gestión Aeronáutica: Estadística Teórica
Facultad Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández

ESTIMADORES
PROPIEDADES. APLICACIÓN.

Un estadístico es cualquier función construida a partir de los valores de la muestra.

Un estimador es un estadístico que se utiliza para estimar - calcular - - aproximar el valor de un parámetro poblacional desconocido.

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:

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

Se pide:

a) Comprobar si los estimadores son insesgados

b) ¿Cuál es el más eficiente?

c) 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 $E(\hat{\theta}) = \theta$

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

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

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

Los tres estimadores son insesgados o centrados.

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

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

$$V\left[\hat{\mu}_2\right] = V\left[\frac{x_1 - 4x_2}{-3}\right] = \frac{1}{9}V\left[x_1 - 4x_2\right] = \frac{1}{9}\left[V(x_1) + 16V(x_2)\right] =$$
$$= \frac{1}{9}\left[17\sigma^2\right] = \frac{17}{9}\sigma^2 = 1{,}89\sigma^2$$

$$V\left[\hat{\mu}_3\right] = V\left[\frac{x_1 + x_2 + x_3 + x_4}{4}\right] = \frac{1}{16} V\left[x_1 + x_2 + x_3 + x_4\right] =$$

$$= \frac{1}{16}\left[V(x_1) + V(x_2) + V(x_3) + V(x_4)\right] = \frac{1}{16}\left[4\sigma^2\right] = \frac{4}{16}\sigma^2 = 0,25\sigma^2$$

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

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

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

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 insesgadez y eficiencia?

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

Solución:

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

$$E(\hat{\theta}_1) = \theta + \underbrace{b(\hat{\theta}_1)}_{\text{Sesgo}}$$

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

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

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

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

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

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

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

$$V(\hat{\theta}_1) = V\left[\frac{x_1 + x_2 + x_3}{4}\right] = \frac{1}{16}\underbrace{\left[V(x_1 + x_2 + x_3)\right]}_{\text{Observaciones independientes}} = \frac{1}{16}\left[V(x_1) + V(x_2) + V(X_3)\right]$$

$$\underset{V(x_i)=4}{=} \frac{1}{16}12 = \frac{12}{16} = \frac{3}{4}$$

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

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: $ECM(\hat{\theta}) = \underbrace{Var(\hat{\theta})}_{\text{Varianza}} + \underbrace{\left[b(\hat{\theta})\right]^2}_{(\text{sesgo})^2}$

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

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

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

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?

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

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:

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

$$E(\hat{\mu}_1) = \mu + b(\hat{\mu}_1) \ \rightarrow \ \text{Sesgo del estimador } \hat{\mu}_1: \ b(\hat{\mu}_1) = E(\hat{\mu}_1) - \mu = \frac{n}{n-1}\mu - \mu = \frac{1}{n-1}\mu$$

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

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:

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

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

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

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

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:

$$\lim_{n\to\infty} E(\hat{\mu}_1) = \mu \quad y \quad \lim_{n\to\infty} V(\hat{\mu}_1) = \lim_{n\to\infty} \frac{n\sigma^2}{(n-1)^2} = 0$$

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

Un estimador es consistente si se aproxima al valor del parámetro cuando el tamaño de la muestra es suficientemente grande.

Universidad Autónoma de Madrid

ESTIMADORES PUNTUALES: INSESGADEZ. ERROR CUADRÁTICO MEDIO

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

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

Solución:

a) Insesgadez:

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

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

$$E(\hat{\mu}_2) = \mu_2 + b(\hat{\mu}_2) \;\rightarrow\; \text{Sesgo: } b(\hat{\mu}_2) = E(\hat{\mu}_2) - \mu = \frac{1}{3}\mu - \mu = -\frac{2}{3}\mu$$

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

Varianza:

$$Var(\mu_1) = \sigma^2$$

$$Var(\mu_2) = Var\left(\frac{3x_1 - 2x_2 + x_3}{6}\right) = \frac{1}{36}Var\left(3x_1 - 2x_2 + x_3\right) = \frac{1}{36}\left[9\,Var(x_1) + 4\,Var(x_2) + Var(x_3)\right] =$$

$$= \frac{1}{36}\left[9\sigma^2 + 4\sigma^2 + \sigma^2\right] = \frac{14}{36}\sigma^2 = \frac{7}{18}\sigma^2$$

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:

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

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

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

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

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:

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

a) Estudiar la insesgadez, la eficiencia relativa y la consistencia de ambos estimadores

b) Elegir uno de los dos en término del error cuadrático medio

Solución:

a) Insesgadez:

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

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

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

$$E(\hat{\mu}_2) = \mu + b(\hat{\mu}_2) \;\rightarrow\; b(\hat{\mu}_2) = E(\hat{\mu}_2) - \mu = \frac{n\mu}{n+1} - \mu = \frac{n\mu - n\mu - \mu}{n+1} = \overbrace{\underbrace{-\frac{\mu}{n+1}}_{\text{insesgado asintoticamente}}}^{\rightarrow 0 \text{ cuando } n \rightarrow \infty}$$

Eficiencia: Sean dos estimadores insesgados $\hat{\theta}_1$ 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: $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$.

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

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

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

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

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 $\lim\limits_{n\to\infty} E(\hat{\theta}) = \theta$ y $\lim\limits_{n\to\infty} V(\hat{\theta}) = 0$

Un estimador es consistente si se aproxima al valor del parámetro cuando el tamaño de la muestra es suficientemente grande.

$$\hat{\mu}_1 \equiv \begin{cases} \lim\limits_{n\to\infty} E(\hat{\mu}_1) = \lim\limits_{n\to\infty} E(\overline{x}) = \mu \\ \lim\limits_{n\to\infty} V(\hat{\mu}_1) = \lim\limits_{n\to\infty} \dfrac{\sigma^2}{n} = 0 \end{cases} \qquad \text{es consistente}$$

$$\hat{\mu}_2 \equiv \begin{cases} \lim\limits_{n\to\infty} E(\hat{\mu}_2) = \lim\limits_{n\to\infty} \left(\mu - \dfrac{1}{n+1}\mu\right) = \mu \\ \lim\limits_{n\to\infty} V(\hat{\mu}_2) = \lim\limits_{n\to\infty} \left[\dfrac{n}{(n+1)^2}\sigma^2\right] = 0 \end{cases} \qquad \text{es consistente}$$

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

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

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

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

El estimador $\hat{\mu}_1$ será elegido sí $ECM(\hat{\mu}_1) \leq ECM(\hat{\mu}_2) \;\rightarrow\; \dfrac{\sigma^2}{n} \leq \dfrac{n\sigma^2+\mu^2}{(n+1)^2}$

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

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

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

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

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:

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

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

donde,

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

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

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

$$\alpha_2 = E(x^2) = \int_{-\infty}^{\infty} x^2\, f(x,\theta)\, dx = \int_0^1 x^2\, f(x,\theta)\, dx = \int_0^1 x^2\, \theta\, x^{\theta-1}\, dx =$$

$$= \int_0^1 \theta\, x^{\theta+1}\, dx = \left[\frac{\theta\, x^{\theta+2}}{\theta+2}\right]_0^1 = \frac{\theta}{\theta+2}$$

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

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

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

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

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

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

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

$$\hat{\theta}_2 = \frac{0{,}7^2+2.0{,}1^2+3.0{,}3^2}{6} = 0{,}13$$

$$\hat{\theta}_3 = \frac{0{,}3-2.0{,}7+4.0{,}1}{6} = -0{,}117 \;\rightarrow\; \text{No puede ser, puesto que } \hat{\theta} > 0$$

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

$$f(\theta_1, x) = \theta_1\, x^{\theta_1-1} = f(0{,}367, x) = 0{,}367\, x^{0{,}367-1} = 0{,}367\, x^{-0{,}633}$$

$$f(\theta_2, x) = \theta_2\, x^{\theta_2-1} = f(0{,}13, x) = 0{,}13\, x^{0{,}13-1} = 0{,}367\, x^{-0{,}87}$$

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.

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

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

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

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

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

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

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 $(\hat{\theta}_1, \cdots, \hat{\theta}_n)$ que maximizan la *función de verosimilitud* de la muestra $(x_1, \cdots, x_n)$ obtenida:

$$L(\theta) = L(X;\theta) = L(x_1, \cdots, x_n; \theta) = P(x_1, \theta) \,.\, P(x_2, \theta) \cdots P(x_n, \theta) \quad \text{caso discreto}$$

$$L(\theta) = L(X;\theta) = L(x_1, \cdots, x_n; \theta) = f(x_1, \theta) \,.\, f(x_2, \theta) \cdots f(x_n, \theta) \quad \text{caso continuo}$$

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 $\hat{\theta} = (\hat{\theta}_1, \cdots, \hat{\theta}_n)$ de la ecuación: $\left.\dfrac{\vartheta \ln L(\theta)}{\vartheta \theta}\right|_{\theta = \hat{\theta}} = 0$ 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: $P(X = x) = \dfrac{\lambda^x}{x!} e^{-\lambda}$ donde, $E(x) = \lambda$, $V(x) = \lambda$

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

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

Función soporte o Log-verosimilitud:

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

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

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

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

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 $\lim\limits_{n\to\infty} E(\hat{\lambda}) = \lambda$ y $\lim\limits_{n\to\infty} V(\hat{\lambda}) = 0$

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

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 Cramér-Rao:

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

La varianza de cualquier estimador siempre es igual o menor que la cota de Cramer-Rao.

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

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

$$V(\hat{\lambda}) \geq CCR = \frac{\left[1 + \dfrac{\vartheta\, b(\hat{\lambda})}{\vartheta\,\lambda}\right]^2}{n \,.\, E\left[\left(\dfrac{\vartheta \ln P(x ; \lambda)}{\vartheta\lambda}\right)^2\right]} \xrightarrow{b'(\hat{\lambda}) = 0} V(\hat{\lambda}) \geq CCR = \frac{1}{n \,.\, E\left[\left(\dfrac{\vartheta \ln P(x ; \lambda)}{\vartheta\lambda}\right)^2\right]}$$

La cota de Cramér-Rao, también llamada desigualdad de Cramér-Rao o desigualdad de información, también se puede expresar:

$$V(\hat{\lambda}) \geq CCR = \frac{1}{n \,.\, E\left[\left(\dfrac{\vartheta \ln P(x ; \lambda)}{\vartheta\lambda}\right)^2\right]} = \frac{1}{I_n(\lambda)} = \frac{1}{n \,.\, I_1(\lambda)} = \frac{1}{-n \,.\, E\left(\dfrac{\vartheta^2 \ln P(x ; \lambda)}{\vartheta\lambda^2}\right)}$$

siendo $I_1(\lambda) = -E\left[\dfrac{\vartheta^2 \ln P(x ; \lambda)}{\vartheta\lambda^2}\right]$ la cantidad de información de Fisher para una muestra

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

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

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

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

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

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

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

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

OTRO PROCEDIMIENTO: $\dfrac{\vartheta \ln P(x, \lambda)}{\vartheta\lambda} = \dfrac{x}{\lambda} - 1 \rightarrow \dfrac{\vartheta^2 \ln P(x, \lambda)}{\vartheta\lambda^2} = -\dfrac{x}{\lambda^2}$

$$I_1(\lambda) = -E\left[\frac{\vartheta^2 \ln P(x\,;\,\lambda)}{\vartheta\lambda^2}\right] = -E\left[\frac{-x}{\lambda^2}\right] = \frac{1}{\lambda^2}E(x) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$V(\hat{\lambda}) \geq CCR = \frac{1}{n\,.\,I_1(\lambda)} = \frac{1}{n\,.\,\frac{1}{\lambda}} = \frac{\lambda}{n}$$

ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD

Una muestra aleatoria $(x_1, \cdots, x_n)$ 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 $L(\theta)$

$$L(\theta) = f_\theta(x_1)\, f_\theta(x_2) \ldots f_\theta(x_n) = \left[\theta^2 x_1 e^{-\theta x_1}\right]\left[\theta^2 x_2 e^{-\theta x_2}\right] \ldots \left[\theta^2 x_n e^{-\theta x_n}\right] =$$

$$= \theta^{2n}(x_1 \ldots x_n)\, e^{-(\theta x_1 + \theta x_2 + \cdots + \theta x_n)} = \theta^{2n}(x_1 \ldots x_n)\, e^{-\theta \sum_{i=1}^{n} x_i}$$

Función soporte o Log-verosimilitud:

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

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

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:

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

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

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:

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

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

Para 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:

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

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

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

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

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

ESTIMADORES PUNTUALES: MÁXIMA VEROSIMILITUD

Sea $(x_1, \cdots, x_n)$ 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:

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

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) $\dfrac{2\theta+1}{\theta-1}$

Solución

a) Función de verosimilitud: $L(\theta) = L(X,\theta) = \prod_{i=0}^{n} f(x_i,\theta)$

$$L(x_1, x_2, \ldots, x_n; \theta) = \left[(\theta+1)\, x_1^{\theta}\right]\left[(\theta+1)\, x_2^{\theta}\right] \ldots \left[(\theta+1)\, x_n^{\theta}\right] = (\theta+1)^n \prod_{i=0}^{n} x_i^{\theta}$$

Función soporte o Log-verosimilitud:

$$\ln L(x_1, x_2, \ldots, x_n; \theta) = \ln\left((\theta+1)^n \prod_{i=0}^{n} x_i^{\theta}\right) = \ln(\theta+1)^n + \ln\left(\prod_{i=0}^{n} x_i^{\theta}\right) = n\ln(\theta+1) + \theta\sum_{i=0}^{n} \ln x_i$$

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

$$\left.\frac{\vartheta \ln L(x_1, x_2, \ldots, x_n; \theta)}{\vartheta\theta}\right|_{\theta=\hat{\theta}} = \frac{n}{\hat{\theta}+1} + \sum_{i=0}^{n} \ln x_i = 0 \quad\rightarrow\quad \frac{n}{\hat{\theta}+1} = -\sum_{i=0}^{n} \ln x_i$$

$$\hat{\theta} = EMV(\theta) = -\frac{n}{\sum_{i=0}^{n} \ln x_i} - 1$$

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,

$$P(0{,}25 \le X \le 0{,}75) = \int_{0,25}^{0,75} f(X,\hat{\theta})\,dx = \int_{0,25}^{0,75} (\hat{\theta}+1)\, x^{\hat{\theta}}\, dx = (\hat{\theta}+1)\, \left.\frac{x^{\hat{\theta}+1}}{(\hat{\theta}+1)}\right|_{0,25}^{0,75} =$$

$$= \left. x^{\hat{\theta}+1} \right|_{0,25}^{0,75} = 0{,}75^{4,027} - 0{,}25^{4,047} = 0{,}31$$

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 $\dfrac{2\theta+1}{\theta-1}$

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

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

ESTIMADORES PUNTUALES: MÁXIMA VERSOSIMILITUD

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:

$$L(X,a) = L(x_1, x_2, \cdots, x_n; a) = (a^2 e^{-a\,x_1}).(a^2 e^{-a\,x_2}) \cdots (a^2 e^{-a\,x_n}) = a^{2n}\, e^{-a \sum\limits_{i=1}^{n} x_i}$$

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

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:

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

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

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

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

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

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 $(x_1, \cdots, x_n)$ 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", $X \sim P(\lambda)$

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

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

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 $\hat{\lambda}$

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

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.

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

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

Ambos estimadores son insesgados.

Varianza de cada estimador:

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

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

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 $(x_1, x_2, \cdots, x_n)$ por medio de la función muestral:

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

Estúdiese la eficiencia del estimador.

Solución:

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

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:

$$E(\hat{\mu}) = E\left[\frac{\sum_{i=1}^{n} i x_i}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E(i x_i) = \frac{1}{n}\sum_{i=1}^{n} i\,E(x_i) = \frac{1}{n}\left[E(x_1) + 2E(x_2) + 3E(x_3) + \cdots + nE(x_n)\right] =$$

$$= \frac{1}{n}\left[\mu + 2\mu + 3\mu + \cdots + n\mu)\right] = \frac{\mu}{n}\left[1 + 2 + 3 + \cdots + n\right] = \frac{\mu}{\cancel{n}} \times \frac{(n+1)\,\cancel{n}}{2} = \frac{(n+1)}{2}\,\mu$$

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

Varianza del estimador:

$$V(\hat{\mu}) = V\left[\frac{\sum_{i=1}^{n} i x_i}{n}\right] = \frac{1}{n^2}\sum_{i=1}^{n} V(i x_i) = \frac{1}{n^2}\sum_{i=1}^{n} i^2\,V(x_i) = \frac{1}{n^2}\left[1^2V(x_1) + 2^2V(x_2) + 3^2V(x_3) + \cdots + n^2V(x_n)\right] =$$

$$= \frac{1}{n^2}\left[1^2\sigma^2 + 2^2\sigma^2 + 3^2\sigma^2 + \cdots + n^2\sigma^2\right] = \frac{\sigma^2}{n^2}\left[1^2 + 2^2 + 3^2 + \cdots + n^2\right] = \frac{\sigma^2}{n^{\cancel{2}}}\,\frac{\cancel{n}\,(n+1)\,(2n+1)}{6}$$

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

Cota de Cramér-Rao: $\displaystyle CCR = \frac{\left[1 + \frac{\vartheta\, b(\hat{\mu})}{\vartheta\,\mu}\right]^2}{n\,.\,E\left[\left(\frac{\vartheta \ln f(x,\mu)}{\vartheta\mu}\right)^2\right]}$, $\displaystyle \left[1 + \frac{\vartheta\, b(\hat{\mu})}{\vartheta\,\mu}\right]^2 = \left[1 + \frac{n-1}{2}\right]^2 = \frac{(n+1)^2}{4}$

La Información de Fisher $I_n(\mu) = n\, I_1(\mu)$ de la muestra se obtiene a partir de la función de densidad poblacional:

$$I_1(\mu) = E\left[\left(\frac{\vartheta \ln L(X,\, \mu)}{\vartheta \mu}\right)^2\right] = -E\left(\frac{\vartheta^2 \ln L(x,\mu)}{\vartheta \mu^2}\right)$$

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

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

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

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

Atendiendo a la aditividad de esta medida: $I(\mu) = n\, I_1(\mu) = \dfrac{n}{\sigma^2}$

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

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

De una población que sigue una distribución Poisson de parámetro λ desconocido, se obtiene una muestra aleatoria simple de tamaño 3, (x_1, x_2, x_3), y se construyen los siguientes estimadores de la varianza poblacional:

$$\hat{V}_1 = \frac{x_1 + x_2}{2} \qquad \hat{V}_2 = x_1 + x_2 - 2x_3 \qquad \hat{V}_3 = \frac{1}{2}\sum_{i=1}^{3}(x_i - \overline{x})^2$$

Se pide:

a) ¿Son insesgados?

b) Calcular el error mínimo cuadrático de $\hat{V}_2$

c) ¿Es la media muestral, $\overline{x} = \frac{1}{3}\sum_{i=1}^{3} x_i$, un estimador insesgado para la varianza poblacional?.

Compararla con el estimador $\hat{V}_1$.

d) ¿Alguno de los estimadores propuestos es eficiente?

Solución:

a) $X \sim P(\lambda) \quad \rightarrow \quad E(X_i) = \lambda \;, \quad Var(X_i) = \lambda$

$$E(\hat{V}_1) = E\left[\frac{x_1 + x_2}{2}\right] = \frac{1}{2}\left[E(x_1) + E(x_2)\right] = \frac{1}{2}(\lambda + \lambda) = \lambda$$

Sesgo: $b(\hat{V}_1) = E(\hat{V}_1) - \lambda = \lambda - \lambda = 0$ es un estimador insesgado o centrado

$$E(\hat{V}_2) = E\left[x_1 + x_2 - 2x_3\right] = E(x_1) + E(x_2) - 2E(x_3) = \lambda + \lambda - 2\lambda = 0$$

Sesgo: $b(\hat{V}_2) = E(\hat{V}_2) - \lambda = 0 - \lambda = -\lambda$ es un estimador sesgado

$\hat{V}_3 = \frac{1}{2}\sum_{i=1}^{3}(x_i - \overline{x})^2$ es la cuasivarianza muestral, estimador insesgado de la varianza poblacional, en consecuencia estimador insesgado de λ

$$\hat{V}_3 = \frac{1}{2}\sum_{i=1}^{3}(x_i - \overline{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{n}{n-1}\,\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n} = \frac{n}{n-1}\,\hat{\sigma}^2$$

Siendo, $E(\hat{\sigma}^2) = \frac{n-1}{n}\,\sigma^2 \quad \rightarrow \quad E(\hat{\sigma}^2) = \frac{2}{3}\lambda$

$$E(\hat{V}_3) = E\left(\frac{3}{2}\,\hat{\sigma}^2\right) = \frac{3}{2}E(\hat{\sigma}^2) = \frac{3}{2} \times \frac{2}{3}\,\lambda = \lambda$$

Por tanto, $\hat{V}_1$ y $\hat{V}_3$ son estimadores insesgados de la varianza poblacional.

b) $ECM(\hat{V}_2) = Var(\hat{V}_2) + b^2(\hat{V}_2)$ siendo $b(\hat{V}_2)$ el sesgo de $\hat{V}_2$

$$Var(\hat{V}_2) = Var\left[x_1 + x_2 - 2x_3\right] = Var(x_1) + Var(x_2) + 4Var(x_3) = \lambda + \lambda + 4\lambda = 6\lambda$$

$$ECM(\hat{V}_2) = 6\lambda + (-\lambda)^2 = 6\lambda + \lambda^2$$

c) La media muestral $\overline{x} = \frac{1}{3}\sum_{i=1}^{3} x_i$ un estimador insesgado de la media poblacional, también el estimador $\hat{V}_1$ es un estimador insesgado de la media poblacional.

$$Var(\hat{V}_1) = Var\left[\frac{x_1 + x_2}{2}\right] = \frac{1}{4}\left[Var(x_1) + Var(x_2)\right] = \frac{1}{4}(\lambda + \lambda) = \frac{\lambda}{2}$$

$$Var(\overline{x}) = Var\left[\frac{1}{3}\sum_{i=1}^{3} x_i\right] = \frac{1}{9}\sum_{i=1}^{3} Var(x_i) = \frac{1}{9}\left[Var(x_1) + Var(x_2) + Var(x_3)\right] = \frac{1}{9}(\lambda + \lambda + \lambda) = \frac{\lambda}{3}$$

Se elige el estimador media muestral por tener menor varianza.

d) 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:

La cota de Cramér-Rao para los estimadores sesgados: $CCR = \dfrac{\left[1 + \dfrac{\vartheta\, b(\hat{\lambda})}{\vartheta\,\lambda}\right]^2}{n \,.\, I_1(\lambda)}$

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \rightarrow \ln P = x \ln\lambda - \ln x! - \lambda \rightarrow \frac{\partial \ln P}{\partial \lambda} = \frac{x}{\lambda} - 1$$

La cantidad de información de Fisher con una muestra: $I_1(\lambda) = -E\left[\dfrac{\partial^2 \ln P}{\partial \lambda^2}\right]$

$$\frac{\partial^2 \ln P}{\partial \lambda^2} = -\frac{x}{\lambda^2} \rightarrow I_1(\lambda) = -E\left[-\frac{x}{\lambda^2}\right] = \frac{1}{\lambda^2}E(x) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

En el caso de los estimadores $\hat{V}_1$ y $\hat{V}_3$ que son insesgados y tener una muestra de tamaño 3, la cota de Cramer-Rao se expresa: $CCR = \dfrac{1}{3 \,.\, I_3(\lambda)} = \dfrac{1}{3 \times \dfrac{1}{\lambda}} = \dfrac{\lambda}{3}$

El estimador media muestral es el eficiente al tener menor varianza.

Cota de Cramer-Rao en el estimador sesgado $\hat{V}_2$: $CCR = \dfrac{\left[1 + (-1)\right]^2}{3 \times \dfrac{1}{\lambda}} = 0$

No es eficiente, se ha visto que la varianza del estimador $\hat{V}_2$ no es nula.

Universidad Autónoma de Madrid

ESTIMADORES PUNTUALES: EFICIENC IA

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

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:

$$CCR = \frac{1}{n\,.\,E\left[\left(\dfrac{\vartheta \ln f(x,\mu)}{\vartheta\mu}\right)^2\right]} = \frac{1}{-n\;E\left[\dfrac{\vartheta^2 \ln P(x\;;\;\lambda)}{\vartheta\lambda^2}\right]}$$

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

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:

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

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

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

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

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

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:

$$V(\hat{\theta}) \geq CCR = \frac{1}{n \,.\, E\left[\left(\dfrac{\vartheta \ln f(x, \mu)}{\vartheta \mu}\right)^2\right]}$$

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

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

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:

$$V(\hat{\theta}) \geq CCR = \frac{1}{n \,.\, I_1(\theta)} = \frac{1}{-n\, E\left[\dfrac{\vartheta^2 \ln L(x\,;\,\theta)}{\vartheta \theta^2}\right]} = \frac{1}{-n\, E\left[\dfrac{\vartheta^2 \ln f(x)}{\vartheta \theta^2}\right]}$$

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

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

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

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

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:

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

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

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

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

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 $E(x_i-\mu)^2 = \sigma^2$

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

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

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

$$= \frac{n}{\sigma^2} - \left.\frac{3\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^4}\right|_{(\bar{x},\,\hat{\sigma}^2)} = \frac{n}{\hat{\sigma}^2} - \frac{3n\hat{\sigma}^2}{\hat{\sigma}^4} = -\frac{2n}{\hat{\sigma}^2}$$

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

ESTIMADORES PUNTUALES: EFICIENCIA

En la distribución $B(n, p)$ se considera como estimador del parámetro p el estadístico $\hat{p} = \dfrac{\bar{x}}{n}$, siendo $\bar{x}$ la media muestral en muestras aleatorias simples de tamaño n.

Hallar la eficiencia del estimador.

Solución

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

El estimador es insesgado o centrado

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

$$CCR = \frac{1}{n\, E\left[\left(\dfrac{\vartheta \ln P(x\,;\,p)}{\vartheta\theta}\right)^2\right]} = \frac{1}{-n\,.\,E\left(\dfrac{\vartheta^2 \ln P(x\,;\,p)}{\vartheta\theta^2}\right)} \qquad I_1(\theta) = -\,E\left[\frac{\vartheta^2 \ln P(x\,;\,p)}{\vartheta\theta^2}\right] \text{ Información de Fisher}$$

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

Función soporte o Log-verosimilitud:

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

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

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

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

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

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 $\hat{p} = \dfrac{\bar{x}}{n}$ es eficiente.

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

$P(X=1)=\dfrac{1-\lambda}{2}$, 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.

$$\overbrace{\begin{aligned}\alpha_1 &= \mu = E(X) = \sum_i x_i P(X=x_i) = (-1)\left(\frac{1-\theta}{2}\right) + (0)\left(\frac{\theta+\lambda}{2}\right) + (1)\left(\frac{1-\lambda}{2}\right) = \frac{\theta-\lambda}{2}\end{aligned}}^{\text{momentos poblacionales}}$$

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

$$\overbrace{a_1 = \bar{x} = \frac{\sum_i x_i}{n} \qquad a_2 = \frac{\sum_i x_i^2}{n}}^{\text{momentos muestrales}}$$

$$\begin{cases} \alpha_1 = a_1 \;\rightarrow\; \dfrac{\theta-\lambda}{2} = \bar{x} \;\rightarrow\; \theta-\lambda = 2\bar{x} \\ \alpha_2 = a_2 \;\rightarrow\; \dfrac{2-\theta-\lambda}{2} = a_2 \;\rightarrow\; -\theta-\lambda = 2a_2 - 2 \end{cases} \rightarrow \begin{cases} \theta-\lambda = 2\bar{x} \\ -\theta-\lambda = 2a_2 - 2 \end{cases} \rightarrow \boxed{\begin{array}{l} \hat{\lambda} = 1-a_2-\bar{x} \\ \hat{\theta} = 1-a_2+\bar{x} \end{array}}$$

Insesgadez

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

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

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

Oposiciones INE Estadísticos Facultativos

Se realiza el experimento de lanzar una moneda tres veces y registrar el número de caras obtenidas. Al repetir 80 veces el experimento se ha obtenido 7 veces ninguna cara, 24 veces una cara, 35 veces dos caras y 14 veces tres caras.

Si p es "la probabilidad de obtener una cara al lanzar la moneda", se pide lo siguiente:

a) Obtenga el estimador de p por el método de los momentos.

b) ¿Es insesgado el estimador obtenido en el apartado anterior?

c) Calcule la varianza del estimador obtenido en el apartado (a).

d) Plantee y resuelva el contraste adecuado para estudiar si la moneda está equilibrada.

(Nota: para un valor $\alpha = 0{,}05$, el valor crítico necesario de la distribución del estadístico es $7{,}81$)

Solución:

a) El experimento aleatorio consiste en lanzar una moneda tres veces, siendo X = "Número de caras obtenidas en los tres lanzamientos", de forma que puede tomar los valores 0, 1, 2 y 3. Repitiendo el proceso 80 veces, se obtiene:

$X = x_i$	0	1	2	3
n_i	7	24	35	14

$p = P(\text{obtener cara al lanzar la moneda})$.

La variable aleatoria $X \sim B(n=3\,,\,p)$, $E(X) = 3p$, $V(X) = 3pq = 3p(1-p)$

Para obtener el estimador por el método de los momentos se necesita igualar los momentos muestrales a los momentos poblacionales.

Primer momento poblacional: $\alpha_1 = E(X) = 3p$ Primer momento muestral: $a_1 = \bar{x}$

$$\alpha_1 = a_1 \;\rightarrow\; 3p = \bar{x} \;\rightarrow\; \hat{p} = \frac{\bar{x}}{3} = \frac{1}{3}\left[\frac{0 \times 7 + 1 \times 24 + 2 \times 35 + 3 \times 14}{7 + 24 + 35 + 14}\right] = 0{,}566$$

b) El estimador es insesgado o centrado cuando: $E(\hat{p}) = p$

$$E(\hat{p}) = E\left(\frac{\bar{x}}{3}\right) = E\left(\frac{1}{3n}\sum_{i=1}^{n} x_i\right) = \frac{1}{3n}\sum_{i=1}^{n} E(x_i) = \frac{1}{3n}\sum_{i=1}^{n} 3p = \frac{1}{3n} \times n \times 3p = p$$

El estimador es insesgado

c) $$\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\frac{\bar{x}}{3}\right) = \mathrm{Var}\left(\frac{1}{3n}\sum_{i=1}^{n} x_i\right) = \frac{1}{9n^2}\sum_{i=1}^{n} \mathrm{Var}(x_i) = \frac{1}{9n^2} \times n \times 3p(1-p) = \frac{p(1-p)}{3n}$$

d) Que la moneda esté equilibrada es equivalente a efectuar un contraste de bondad de ajuste, estableciendo las hipótesis: $H_0 : p = \frac{1}{2}$, $H_1 : p \neq \frac{1}{2}$

Para un nivel de significación (o riesgo) α, se acepta H_0 :

$$\chi^2_{(k-1)} = \overbrace{\sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i}}^{\text{estadístico observado}} < \overbrace{\chi^2_{\alpha,(k-1)}}^{\text{estadístico teórico}} \qquad \sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i} = \sum_{i=1}^{k} \frac{n_i^2}{e_i} - n \qquad e_i = n\,p_i$$

Lanzamiento de tres monedas: $\Omega = \{(c,c,c),(c,c,e),(c,e,c),(e,c,c),(c,e,e),(e,c,e),(e,e,c),(e,e,e)\}$

$$P(X=0) = \frac{1}{8} \qquad P(X=1) = \frac{3}{8} \qquad P(X=3) = \frac{3}{8} \qquad P(X=3) = \frac{1}{8}$$

Frecuencias esperada: $e_i = n\,p_i = 80\,P(X = x_i)$

$X = x_i$	0	1	2	3
n_i	7	24	35	14
$e_i = np_i$	10	30	30	10

Estadístico observado:

$$\chi^2_3 = \sum_{i=1}^{4} \frac{n_i^2}{e_i} - n = \sum_{i=1}^{4} \frac{n_i^2}{np_i} - n = \frac{7^2}{10} + \frac{24^2}{30} + \frac{35^2}{30} + \frac{14^2}{10} - 80 = 4,5333$$

Estadístico teórico: $\chi^2_{0,05\,,\,3} = 7,81$

El estadístico de contraste (bondad de ajuste) es menor que el estadístico teórico, con un nivel de significación 0,05 , $(\chi^2_3 = 5,4533 < 7,18 = \chi^2_{0,05\,,\,3})$, concluyendo que la moneda está equilibrada.

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

$$f(x) = \begin{cases} \dfrac{1}{a}\, e^{-\frac{x}{a}} & \text{si } x > 0,\ a > 0 \\ 0 & \text{en el resto} \end{cases}$$

Para una muestra de tamaño n, calcular:

a) Estimador máximo verosímil de a.

b) Estimador de a por el método de los momentos.

c) ¿Son eficientes?

Solución

a) Función de verosimilitud:

$$L(X,a) = L(x_1, x_2, \cdots, x_n; a) = \left(\frac{1}{a} e^{-\frac{x_1}{a}}\right) . \left(\frac{1}{a} e^{-\frac{x_2}{a}}\right) \cdots \left(\frac{1}{a} e^{-\frac{x_n}{a}}\right) = \frac{1}{a^n} e^{-\frac{1}{a}\sum_{i=1}^{n} x_i}$$

Función soporte o Log-verosimilitud: $\ln L(X,a) = \ln\left(\dfrac{1}{a^n} e^{-\frac{1}{a}\sum_{i=1}^{n} x_i}\right) = -n \ln a - \dfrac{1}{a}\displaystyle\sum_{i=1}^{n} x_i$

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:

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

b) La expresión de la función de densidad desvela que se trata de una variable aleatoria exponencial, es decir, $X \sim Exp\left(\lambda = \dfrac{1}{a}\right)$ siendo, por tanto, $E(X) = \dfrac{1}{1/a} = a$, $Var(X) = \dfrac{1}{(1/a)^2} = a^2$

Para obtener el estimador de "a" por el método de los momentos se necesita igualar los momentos muestrales a los momentos poblacionales.

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

Se comprueba que: $E[X] = \displaystyle\int_0^\infty x f(x) dx = \underbrace{\int_0^\infty x \frac{1}{a} e^{-\frac{x}{a}} dx}^{\text{integración por partes}} = \left. -\frac{(x+a)}{e^{\frac{x}{a}}}\right|_0^\infty = a$

$$u = x \qquad du = dx$$

$$dv = \frac{1}{a} e^{-\frac{x}{a}} dx \qquad v = \int \frac{1}{a} e^{-\frac{x}{a}} dx = -e^{-\frac{x}{a}}$$

$$\int x \frac{1}{a} e^{-\frac{x}{a}} dx = -x e^{-\frac{x}{a}} + \int e^{-\frac{x}{a}} dx = -x e^{-\frac{x}{a}} - a e^{-\frac{x}{a}} = -(x+a)\, e^{-\frac{x}{a}}$$

c) Se trata de comprobar que la media muestral es un estimador eficiente de "a"

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).

Insesgadez: $$E(\overline{x}) = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\, n\, E(x_i) = E(x_i) = a$$

La media muestal es un estimador insesgado o centrado de "a"

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

$$CCR = \frac{1}{I(a)} \qquad I(a) = -\,E\left[\frac{\vartheta^2 \ln L(x\,;\,a)}{\vartheta a^2}\right] \quad \text{Información de Fisher}$$

$$\frac{\vartheta \ln L(X,a)}{\vartheta a} = -\frac{n}{a} + \frac{1}{a^2}\sum_{i=1}^{n} x_i \quad \to \quad \frac{\vartheta^2 \ln L(X,a)}{\vartheta a^2} = \frac{n}{a^2} - \frac{2}{a^3}\sum_{i=1}^{n} x_i$$

$$I_n(a) = -E\left[\frac{\vartheta^2 \ln L(X,a)}{\vartheta a^2}\right] = -E\left[\frac{n}{a^2} - \frac{2}{a^3}\sum_{i=1}^{n} x_i\right] = -\frac{n}{a^2} + \frac{2}{a^3}\sum_{i=1}^{n} E(x_i) = -\frac{n}{a^2} + \frac{2\,n\,a}{a^3} = \frac{n}{a^2}$$

$$CCR = \frac{1}{I(a)} = \frac{1}{\dfrac{n}{a^2}} = \frac{a^2}{n}$$

Por otra parte, la varianza de la media muestral:

$$Var(\overline{x}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i) = \frac{1}{n^2}\, n\, Var(x_i) = \frac{1}{n^2}\, n\, \frac{1}{(1/a)^2} = \frac{a^2}{n}$$

Se ha probado que la media muestral es un estimador eficiente de "a".

Sea $X_1, \cdots, X_n$ una muestra aleatoria simple de la distribución

$f_\theta(x) = \theta e^{-\theta x} \quad , \quad x > 0$

a) Calcular el estimador de máxima verosimilitud para $P(X > t)$

b) Supongamos que se miden 10 componentes cuya duración sigue la distribución anterior y que si se suma el tiempo de funcionamiento de todas ellas se contabiliza un total de 1200 horas. A partir de esta muestra, estimar la probabilidad de que una componente vaya a durar menos de 130 horas.

c) Probar si el estadístico $T(X_1, \cdots, X_n) = \dfrac{n-1}{\sum_{i=1}^{n} X_i}$ es centrado para θ

d) Probar si el estadístico $\sum_{i=1}^{n} X_i$ es estadístico suficiente y completo. ¿Es $T(X_1, \cdots, X_n)$ estimador centrado uniformemente de mínima varianza?.

Solución

a) $F_\theta(x) = \int_0^x f_\theta(t)\,dt = \int_0^x \theta e^{-\theta t}\,dt = -e^{-\theta t}\Big|_0^x = 1 - e^{-\theta x}$

$P(X > t) = 1 - P(X \le t) = 1 - F(t) = 1 - (1 - e^{-\theta t}) = e^{-\theta t}$

En consecuencia, hay que encontrar el estimador de máxima verosimilitud de $g(\theta) = e^{-\theta t}$

Función de verosimilitud:

$$L(\theta) = f_\theta(x_1) . f_\theta(x_2) \ldots f_\theta(x_n) = (\theta e^{-\theta x_1})\,(\theta e^{-\theta x_2}) \ldots (\theta e^{-\theta x_n}) = \theta^n\, e^{-\theta \sum_{i=1}^{n} x_i} \qquad \forall x_i > 0$$

Función soporte o Log-verosimilitud: $\ln L(\theta) = \ln\left[\theta^n\, e^{-\theta \sum_{i=1}^{n} x_i}\right] = n \ln\theta - \theta \sum_{i=1}^{n} x_i$

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

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

Con lo que el estimador de máxima verosimilitud de θ es el inverso de la media muestral.

Considerando que el estimador de máxima verosimilitud es invariante ante transformaciones biunívocas y la función $e^{-\theta t}$ es decreciente en θ, la estimación máximo verosímil de $g(\theta) = e^{-\theta t}$ es $\widehat{g(\theta)} = e^{-\frac{t}{\overline{x}}}$

b) Se tiene: $n=10$, $\bar{x} = \dfrac{1}{10}\sum_{i=1}^{10} x_i = \dfrac{1}{10}1200 = 120$, $P(X>t) = 1 - P(X\le t) = e^{-\frac{t}{\bar{x}}}$

$$P(X<130) = 1 - P(X\ge 130) = 1 - P(X>130) = 1 - e^{-\frac{130}{\bar{x}}} = 1 - e^{-\frac{130}{120}} = 0{,}6615$$

c) Insesgadez del estadístico T :

$$E(T) = E\left[(n-1)/\sum_{i=1}^{n} X_i\right] = (n-1)\,E\left[\frac{1}{Y}\right] \qquad Y = \sum_{i=1}^{n} X_i \quad , \quad X_i \sim Exp(\theta) \equiv \Gamma(1,\theta) \quad \to \quad Y\sim\Gamma(n,\theta)$$

Momentos de la variable $Y\sim\Gamma(n,\theta)$: $E(Y) = \dfrac{n}{\theta}$ $\quad V(Y) = \dfrac{n}{\theta^2}$ $\quad E\left[\dfrac{1}{Y}\right] = \dfrac{\theta}{n-1}$

$$E(T) = E\left[(n-1)/\sum_{i=1}^{n} X_i\right] = E\left[\frac{n-1}{Y}\right] = (n-1)\,E\left[\frac{1}{Y}\right] = (n-1)\,\frac{\theta}{n-1} = \theta$$

El estadístico T es centrado o insesgado para θ.

d) El teorema de factorización de Fisher-Neyman identifica estadísticos suficientes. Establece que dada una muestra aleatoria $(x_1, \cdots, x_n)$ de una población X con función de densidad f_θ, un estadístico $\hat{\theta}$ es suficiente para θ sí y sólo sí:

$$f_\theta(x_1, \cdots, x_n) = g\left[\hat{\theta}(x_1, \cdots, x_n), \theta\right].\ h(x_1, \cdots, x_n)$$

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

$$L(\theta) = g(\hat{\theta},\theta)\,.\,h(x_1, \cdots, x_n)$$

Función verosimilitud: $L(\theta) = f_\theta(x_1)\,.\,f_\theta(x_2) \ldots f_\theta(x_n) = (\theta e^{-\theta x_1})\,(\theta e^{-\theta x_2}) \ \ldots\ (\theta e^{-\theta x_n}) = \theta^n e^{-\theta\sum_{i=1}^{n} x_i}$

$h(x_1, \cdots, x_n) = 1$

La Función verosimilitud $L(\theta)$ depende de la muestra sólo a través del estadístico $\sum_{i=1}^{n} x_i$ y que

$h(x_1, \cdots, x_n)$ no depende del parámetro, quedando demostrada la suficiencia del estadístico.

El estadístico $\sum_{i=1}^{n} x_i$ es completo sí $E_\theta\left[g\left(\sum_{i=1}^{n} x_i\right)\right] = 0 \ \Rightarrow\ g\left(\sum_{i=1}^{n} x_i\right) = 0$

Como $\sum_{i=1}^{n} x_i \sim \Gamma(n,\theta)$ $\to$ $E_\theta\left[g\left(\sum_{i=1}^{n} x_i\right)\right] = \int_0^\infty f(y)\,\dfrac{\theta^n}{(n-1)!}\,y^{n-1}\,e^{-\theta y}\,dy = 0 \quad \forall\theta>0$

La transformada de Laplace $\int_0^\infty f(y)\,y^{n-1}\,e^{-\theta y}\,dy$ de la función $f(y)\,y^{n-1} = 0$ y, por el teorema de unicidad de la transformada de Laplace, $f(y)\,y^{n-1} = 0$ o bien $f(y) = 0$

El estadístico $\sum_{i=1}^{n} x_i$ es suficiente y completo.

Por el teorema de Lehmann - Scheffé, siendo el estadístico T centrado y función del estadístico suficiente y completo $\sum_{i=1}^{n} x_i$.

T es un estadístico ECUMV (estadístico centrado uniforme de mínima varianza).

El tiempo de vida de un tipo de bombillas se sabe que es una variable aleatoria que sigue una distribución exponencial de parámetro $\dfrac{1}{\theta}$ con $\theta > 0$

Para estimar el tiempo medio de vida se extrae una muestra aleatoria simple de n bombillas y se consideran dos estimadores. Uno es la media muestral, $T_1(X_1, \cdots, X_n) = \dfrac{1}{n}\sum_{i=1}^{n} X_i$, y el otro se define como $T_2(X_1, \cdots, X_n) = \text{mín}(X_1, \cdots, X_n)$. Se pide:

a) Calcular la media, el sesgo y la varianza del estimador T_1.

b) Calcular la media, el sesgo y la varianza del estimador T_2.

c) ¿Es T_2 insesgado?. Si no lo es, define un estimador insesgado T_3 que sea función de él.

Solución:

a) $x_i \sim \text{Exp}\left(\dfrac{1}{\theta}\right)$ con lo que $E(x_i) = \theta$, $\text{Var}(x_i) = \theta^2$, $E\left[x_i^2\right] = 2!\,\theta^2 = 2\theta^2$

$$X \sim \text{Exp}(\lambda): \quad f(x) = \lambda e^{-\lambda x} \quad F(x) = 1 - e^{-\lambda x} \quad x \geq 0$$

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}, \quad E(X^n) = \frac{n!}{\lambda^n}$$

$$E[T_1] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \frac{1}{n}\, n\,\theta = \theta$$

Es un estimador centrado o insesgado: $b[T_1] = \text{Sesgo}[T_1] = 0$

$$\text{Var}[T_1] = \text{Var}\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}[x_i] = \frac{1}{n^2}\, n\,\theta^2 = \frac{\theta^2}{n}$$

b) Para el cálculo de la media, sesgo u varianza del estimador $T_2(X_1, \cdots, X_n) = \text{mín}(X_1, \cdots, X_n)$ lo primero que hay que obtener es su función de densidad:

$$X \sim \text{Exp}\left[\frac{1}{\theta}\right]: \quad F(x) = 1 - e^{-\frac{x}{\theta}} \quad \rightarrow \quad f(x) = \frac{dF(x)}{dx} = \frac{1}{\theta}\, e^{-\frac{x}{\theta}} \quad x > 0$$

Sea $X^{\bullet} = \text{mín}(X_1, \cdots, X_n) = T_2(X_1, \cdots, X_n)$

$$F_{X^{\bullet}} = P(X^{\bullet} \leq x) = 1 - P(X^{\bullet} \geq x) = 1 - P(X_1 \geq x, X_2 \geq x, \cdots, X_n \geq x) = 1 - \prod_{i=1}^{n} P(X_i \geq x) =$$

$$= 1 - \prod_{i=1}^{n}\left[1 - P(X_i < x)\right] = 1 - \left[1 - F(x)\right]^n$$

$$f_{T_2}(x) = \frac{dF_{X_\bullet}}{dx} = n\, f(x)\left[1 - F(x)\right]^{(n-1)} = n\,\frac{1}{\theta}\,e^{-\frac{x}{\theta}}\left[1 - \left(1 - e^{-\frac{x}{\theta}}\right)\right]^{(n-1)} = \frac{n}{\theta}\,e^{-\frac{x}{\theta}}\,e^{-\frac{x(n-1)}{\theta}} = \frac{n}{\theta}\,e^{-\frac{nx}{\theta}}$$

$$f_{T_2}(x) = \begin{cases} \dfrac{n}{\theta}\,e^{-\frac{nx}{\theta}} & x > 0 \\ 0 & \text{en el resto} \end{cases} \qquad T_2(X_1, \cdots, X_n) \sim \text{Exp}\left[\frac{n}{\theta}\right]$$

$$E(T_2) = \frac{1}{\frac{n}{\theta}} = \frac{\theta}{n} \quad \rightarrow \quad \text{Sesgo: } b(T_2) = \frac{\theta}{n} - \theta = \left(\frac{1-n}{n}\right)\theta$$

$$Var(T_2) = \frac{1}{\left(\frac{n}{\theta}\right)^2} = \frac{\theta^2}{n^2}$$

c) El estimador T_2 es sesgado: $E(T_2) = \dfrac{\theta}{n}$

Considerando el estimador $T_3 = n\,T_2$ se obtiene un estimador insesgado:

$$E(T_3) = E(nT_2) = n\,E(T_2) = n\,\frac{\theta}{n} = \theta$$

$$Var(T_3) = Var(nT_2) = n^2\,Var(T_2) = n^2\,\frac{\theta^2}{n^2} = \theta^2$$

El estimador T_1 es más eficiente que el estimador T_3 por no tener menor varianza cuando $n > 1$:

$$Var\left[T_1\right] = \frac{\theta^2}{n} < Var\left[T_3\right] = \theta^2$$

En caso que $n = 1$ son igual de eficientes.

El tiempo de realización en minutos de una determinada tarea dentro de un proceso industrial es una variable aleatoria con función de densidad

$$f_\theta(x) = \frac{x}{\theta^2} e^{-x/\theta} \;,\quad x > 0 \;\text{ donde }\; \theta > 0$$

a) Calcular el estimador máximo-verosímil de θ para una muestra aleatoria simple de tamaño n.

b) Calcular el estimador máximo-verosímil de $E[X]$ para una muestra aleatoria simple de tamaño n.

c) Mediante un muestreo aleatorio simple se han recogido los siguientes 15 tiempos de realización de la tarea:

5,56	2,23	0,58	1,37	0,21	1,98	2,44	2,71
10,12	4,69	3,47	1,73	3,51	1,19	0,97	

Obtener la estimación máximo-verosímil del tiempo medio de realización del proceso.

d) Para la muestra del apartado anterior, dar una aproximación de la varianza asintótica del estimador de máxima verosimilitud de θ.

Solución:

a) Función de verosimilitud:

$$L(\theta) = f_\theta(x_1)\,.\,f_\theta(x_2)\,\ldots\,f_\theta(x_n) = \left(\frac{x_1}{\theta^2} e^{-x_1/\theta}\right)\left(\frac{x_2}{\theta^2} e^{-x_2/\theta}\right)\ldots\left(\frac{x_n}{\theta^2} e^{-x_n/\theta}\right) =$$

$$= \frac{1}{\theta^{2n}} \prod_{i=1}^{n} x_i \; e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i} \qquad \forall x_i > 0 \quad \theta > 0$$

Función soporte o Log-verosimilitud:

$$\ln L(\theta) = \ln\left[\frac{1}{\theta^{2n}} \prod_{i=1}^{n} x_i \; e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}\right] = -2n\ln\theta + \sum_{i=1}^{n}\ln(x_i) - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

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

$$\left.\frac{\vartheta \ln L(X,\theta)}{\vartheta\theta}\right|_{\theta=\hat{\theta}} = \left. -\frac{2n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i \right|_{\theta=\hat{\theta}} = 0 \;\rightarrow\; \hat{\theta} = \frac{1}{2n}\sum_{i=1}^{n} x_i = \frac{\overline{x}}{2}$$

Se comprueba que es un máximo: $\left.\dfrac{\vartheta^2 \ln L(X,\theta)}{\vartheta\theta^2}\right|_{\theta=\hat{\theta}} < 0$

$$\left.\frac{\vartheta^2 \ln L(X,\theta)}{\vartheta\theta^2}\right|_{\theta=\hat{\theta}} = \left.\frac{\vartheta\left(\frac{2n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i\right)}{\vartheta\theta}\right|_{\theta=\hat{\theta}} = -\frac{2n}{\theta^2} - \left.\frac{2}{\theta^3}\sum_{i=1}^{n} x_i\right|_{\theta=\hat{\theta}} \;\rightarrow\; \hat{\theta} = -\frac{\sum_{i=1}^{n} x_i}{n} = -\overline{x} < 0$$

$$\text{EMV}(\theta) = \frac{\overline{x}}{2}$$

b) Siendo X una variable aleatoria continua: $E(X) = \int_0^\infty x\, f(x)\, dx = \int_0^\infty x\, \frac{x}{\theta^2}\, e^{-x/\theta}\, dx$

$$E(X) = \int_0^\infty \frac{x^2}{\theta^2}\, e^{-x/\theta}\, dx = -\frac{1}{\theta}\int_0^\infty \underbrace{x^2}_{u}\, \underbrace{\left(-\frac{1}{\theta}\right) e^{-x/\theta}\, dx}_{dv} = x^2 e^{-x/\theta}\Big|_0^\infty - \int_0^\infty \underbrace{2x}_{u}\, \underbrace{\left(-\frac{1}{\theta}\right) e^{-x/\theta}\, dx}_{dv} =$$

$$= x^2 e^{-x/\theta}\Big|_0^\infty - \left(2x\, e^{-x/\theta}\Big|_0^\infty - \int_0^\infty 2 e^{-x/\theta}\, dx\right) =$$

$$= x^2 e^{-x/\theta}\Big|_0^\infty - \left(2x\, e^{-x/\theta}\Big|_0^\infty + 2\theta \int_0^\infty \left(-\frac{1}{\theta}\right) e^{-x/\theta}\, dx\right) =$$

$$= x^2 e^{-x/\theta}\Big|_0^\infty - \left(2x\, e^{-x/\theta}\Big|_0^\infty + 2\theta\, e^{-x/\theta}\Big|_0^\infty\right) = (x^2 - 2x - 2\theta)\, e^{-x/\theta}\Big|_0^\infty = 2\theta$$

$$\text{EMV}\big[E(X)\big] = 2\theta = 2\,\frac{\overline{x}}{2} = \overline{x}$$

c) $\overline{x} = \frac{1}{15}\sum_{i=1}^{15} x_i = \frac{42{,}76}{15} = 2{,}8507$ minutos

$$\text{EMV}\left[\widehat{E(X)}\right] = \overline{x} = 2{,}8507 \text{ minutos}$$

d) Un estimador de máxima versoimilitud (EMV) es asintóticamente normal y su varianza asintótica verifica la cota de Cramér-Rao:

$$\text{CCR} = \frac{1}{-n\,.\,E\left(\dfrac{\vartheta^2 \ln P(x\,;\,p)}{\vartheta\theta^2}\right)} \qquad \hat{\theta} = \frac{1}{2n}\sum_{i=1}^{n} x_i \qquad I_1(\theta) = -E\left[\frac{\vartheta^2 \ln P(x\,;\,p)}{\vartheta\theta^2}\right] \text{ Información de Fisher}$$

$$\frac{\vartheta^2 \ln L(X,\theta)}{\vartheta\theta^2} = -\frac{2n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} x_i = -\frac{2}{\theta^2}\left(n - \frac{1}{\theta}\sum_{i=1}^{n} x_i\right) = -\frac{2}{\theta^2}\left(n - \frac{2n}{\sum_{i=1}^{n} x_i}\sum_{i=1}^{n} x_i\right) = -\frac{2n}{\theta^2}$$

$$\text{CCR} = \frac{1}{-n\,.\,E\left(\dfrac{\vartheta^2 \ln P(x\,;\,p)}{\vartheta\theta^2}\right)} = \frac{1}{-n\,.\,E\left(\dfrac{-2n}{\theta^2}\right)} = \frac{\theta^2}{2n^2} = \frac{(\overline{x}/2)^2}{2n^2}$$

$$\text{CCR} = \frac{(2{,}8507/2)^2}{2 \times 15} = 0{,}0677$$

ESTIMADORES PUNTUALES: SUFICIENCIA. VEROSIMILITUD. MÉTODO DE LOS MOMENTOS

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

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

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.

El Teorema de Factorización de Fisher-Neyman identifica estadísticos suficientes establece que dada una muestra aleatoria $(x_1, \cdots, x_n)$ de una población X con función de densidad f_θ (caso continuo), un estadístico $\hat{\theta}$ es suficiente para θ sí y sólo sí:

$$f_\theta(x_1, \cdots, x_n) = g\left[\hat{\theta}(x_1, \cdots, x_n), \theta\right] . h(x_1, \cdots, x_n)$$

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

$$L(\theta) = g(\hat{\theta}, \theta) . h(x_1, \cdots, x_n)$$

Función de verosimilitud:

$$L(\theta) = f_\theta(x_1) . f_\theta(x_2) \ldots f_\theta(x_n) = (\theta x_1^{\theta-1})(\theta x_2^{\theta-1}) \ldots (\theta x_n^{\theta-1}) = \theta^n (x_1 \cdots x_n)^{\theta-1}$$

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

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

Función soporte o Log-verosimilitud:

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

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

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

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

Los taxis en servicio de una ciudad están numerados del 1 al N y se desea conocer cuántos taxis hay.

Para ello se observa una muestra de n taxis y se apuntan sus números. Se pide:

a) Obtener un estimador por el método de los momentos. ¿Es insesgado?

b) Obtener un estimador por el método de máxima verosimilitud.

Solución:

a) En la variable aleatoria X = "Número de taxis", todos los taxis tienen igual probabilidad.

Función de densidad: $f(k) = P(X=k) = \dfrac{1}{N} \qquad k = 1, 2, \ldots, N$

$$\text{Esperanza: } E(X) = \sum_{i=1}^{N} x_i\, P(X=x_i) = \sum_{i=1}^{N} i \times \frac{1}{N} = \frac{1}{N}(1+2+\ldots+N) = \frac{1}{N}\,\frac{(1+N)\,N}{2} = \frac{1+N}{2}$$

$$\text{Aplicando el méodo de los momentos: } E(X) = \frac{1+N}{2} = \bar{x} \;\rightarrow\; \hat{N} = 2\bar{x} - 1$$

$$\text{Insesgadez: } E(\hat{N}) = E(2\bar{x} - 1) = \frac{2}{N}\sum_{i=1}^{N} E(x_i) - 1 = \frac{2}{N}\sum_{i=1}^{N} E(x_i) - 1 = \frac{2}{N}\,N\,\frac{1+N}{2} - 1 = N$$

Luego es un estimador insesgado.

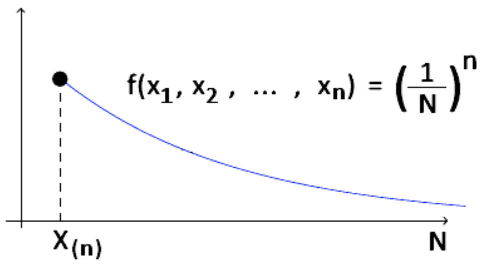
b) No se puede aplicar el método de máxima verosimilitud de forma habitual al no verificarse las condiciones regulares, dado que el soporte de la función depende del parámetro.

$$f(x_1, x_2, \cdots, x_n) = \left[\frac{1}{N}\right]\left[\frac{1}{N}\right]\cdots\left[\frac{1}{N}\right] = \left(\frac{1}{N}\right)^n \times I_{X_{(n)} < N}$$

$$\text{siendo: } X_{(n)} = \max_{1 \le i \le n} X_i \;, \quad I_{X_{(n)} < N} = \begin{cases} 1 & X_{(n)} < N \\ 0 & \text{resto} \end{cases}$$

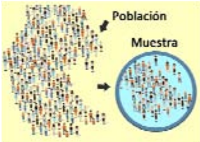
La función se hace máxima cuando N es lo menor posible, pero N siempre es mayor que $X_{(n)}$.

Por tanto, $EMV(\hat{N}) = X_{(n)}$



Universidad Autónoma de Madrid

ESTIMADORES PUNTUALES: MÉTODO DE LOS MOMENTOS. INSESGADEZ



Una muestra aleatoria $(x_1, \cdots, x_n)$ 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: $E(X) = \overline{x}$

$$\overline{x} = E[X] = \int_\theta^\infty x f_\theta(x)\,dx = \overbrace{\int_\theta^\infty x\, e^{-x+\theta}\,dx}^{\text{integración por partes}} = 1 + \theta \;\to\; \hat{\theta} = \overline{x} - 1$$

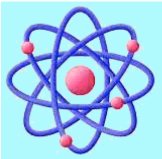
$$\int \underbrace{x}_{u}\, \underbrace{e^{-x+\theta}dx}_{dv} = \underbrace{x}_{u}\,\underbrace{(-e^{-x+\theta})}_{v} - \int \underbrace{-e^{-x+\theta}}_{v}\,\underbrace{dx}_{du} = -x e^{-x+\theta} - e^{-x+\theta} = -(1+x)\,e^{-x+\theta} = -e^{-\theta}\left(\frac{1+x}{e^{x}}\right)$$

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

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

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

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:

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

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

a) Obtener el estimador θ por el método de los momentos

b) Calcular la varianza de este estimador y demostrar que es consistente

Solución

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

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

b) $V(\hat{\theta}) = V(3\,\overline{x}) = 9\,V(\overline{x}) = 9\,\dfrac{V(X)}{n} = \dfrac{9}{n}\,V(X)$

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

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

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

$$\lim_{n\to\infty} E(\hat{\theta}) = \lim_{n\to\infty} E(3\,\overline{x}) = \lim_{n\to\infty} 3E(\overline{x}) = 3E(X) = 3\,\frac{\theta}{3} = \theta$$

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

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

Universidad Autónoma de Madrid

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