Universidad Rey Juan Carlos

Tema 3. estimación Puntual

Estadística II. Grado de Economía y ADE

Dr. Francisco Rabadán Pérez

www.pacorabadan.com

Esquema

1. Introducción: concepto de estimador
2. Propiedades (deseables) de los estimadores $\hat{\boldsymbol{\theta}}$
   1. **Consistencia.**
   2. **Insesgadez**
   3. **Eficiencia (Cota de Cramer-Rao)**
   4. **Suficiencia**
   5. Invarianza
   6. Robustez

1. Introducción

- Población: $\xi: F(X; \theta_1, \theta_2, \ldots, \theta_m)$. Tal que θ_i es PARÁMETRO: **No es variable aleatoria**
- Muestra: $X = \{x_1, x_2, \ldots, x_n\}$, de donde obtenemos el estadístico T(X) que usamos como ESTIMADOR DE θ, es decir

$T(X) = \hat{\theta}$, que **si es variable aleatoria** (número concreto para cada muestra).

Estimar es sinónimo de aproximación.

$$Si\ \theta\ es \begin{cases} Valor\ Exacto \rightarrow Estimación\ Puntual \\ Intervalo \rightarrow Estimación\ por\ intervalos \end{cases}$$

Universidad Rey Juan Carlos

1. Error Cuadrático Medio (ECM)

$$ECM(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

Minimizar el ECM significa que $\hat{\theta}$ se aproxima a θ en media.

$$E(\hat{\theta} - \theta)^2 = E[\hat{\theta} - \theta + E(\hat{\theta}) - E(\hat{\theta})]^2 = E\left[\left(\hat{\theta} - E(\hat{\theta})\right) - \left(\theta - E(\hat{\theta})\right)\right]^2 =$$

$$= E[\hat{\theta} - E(\hat{\theta})]^2 + E[\theta - E(\hat{\theta})]^2 - 2E\left[\left(\hat{\theta} - E(\hat{\theta})\right)\left(\theta - E(\hat{\theta})\right)\right] =$$

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

$- 2E\left(\hat{\theta} - E(\hat{\theta})\right)E[\theta - E(\hat{\theta})] = 0$ porque $E\left(\hat{\theta} - E(\hat{\theta})\right) = E(\hat{\theta}) - E(\hat{\theta}) = 0$

Y nos queda ... $\boldsymbol{ECM(\hat{\theta}) = \mathrm{V}(\hat{\theta}) + E[\theta - E(\hat{\theta})]^2}$

siendo el sesgo: $\textbf{Sesgo} = E[\theta - E(\hat{\theta})]^2$

Universidad Rey Juan Carlos

2. Propiedades (deseables) de los estimadores $\hat{\boldsymbol{\theta}}$

Resumen

1. **Consistencia:** $\hat{\theta} \xrightarrow[n \to \infty]{} \theta$ (convergencia en probabilidad)

2. **Insesgadez:** $\boldsymbol{E(\hat{\theta}) = \theta} \to ECM(\hat{\theta}) = V(\hat{\theta})$

3. **Eficiencia:** $\mathbf{V(\hat{\theta})}$ **es mínima**

4. **Suficiencia:** $\hat{\theta}$ contiene en sí toda la información proporcionada por la muestra.

#yomequedoencasa

Universidad Rey Juan Carlos

2.1. **Consistencia**: $\hat{\theta} \xrightarrow[n\to\infty]{} \theta$ (convergencia en probabilidad)

- Def: Principio lógico de que a medida que aumentamos la muestra, los valores que obtenemos de la estimación se van aproximando cada vez más a los valores reales de los parámetros poblacionales.
- Un estimador es CONSISTENTE cuando converge en probabilidad al verdadero valor del parámetro.

$$\hat{\theta} \xrightarrow[n\to\infty]{} \theta \leftrightarrow \begin{cases} P\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) \xrightarrow[n\to\infty]{} 1 \\ P\left(\left|\hat{\theta} - \theta\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0 \end{cases}$$

- Sólo se puede comprobar (por la desigualdad de Chebychev) que $\hat{\theta}$ es consistente si $\mathrm{E}(\hat{\theta}) = \theta$
- **INCONVENIENTE**: Lo único que nos dice es que utilicemos muestras grandes, pero **no cuanto mejora** la aproximación de $\hat{\theta}$ a θ al aumentar n.

#yomequedoencasa

2.1. **Consistencia**: $\hat{\theta} \xrightarrow[n\to\infty]{} \theta$ (convergencia en probabilidad)

Ejemplo

$$\xi : N(\mu, \sigma) \text{ y } \sigma^2 < \infty \xRightarrow[m.a.s(n)]{} X = \{x_1, x_2, \dots, x_n\}$$

¿ $\hat{\mu} = \bar{x}$ es consistente?

$\bar{x} = \hat{\mu}$ es consistente si $\bar{x} \xrightarrow[P]{} \hat{\mu} \Leftrightarrow P(|\bar{x} - \mu| < \varepsilon) \xrightarrow[n\to\infty]{} 1$, **<u>comprobamos</u>**:

1. $E[\bar{x}] = \mu$ -> es cierto (lo vimos en el tema 1)
2. Chebychev:

$$\boxed{P(|\bar{x} - \mu| < \varepsilon) \geq 1 - \frac{V(\bar{x})}{\varepsilon^2} = 1 - \frac{\sigma^2/n}{\varepsilon^2} \xRightarrow[n\to\infty]{} 1}$$

Luego es consistente

#yomequedoencasa

2.2. **Insesgadez**: $\mathrm{E}\left(\hat{\theta}\right) = \theta \rightarrow ECM\left(\hat{\theta}\right) = V\left(\hat{\theta}\right)$

- Se basa en el principio lógico de que los valores medios tienen una gran representatividad.
- $\hat{\theta}$ es insesgado si $\mathrm{E}\left(\hat{\theta}\right) = \theta$, si esto no se cumple, aparece el sesgo

$$\mathrm{E}(\hat{\theta}) = \theta + \overbrace{\mathrm{b}(\theta)}^{sesgo} \begin{cases} sesgo > 0 \rightarrow sobreestimar \\ sesgo < 0 \rightarrow infravalorar \end{cases}$$

- Hay que ver que pasa con el sesgo cuando el tamaño de la muestra tiende a infinito…

- $\mathrm{E}\left(\hat{\theta}\right) = \theta \rightarrow \hat{\theta}$ es insesgado
- $\hat{\theta}$ es asintóticamente insesgado si $\begin{cases} 1.\ \mathrm{E}\left(\hat{\theta}\right) = \theta + \overbrace{\mathrm{b}(\theta)}^{sesgo} \\ 2.\ \mathrm{b}(\theta) \xrightarrow[n\rightarrow\infty]{} 0 \end{cases}$

Nota: en estadística solemos considerar que $n \rightarrow \infty$ cuando $n \geq 30$.

Universidad Rey Juan Carlos

2.2. **Insesgadez**: $\mathrm{E}(\hat{\theta}) = \theta \rightarrow ECM(\hat{\theta}) = V(\hat{\theta})$

Ejemplo: Sea $\xi: N(\mu, \sigma)$ ¿ $\bar{x}$ es estimador insesgado de μ?

$$E(\bar{x}) = E\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) = \frac{1}{n} n E(x_i) = \frac{1}{n} n\mu = \mu$$

Luego $\bar{x}$ es insesgado porque $E(\bar{x}) = \mu$ (tema 1)

2.2. **Insesgadez**: $\mathrm{E}\left(\hat{\theta}\right) = \theta \rightarrow ECM\left(\hat{\theta}\right) = V\left(\hat{\theta}\right)$

Ejemplo 2: Sea $\xi : N(\mu, \sigma)$ ¿ s^2 es estimador insesgado de $\boldsymbol{\sigma^2}$?

$$E(s^2) = E\left[\frac{\sum x_i^2}{n} - \bar{x}^2\right] = \frac{\boldsymbol{n-1}}{\boldsymbol{n}}\boldsymbol{\sigma^2} = \boldsymbol{\sigma^2} - \underbrace{\boldsymbol{\sigma^2}/\boldsymbol{n}}_{\textbf{sesgo}}$$

Luego es sesgado, pero como

$$\boldsymbol{b(x)} = \frac{\boldsymbol{\sigma^2}}{\boldsymbol{n}} \xrightarrow[\boldsymbol{n}\to\infty]{} \boldsymbol{0}$$

también es asintóticamente insesgado.

Nota: en estadística solemos considerar que $\boldsymbol{n} \rightarrow \infty$ cuando $\boldsymbol{n} \geq \mathbf{30}$.

2.2. **Insesgadez**: $\mathrm{E}\left(\widehat{\theta}\right) = \theta \rightarrow ECM\left(\hat{\theta}\right) = V\left(\hat{\theta}\right)$

Ejemplo 3: Sea $\xi: N(\mu, \sigma)$ ¿ s_1^2 es estimador insesgado de $\boldsymbol{\sigma}^2$?

$$\boldsymbol{s_1^2} = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{n}{n-1}\frac{\sum(x_i - \bar{x})^2}{n} = \frac{\boldsymbol{n}}{\boldsymbol{n-1}}\boldsymbol{s}^2$$

$$E(s_1^2) = \frac{\boldsymbol{n}}{\boldsymbol{n-1}}\boldsymbol{E}(s^2) = \frac{\boldsymbol{n}}{\boldsymbol{n-1}}\frac{\boldsymbol{n-1}}{\boldsymbol{n}}\boldsymbol{\sigma}^2 = \boldsymbol{\sigma}^2$$

Luego es, insesgado

$$\boldsymbol{b(x) = 0}$$

2.3. **Eficiencia**: $\mathbf{V}(\hat{\boldsymbol{\theta}})$ es mínima

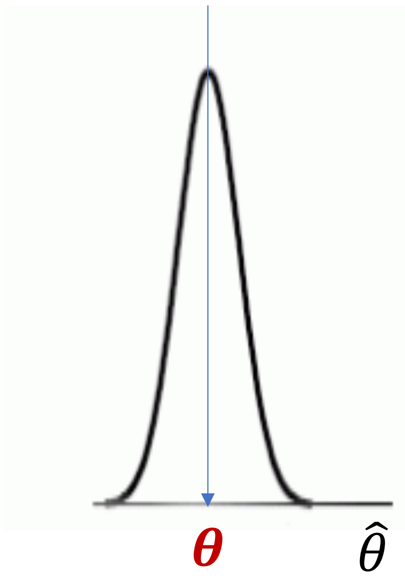
$$\hat{\theta}\, es\ eficiente\ si \begin{cases} \hat{\theta}\ es\ insesgado \\ y \\ V(\hat{\theta})\ es\ mínima \end{cases}$$

Para ver si la varianza del estimador es mínima se calcula la **COTA DE CRAMER-RAO**, que es un valor *único para cada tipo de distribución poblacional y para cada parámetro*.

La **Cota de Cramer-Rao (CCR),** índica la *varianza mínima que debe poseer cualquier estimador del parámetro poblacional.*

$$V(\hat{\theta}) \geq CCR = \frac{1}{nE\left[\frac{\partial \ln[f(X,\theta)]}{\partial \theta}\right]^2}$$

Objetivo: pretende que la distribución de $\hat{\theta}$ esté muy concentrada en θ, es decir, que haya la menor dispersi posible.



2.3. **Eficiencia**: $\min \mathbf{V}(\hat{\boldsymbol{\theta}})$

$$V(\hat{\theta}) \geq CCR = \frac{1}{nE\left[\frac{\partial \ln[f(X,\theta)]}{\partial \theta}\right]^2}$$

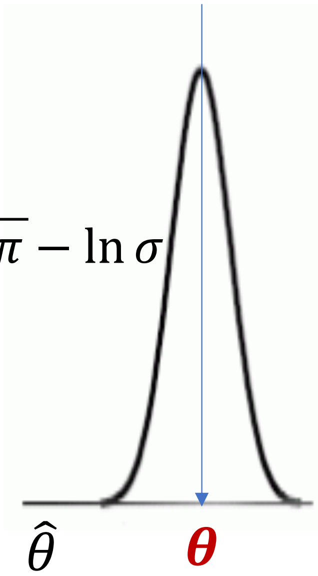
Universidad Rey Juan Carlos

Ejemplo1: Sea $\xi: N(\mu, \sigma)$ ¿ $\bar{x}$ es eficiente?

1º ¿ Insesgado?: si, lo hemos visto. $E(\bar{x}) = \mu$

2º. Cota de Cramer Rao:

1. ***Función de densidad* (o de cuantía):** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
2. $\ln[f(x)] = -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \ln[e] - \ln[\sqrt{2\pi}] - \ln[\sigma] = -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 = \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 - \ln\sqrt{2\pi} - \ln\sigma$
3. $\frac{\partial \ln[f(x)]}{\partial \mu} = \frac{2}{2\sigma^2}(x-\mu) = \frac{(x-\mu)}{\sigma^2}$
4. $E\left[\frac{\partial \ln[P(\xi=x)]}{\partial \mu}\right]^2 = \mathrm{E}\left(\frac{(x-\mu)}{\sigma^2}\right)^2 = \frac{1}{(\sigma^2)^2}\mathrm{E}(x-\mu)^2 = \frac{1}{(\sigma^2)^2}\sigma^2 = \frac{1}{\sigma^2}$
5. $CCR = \frac{1}{nE\left[\frac{\partial \ln[f(X,\theta)]}{\partial \mu}\right]^2} = \frac{1}{n\frac{1}{\sigma^2}} = \frac{\sigma^2}{n}$



Como $CCR = \frac{\sigma^2}{n} = \mathrm{V}(\bar{x}) = \frac{\sigma^2}{n}$ y además $\bar{x}$ Insesgado $\rightarrow$ $\bar{x}$ ES EFICIENTE

#yomequedoencasa

Universidad Rey Juan Carlos

2.3. **Eficiencia**: $\min \mathbf{V}(\widehat{\boldsymbol{\theta}})$

Ejemplo: Sea $\xi: B(5,\pi) \rightarrow m.a.s.(5)$ ¿ $\hat{\pi} = \frac{\bar{x}}{5}$ es eficiente?

1º $\hat{p}$ es Insesgado. $\boldsymbol{E}(\widehat{\boldsymbol{\pi}}) = E\left(\frac{\bar{x}}{5}\right) = \frac{1}{5}E\left(\frac{x_1+x_2+\cdots+x_n}{5}\right) = \frac{1}{5}\frac{5*E(\xi)}{5} = \frac{1}{5}n\pi = \frac{1}{5}5\pi = \boldsymbol{\pi}$

$$\boldsymbol{V}(\widehat{\boldsymbol{\pi}}) = V\left(\frac{\bar{x}}{5}\right) = \frac{1}{5^2}V\left(\frac{x_1+x_2+\cdots+x_n}{5}\right) = \frac{1}{5^2}\frac{1}{5^2}V(x_1+x_2+\cdots+x_n) = \frac{1}{5^4}5(V(\xi)) = \frac{\pi(1-\pi)}{5^2}$$

2º. Cota de Cramer Rao:

1. ***Función* de cuantía:** $\mathbf{P}(\xi = \boldsymbol{x}) = \binom{5}{x}\boldsymbol{\pi}^{x}(\mathbf{1}-\boldsymbol{\pi})^{5-x}$
2. $\ln[\boldsymbol{P}(\xi = \boldsymbol{x})] = \ln\binom{5}{x} + \mathrm{x}\ln\boldsymbol{\pi} + (\mathbf{5}-\boldsymbol{x})\ln(1-\pi)$
3. $\frac{\partial \ln[P(\xi=x)]}{\partial\pi} = \frac{x}{\pi} - \frac{5-x}{1-\pi} = \cdots = \frac{x-5\pi}{\pi(1-\pi)}$
4. $\boldsymbol{E}\left[\frac{\partial \ln[P(\xi=x)]}{\partial\pi}\right]^2 = \boldsymbol{E}\left[\frac{x-5\pi}{\pi(1-\pi)}\right]^2 = \frac{1}{\pi^2(1-\pi)^2}\boldsymbol{E}[x-5\pi]^2 = \frac{V(\xi)}{\pi^2(1-\pi)^2} = \frac{5\pi(1-\pi)}{\pi^2(1-\pi)^2} = \frac{5}{\pi(1-\pi)}$
5. $\boldsymbol{CCR} = \frac{1}{nE\left[\frac{\partial \ln[P(\xi=x)]}{\partial\pi}\right]^2} = \frac{1}{5\frac{5}{\pi(1-\pi)}} = \frac{\pi(1-\pi)}{5^2}$

Como $CCR = \mathrm{V}(\widehat{\boldsymbol{p}})$ y además $\widehat{\boldsymbol{p}}$ es Insesgado $\rightarrow$ $\widehat{\boldsymbol{p}}$ ES EFICIENTE

2.3. **Eficiencia**: $\min \mathbf{V}(\hat{\boldsymbol{\theta}})$

- Diremos que un estimador $\hat{\theta}$ es asintóticamente eficiente si su varianza converge a la CCR cuando el tamaño de la muestra tiende a infinito

$$\hat{\theta}\ asintoticamente\ eficiente\ si\ V(\hat{\theta}) \xrightarrow[n\to\infty]{} CCR$$

2.4. **Suficiencia**: $\hat{\theta}$ contiene en sí toda la información proporcionada por la muestra.

- $\hat{\theta}$ es suficiente cuando contiene toda la información que hay en la muestra respecto al parámetro poblacional desconocido.
- Si la población es $\xi: N(\mu, \sigma)$, son estimadores suficientes:

$$\hat{\mu} = \bar{x}$$

$$\widehat{\sigma^2} = S^2$$

Más en aula virtual URJC y (Martín Pliego ,2005)

Universidad Rey Juan Carlos

- $\hat{\theta}$ es invariante si no varía ante transformaciones.

$$\hat{\sigma} = s \xrightarrow[transformación:\ \wedge 2]{} \widehat{\sigma^2} = S^2$$

Más en aula virtual URJC y (Martín Pliego ,2005)

2.6. **Robustez**: $\hat{\theta}$ en relación a los supuestos del modelo (hipótesis iniciales)

- $\hat{\theta}$ es robusto si no le afectan significativamente alteraciones en los supuestos de un modelo estadístico (hipótesis iniciales).
- Ejemplos de supuestos:
  - Igualdad de varianzas (necesario para la prueba t de igualdad de medias)
  - Población: Normal, Binomial, etc...
- Veremos bastantes más en las asignaturas de Estadística Superior y en Econometría

Bibliografía

- Martín Pliego, Fundamentos de Inferencia Estadística, Editorial AC, 3ªEd, 2005

- Otros recursos en www.pacorabadan.com