Inferencia Estadística

Marisol García Peña

Departamento de Matemáticas
Pontificia Universidad Javeriana

Bogotá, 2022

Ejemplo

Sea $X_1, \ldots, X_n$ una m.a. de alguna distribución tal que $E[X_i] = \mu$ y $V[X_i] = \sigma^2$, $i = 1, \ldots, n$. Considere los estimadores de μ, $T_1 = \overline{X}$ y $T_2 = \sum_{i=1}^{n} \frac{X_i}{n+1}$.

Obetener los ECM de T_1 y T_2 y demuestre que $ECM[T_2] < ECM[T_1]$ para algunos valores de μ mientras que la proposición inversa es cierta para otros valores de μ.

Estimación puntual (cont.)

Insesgadez/Insesgamiento

Un estimador $T(X_1, \ldots, X_n)$ es definido como estimador insesgado de $\tau(\theta)$ si y sólo si

$$E_\theta[\widehat{\theta}] = E_\theta[T(X_1, \ldots, X_n)] = \tau(\theta)$$

Un estimador es insegado si su valor esperado es igual a $\tau(\theta)$.

Ejemplo

Sea $X_1, \ldots, X_n$ una m.a. de una distribución $N(\mu, \sigma^2)$.

- Muestre que $\overline{X} = \sum_{i=1}^{n} \frac{x_i}{n}$ es un estimador insesgado de μ.
- Muestre que $S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$ es un estimador insesgado de σ^2.
- Muestre que $\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$ es un estimador sesgado de σ^2.

Estimación puntual (cont.)

Ejemplo

Sea $X_1, \ldots, X_n$ una m.a. i.i.d $U(0, \beta)$. El EMV de β es $\widehat{\beta}_{MV} = X_{(n)} = $ máx. ¿$\widehat{\beta}$ es un estimador insesgado?

La f.d.p de $\widehat{\beta}$ es $f_{\text{máx}}(x) = \frac{n}{\beta^n} x^{n-1}$. Su valor esperado es

$$E[\widehat{\beta}] = E[X_{(n)}] = \int_0^{\beta} x \frac{n}{\beta^n} x^{n-1} dx = \frac{n}{n+1} \beta$$

Entonces $\widehat{\beta}_{MV} = X_{(n)}$ es un estimador sesgado de β. Pero $\frac{n+1}{n} X_{(n)}$ es insegado.

Si se usa el estimador del método de momentos $\widehat{\beta} = 2\overline{X}$. Su valor esperado es $E[\widehat{\beta}] = E[2\overline{X}] = 2\frac{\beta}{2} = \beta$. Esto indica que el estimador de β por el método de momentos es insesgado.

Estimación puntual (cont.)

- Si $T(X_1, \ldots, X_n)$ es un estimador insesgado para $\tau(\theta) \Longrightarrow h(T)$ en general no es un estimador insesgado de $h(\tau(\theta))$.
- El estimador insesgado que minimiza el ECM se conoce como **MVUE-Minimum Variance Unbiased Estimator** de $\tau(\theta)$.

Estimador asintóticamente insesgado

Sea $X_1, \ldots, X_n$ una m.a. con función de densidad, y sea $T(X_1, \ldots, X_n)$ un estimador puntual de $\tau(\theta)$. Se dice que T es un estimador asintóticamente insesgado si

$$\lim_{n \to \infty} E[T(X_1, \ldots, X_n)] = \tau(\theta)$$

Estimación puntual (cont.)

Ejemplo

En el caso del estimador de β, $X_{(n)}$

$$\lim_{n\to\infty} E[X_{(n)}] = \lim_{n\to\infty} \frac{n}{n+1}\beta = \beta.$$

Para el estimador de σ^2 de una $N(\mu, \sigma^2)$, se tiene

$$\lim_{n\to\infty} E[\widehat{\sigma}^2] = \lim_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2.$$

Estos estimadores son asintóticamente insesgados.

Si se tienen 2 estimadores insesgados de $\tau(\theta)$. ¿Cuál se debe seleccionar?

Eficiencia

Sean $T_1(X_1, \ldots, X_n)$ y $T_2(X_1, \ldots, X_n)$ dos estimadores insesgados de $\tau(\theta)$. Se dice que T_1 es **más eficiente** que T_2 si se verifica que

$$V[T_1(X_1, \ldots, X_n)] < V[T_2(X_1, \ldots, X_n)]$$

Es decir, un estimador más eficiente que otro tiene menor varianza/dispersión.

Estimación puntual (cont.)

Eficiencia relativa

Sean $T_1(X_1, \ldots, X_n)$ y $T_2(X_1, \ldots, X_n)$ dos estimadores insesgados de $\tau(\theta)$. Se define la **eficiencia relativa de** T_1 **respecto a** T_2 como

$$E.R.[T_1, T_2] = \frac{V[T_2]}{V[T_1]}$$

Si $E.R.[T_1, T_2] > 1 \Longrightarrow T_1$ es más eficiente que T_2, caso contrario T_2 será más eficiente que T_1.

Ejemplo

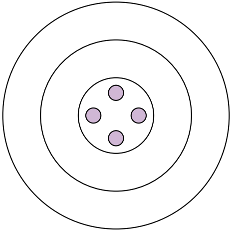
Sean $X_1, \ldots, X_{n_1}$ y $Y_1, \ldots, Y_{n_2}$ dos muestras aleatorias independientes de una población $N(\mu, \sigma^2)$ con medias $\overline{X}_1$ y $\overline{X}_2$ respectivamente.

Un investigador pretende estimar la media poblacional μ y propone 2 estimadores alternativos

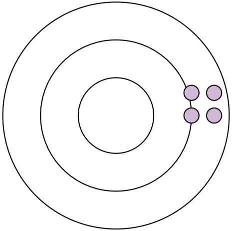
$$\widehat{\mu} = \frac{1}{2}(\overline{X}_1 + \overline{X}_2) \quad \text{y } \widetilde{\mu} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$$

¿Cuál de esos estimadores es más adecuado para estimar μ?

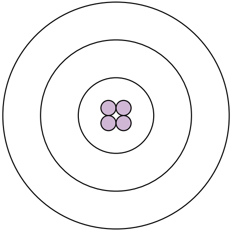
Estimación puntual (cont.)



Insesgado



Eficiente



Insesgado y Eficiente

Consistencia y BAN

- *ECM* y la propiedad de insesgamiento $\Longrightarrow$ n fijo.
- 2 conceptos $\Longrightarrow$ aumento de n.
- Notación, $T(X_1, \ldots, X_n) \Longrightarrow T_n(X_1, \ldots, X_n)$.
- Secuencia de estimadores, $T_1(X_1), T_2(X_1, X_2), \ldots, T_n(X_1, \ldots, X_n)$.
- Buena secuencia de estimadores $\Longrightarrow$ Los valores de los estimadores $\Longrightarrow$ acerca a la cantidad estimada cuando $n \Uparrow$

Consistencia de error cuadrático medio

Sea $T_1, T_2, \ldots, T_n$ una secuencia de estimadores de $\tau(\theta)$, donde $T_n(X_1, \ldots, X_n)$ está basado en una muestra de tamaño n.

La secuencia de estimadores se define como secuencia de estimadores de $\tau(\theta)$ *consistente de error cuadrático medio*, si y sólo si

$$\lim_{n\to\infty} E[[T_n - \tau(\theta)]^2] = 0, \forall\theta$$

Consistencia de error cuadrático medio $\Longrightarrow$ sesgo y varianza de T_n se aproximan a 0 ya que $E[[T_n - \tau(\theta)]^2] = V[T_n] + \{\tau(\theta) - E[T_n]\}^2$.

Ejemplo

En un muestreo de una densidad con media μ y varianza σ^2, sea $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ una secuencia de estimadores de μ y $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ una secuencia de estimadores de σ^2.

$E[(\overline{X}_n - \mu)^2] = V[\overline{X}] = \frac{\sigma^2}{n} \longrightarrow 0$ cuando $n \longrightarrow \infty$; en este caso, la secuencia $\{\overline{X}_n\}$ es una secuencia de estimadores de μ consistente de ECM.

$E[(S_n^2 - \sigma^2)^2] = V[S_n^2] = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right) \longrightarrow 0$ cuando $n \longrightarrow \infty$, la secuencia $\{S_n^2\}$ es una secuencia de estimadores de σ^2 consistente de ECM.

Si $T_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$, la secuencia $\{T_n\}$ también es una secuencia de estimadores de σ^2 consistente de ECM.

Estimación puntual (cont.)

Consistencia simple

Sea $T_1, \ldots, T_n$ una secuencia de estimadores de $\tau(\theta)$, donde $T_n(X_1, \ldots, X_n)$. La secuencia $\{T_n\}$ es definida como una secuencia de estimadores de $\tau(\theta)$ *consistente simple (débil)* si para cada $\varepsilon > 0$ se satisface:

$$\lim_{n\to\infty} P[|T_n - \tau(\theta)| \leq \varepsilon] = 1; \forall\theta, \forall\varepsilon > 0$$

o equivalentemente,

$$\lim_{n\to\infty} P[|T_n - \tau(\theta)| > \varepsilon] = 0; \forall\theta, \forall\varepsilon > 0$$

Convergencia en probabilidad.

Para una cantidad aceptable de error $\varepsilon > 0$, la probabilidad de que el error real empeore más que ε tiende a cero.

Estimación puntual (cont.)

- La consistencia de un estimador garantiza que si se tiene un tamaño de muestra grande cualquier estimación particular va a estar cerca del valor de $\tau(\theta)$.

- Si un estimador es un estimador consistente de ECM también es un estimador consistente simple, lo contrario no necesariamente es cierto.

Ejemplo

Sea $X_1, \ldots, X_n$ denota una muestra aleatoria de $N(\mu, 1)$ y sea $\overline{X}_n$ la media muestral. Muestre que $\overline{X}_n$ es un estimador consistente para μ.

Ejemplo

Sea $X_1, \ldots, X_n$ denota una muestra aleatoria de una distribución Bernoulli $(1, \theta)$. Muestre que $\overline{X}$ es un estimador consistente para θ, $0 < \theta < 1$.

Estimación puntual (cont.)

Ejemplo

Sea $X_1, \ldots, X_n$ denota una muestra aleatoria de $N(\mu, \sigma^2)$ y sea S_n^2 la varianza muestral. Muestre que S_n^2 es un estimador consistente para σ^2.

Consistencia

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una f.d.p. Sea $\{T_n\}$ una secuencia de estimadores de $\tau(\theta)$, tal que

$$\lim_{n\to\infty} E[T(X_1, \ldots, X_n)] = \tau(\theta)$$
$$\lim_{n\to\infty} V[T(X_1, \ldots, X_n)] = 0$$

Entonces T_n es consistente para $\tau(\theta)$.

Estimación puntual (cont.)

Recordando

Por el TCL. Sea $\{X_n\}$ una secuencia de v.a's iid con $E[X_i] = \mu$ y $V[X_i] = \sigma^2$, $0 < \sigma^2 < \infty$ entonces

$$\frac{T_n - E[T_n]}{S_n} \xrightarrow[n\to\infty]{D} N(0,1)$$

donde $T_n = \sum_{i=1}^n X_i, S_n^2 = V[T_n]$

$$\begin{aligned} E[T_n] &= E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i] = n\mu \\ V[T_n] &= V[\sum_{i=1}^n X_i] = \sum_{i=1}^n V[X_i] = n\sigma^2 \end{aligned}$$

De lo anterior, tenemos que

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \xrightarrow[n\to\infty]{D} N(0, 1)$$

$$\sqrt{n}(\overline{X} - \mu) \xrightarrow[n\to\infty]{D} N(0, \sigma^2)$$

Ejemplo

Sean $X_1, \ldots, X_n$ v.a's iid tales que $X \sim U(0, \theta), \theta > 0$. Demuestre que $\sqrt{n}(\overline{X} - \frac{\theta}{2}) \xrightarrow[n \to \infty]{D} N(0, \theta^2/12)$, donde $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$.

$$E[X] = \frac{\theta}{2} \quad E[X^2] = \frac{\theta^2}{3} \quad V[X] = E[X^2] - (E[X])^2 = \frac{\theta^2}{3} - \frac{\theta^2}{4} = \frac{\theta^2}{12}$$

Por TCL, se tiene

$$\sqrt{n}(\overline{X} - \theta/2) \xrightarrow[n \to \infty]{D} N(0, \theta^2/12)$$

Estimación puntual (cont.)

Método Delta

Sea $\{T_n\}$ una secuencia de v.a's, $\mu \in \mathbb{R}$ y g una función real de variable real, derivable en un intervalo que contiene el punto μ, con $g'(\mu) \neq 0$.

Se $\sqrt{n}(T_n - \mu) \xrightarrow[n\to\infty]{D} N(0, \sigma^2)$, entonces

$$\sqrt{n}(g(T_n) - g(\mu)) \xrightarrow[n\to\infty]{D} N(0, (g'(\mu))^2\sigma^2)$$

Ejemplo

Sea $X_1, \ldots, X_n$ iid con $E[X] = \mu$ y $V[X] = \sigma^2 < \infty$. Muestre que

1. $\sqrt{n}(\overline{X}^2 - \mu^2) \xrightarrow[n\to\infty]{D} N(0, 4\sigma^2\mu^2)$.
2. $\sqrt{n}\left(\frac{1}{\overline{X}} - \frac{1}{\mu}\right) \xrightarrow[n\to\infty]{D} N\left(0, \frac{\sigma^2}{\mu^4}\right)$

Estimación puntual (cont.)

Mejores estimadores asintóticamente normales - BAN (best asymptotically normal estimators)

Una secuencia de estimadores $T_1^*, \ldots, T_n^*$ de $\tau(\theta)$ se define como el mejor asintóticamente normal (BAN) si y solo si se satisfacen las siguientes 4 condiciones:

1. La distribución de $\sqrt{n}[T_n^* - \tau(\theta)] \xrightarrow[n\to\infty]{D} N\left(0, \sigma^{*2}(\theta)\right)$.
2. Para cada $\varepsilon > 0$, $\lim_{n\to\infty} P[|T_n^* - \tau(\theta)| > \varepsilon] = 0$.
3. Sea $\{T_n\}$ una secuencia de estimadores consitente simple para la cual la distribución de $\sqrt{n}[T_n - \tau(\theta)] \xrightarrow[n\to\infty]{D} N\left(0, \sigma^2(\theta)\right)$.
4. $\sigma^2(\theta)$ no es menor que $\sigma^{*2}(\theta) \forall\theta$ en un intervalo abierto.

- BAN a veces se reemplaza por CANE - Consistent asymptotically normal efficient.

- Estimadores BAN $\Longrightarrow$ son consitentes por (2).

Estimación puntual (cont.)

Consistencia y normalidad asintótica para EMV

Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una función de densidad o de probabilidad $f(x;\theta)$ que satisface ciertas condiciones de regularidad y sea $\widehat{\Theta}_n$ el estimador de máxima verosimilitud de θ, entonces:

1. $\widehat{\Theta}_n \xrightarrow[n\to\infty]{D} N\left(\theta, \dfrac{1}{nE\left\{\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right\}}\right).$

2. La secuencia de estimadores de máxima verosimilitud $\widehat{\Theta}_1, \ldots, \widehat{\Theta}_n$ es BAN y por tanto consistente.

3. $\sqrt{n}(\widehat{\Theta}_n - \theta) \xrightarrow[n\to\infty]{D} N\left(0, \dfrac{1}{E\left\{\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right\}}\right).$

Estimación puntual (cont.)

Ejemplo

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución exponencial negativa $f(x;\theta) = \theta e^{-\theta x} I_{[0,\infty)}(x)$. El estimador de máxima verosimilitud de θ es $\frac{1}{\overline{X}_n}$.

De acuerdo al teorema anterior,

$$\frac{1}{\overline{X}_n} \xrightarrow[n\to\infty]{D} N\left(\theta, \frac{1}{nE\left\{\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right\}} = \frac{\theta^2}{n}\right)$$

Estimación puntual (cont.)

- Se ha considerado la estimación de $\tau(\theta)$, una función de θ y no la estimación simplemente de θ.
- La EMV de $\tau(\theta)$ está dada por $\tau(\widehat{\Theta})$, donde $\widehat{\Theta}$ es el EMV de θ.
- Si $\tau(\bullet)$ es diferenciable, entonces

$$\tau(\widehat{\Theta}) \xrightarrow[n \to \infty]{D} N\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nE\left\{\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right\}}\right)$$

La varianza de la última expresión se conoce como Cota inferior de Cramér-Rao.

Información de Fisher - $I_F(\theta)$

Sea X una v.a. con función de densidad o de probabilidad $f(x;\theta)$. La información de Fisher o llamada información total de Fisher de una v.a. X está dada por

$$I_F(\theta) = E\left\{\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]^2\right\}$$

Se existe $\frac{\partial^2}{\partial\theta^2}\log f(X;\theta), \forall\theta$, entonces

$$I_F(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\right]$$

Estimación puntual (cont.)

Ejemplo

Sea X una v.a. con distribución Bernoulli $(1,\theta)$. Encuentre la información de Fisher de la v.a. X. $P[X=x]=\theta^{x}(1-\theta)^{1-x}; x=0,1.$

$$\log P[X=x] = x\log(\theta)+(1-x)\log(1-\theta)$$
$$\frac{\partial}{\partial\theta}\log P[X=x] = \frac{x}{\theta}-\frac{(1-x)}{1-\theta}$$
$$\frac{\partial^2}{\partial\theta^2}\log P[X=x] = -\frac{x}{\theta^2}-\frac{(1-x)}{(1-\theta)^2}$$

$$\begin{aligned}-E\left[\frac{\partial^2}{\partial\theta^2}\log P[X=x]\right] &= -E\left[-\frac{x}{\theta^2}-\frac{(1-x)}{(1-\theta)^2}\right] = E\left[\frac{x}{\theta^2}\right]+E\left[\frac{1-x}{(1-\theta)^2}\right]\\ &= \frac{E[x]}{\theta^2}+\frac{E[1-x]}{(1-\theta)^2} = \frac{\theta}{\theta^2}+\frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta}+\frac{1}{1-\theta}\\ &= \frac{1}{\theta(1-\theta)}\end{aligned}$$

Ejemplo

Sea X una v.a. con distribución Poisson (λ). Encuentre la información de Fisher de la v.a. X. Recuerde que $P[X = x] = \frac{\lambda^x e^{-\lambda}}{x!}$.