

Tarea 2 inferencia estadística: Estimación puntual y por intervalos

1

1. Elabora los siguientes ejercicios del cuadernillo de ejercicios de Estadística II

- 1 • ~~3.1.1~~
- 2 • ~~3.1.7~~
- 3 • ~~3.1.11~~
- 4 • ~~3.1.15~~
- 5 • ~~3.1.18~~ adicionalmente calcula el EMV de θ , la CICR(θ) y la distribución asintótica de θ_{EMV}
- 6 • ~~3.1.19~~ ojo con el inciso b) recuerden que el sesgo se define como la esperanza del estimador menos el parámetro a estimar (la varianza en este caso)
- 7 • ~~3.1.21~~
- 8 • 4.1.2
- 9 • 4.1.6
- 10 • 4.1.9
- 11 • 4.1.7
- 12 • 4.1.20
- 13 • 4.2.1
- 14 • 4.2.2

3.11

PROPIEDADES DE ESTIMADORES

3.1.1. Responda falso o verdadero a las siguientes afirmaciones. Justifique su respuesta.

- a) Un estimador es cualquier característica numérica de la población.
- b) Es posible obtener estimadores diferentes para un mismo parámetro poblacional.
- c) El error de estimación y el sesgo son conceptos que se refieren a lo mismo.
- d) Se dice que un estadístico es un estimador eficiente del parámetro de una población si es insesgado y su varianza es mínima.
- e) El que un estimador $\hat{\theta}$ sea insesgado, significa que cualquier estimación de θ siempre va a ser igual al parámetro poblacional.
- f) Si $ECM(\hat{\theta}_1) < ECM(\hat{\theta}_2)$ eso significa que $V(\hat{\theta}_1) < V(\hat{\theta}_2)$.
- g) Si $ECM(\hat{\theta}_1) < ECM(\hat{\theta}_2)$ eso significa que $\hat{\theta}_1$ es insesgado y $\hat{\theta}_2$ no lo es.

a) FALSO

Un estimador es un estadístico, osea una función de la muestra, que usa para estimar un parámetro desconocido de la población.

b) VERDADERO.

Dependiendo del método que se use (momentos o máxima verosimilitud) puede haber resultados diferentes. Para determinar qué resultado es el "mejor" analizamos el sesgo, error cuadrático medio, eficiencia relativa y consistencia.

c) VERDADERO

Sí siempre y cuando nos refiramos a un error de estimación promedio

Un error de estimación es la diferencia en v. absoluta entre un estimador y el valor real del parámetro.

d) VERDADERO

Así se define la eficiencia de un estadístico

e) FALSA

Un estimador es insesgado cuando LA MEDIA de tu estimador es igual al parámetro, no si el estimador es igual al parámetro.

$$B(\hat{\theta}) = 0 \Leftrightarrow \frac{E(\hat{\theta}) - \theta}{E(\hat{\theta})} = 0$$

f) FALSO

No necesariamente

$$\text{Si } ECM(\hat{\theta}_1) > ECM(\hat{\theta}_2) \Rightarrow \text{Var}(\hat{\theta}_1) + B^2(\theta_1) > \text{Var}(\hat{\theta}_2) + B^2(\theta_2)$$

g) FALSO

No necesariamente $B_1(\theta_1) = 0 \wedge B_2(\theta_2) > 0$, esa solo es una de las posibles causas.

3.1.7

3.1.7. Sea X_1, X_2, \dots, X_n una muestra aleatoria de una población con media μ y varianza σ^2 . Considere los tres estimadores siguientes para μ :

$$\hat{\mu}_1 = \frac{1}{2}(X_1 + X_2), \quad \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{X_2 + \dots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n, \quad \hat{\mu}_3 = \bar{X}$$

a) Determine si son insesgados.

b) Encuentre la varianza de cada estimador e identifique cuál es el más eficiente.

c) Determine la eficiencia relativa de $\hat{\mu}_3$ con respecto a $\hat{\mu}_2$ y $\hat{\mu}_1$, respectivamente.

- Sesgo :

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta \quad \text{Fórmula}$$

a) $\hat{\mu}_1$: $B(\hat{\mu}_1) = E(\hat{\mu}_1) - \mu$

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

$$= \frac{1}{2}E[x_1 + x_2] - \mu$$

$$= \frac{1}{2}E[x_1] + \frac{1}{2}E[x_2] - \mu \Rightarrow E[x_1] = \mu$$

$$= \frac{1}{2}\mu + \frac{1}{2}\mu - \mu = 0$$

$$E[x_2] = \mu$$

\therefore es insesgado $\hat{\mu}_1$

$$\begin{aligned}
 \hat{\mu}_2 &: B(\hat{\mu}_2) = E[\hat{\mu}_2] - \mu \\
 &= E\left[\frac{x_1}{4} + \frac{x_2 + \dots + x_{n-1}}{2(n-2)} + \frac{x_n}{4}\right] - \mu \\
 &= \frac{1}{4}E[x_1] + \frac{E[x_2 + \dots + x_{n-1}]}{2(n-2)} + \frac{1}{4}E[x_n] - \mu
 \end{aligned}$$

donde $E[x_1] = E[x_2] = E[\dots] = \mu$

$$\begin{aligned}
 &= \frac{1}{4}\mu + \frac{(n-2)\mu}{2(n-2)} + \frac{1}{4}\mu - \mu \\
 &= \frac{1}{4}\mu + \frac{1}{2}\mu + \frac{1}{4}\mu - \mu = 0
 \end{aligned}$$

∴ es insesgado $\hat{\mu}_2$

$$\begin{aligned}
 \hat{\mu}_3 &: B(\hat{\mu}_3) = E[\hat{\mu}_3] - \mu \\
 &= E[\bar{x}] - \mu \quad \text{donde } \bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} \\
 &= \frac{E[x_1 + x_2 + \dots + x_n]}{n} - \mu \\
 &= \frac{\cancel{n}}{n}\mu - \mu = 0
 \end{aligned}$$

∴ es insesgado $\hat{\mu}_3$

Varianza muestral :

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

Fórmula

$$b) \frac{\text{var}(\hat{\mu}_1)}{\text{var}(\hat{\mu}_1)} = \text{var}\left(\frac{x_1 + x_2}{2}\right) = \frac{1}{4}\text{var}(x_1 + x_2) \xrightarrow{\substack{\text{ident.} \\ \text{distr.}}} \frac{\text{var}(x_1) + \text{var}(x_2)}{4} \xrightarrow{\substack{\text{indep.}}} \frac{\text{var}(x) + \text{var}(x)}{4} = \frac{1}{2}\text{var}(x)$$

donde $\text{var}(x) = \sigma^2$

$$\Rightarrow \text{var}(\hat{\mu}_1) = \frac{\sigma^2}{2}$$

$$\text{var}(\hat{\mu}_2) = \text{var}\left(\frac{x_1}{4} + \frac{x_2 + \dots + x_{n-1}}{2(n-2)} + \frac{x_n}{4}\right)$$

$$\begin{aligned}
 &\stackrel{\text{ident.}}{=} \frac{1}{16}\text{var}(x_1) + \frac{1}{4(n-2)^2}[\text{var}(x_2) + \text{var}(\dots) + \text{var}(x_{n-1})] + \frac{1}{16}\text{var}(x_n)
 \end{aligned}$$

$$\stackrel{\text{indep.}}{=} \frac{1}{16}\text{var}(x) + \frac{(n-2)\text{var}(x)}{4(n-2)^2} + \frac{1}{16}\text{var}(x) = \text{var}(x) \left(\frac{1}{8} + \frac{1}{4(n-2)} \right)$$

$$= \text{Var}(x) \left(\frac{4(n-2) + 8}{32(n-2)} \right) = \text{Var}(x) \cdot \frac{4n-8+8}{32(n-2)} = \text{Var}(x) \frac{n}{8(n-2)}$$

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

$$\Rightarrow \text{Var}(\hat{\mu}_2) = \frac{\sigma^2 n}{8(n-2)}$$

$$\text{Var}(\hat{\mu}_3) = \text{Var}(\bar{x}) = \text{Var}\left(\frac{\sum x_i}{n}\right) = \frac{1}{n^2} \text{Var}\left(\sum x_i\right) = \frac{\sum \text{Var}(x_i)}{n^2} = \frac{n \text{Var}(x)}{n^2} = \frac{\text{Var}(x)}{n^2}$$

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

$$\Rightarrow \text{Var}(\hat{\mu}_3) = \frac{\sigma^2}{n}$$

c) Eficiencia relativa

$$Ef(\hat{\theta}_1, \hat{\theta}_2) = \frac{ECM(\hat{\theta}_1)}{ECM(\hat{\theta}_2)}, \quad Ef > 1 \Rightarrow \hat{\theta}_2 \text{ es mejor}$$

$$ECM(\hat{\theta}) = \text{Var}(\hat{\theta}) + B^2(\hat{\theta}^2)$$

Fórmulas

$$ECM(\hat{\mu}_1) = \frac{\sigma^2}{2} + 0$$

$$\rightarrow Ef(\hat{\mu}_1, \hat{\mu}_3) = \frac{ECM(\hat{\mu}_1)}{ECM(\hat{\mu}_3)} = \frac{\left(\frac{\sigma^2}{2}\right)}{\left(\frac{\sigma^2}{n}\right)} = \frac{n}{2}$$

$$ECM(\hat{\mu}_2) = \frac{\sigma^2 n}{8(n-2)} + 0$$

si $n > 2 \Rightarrow Ef > 1 \Rightarrow \hat{\mu}_3 \text{ es mejor}$

$$ECM(\hat{\mu}_3) = \frac{\sigma^2}{n} + 0$$

$$Ef(\hat{\mu}_2, \hat{\mu}_3) = \frac{ECM(\hat{\mu}_2)}{ECM(\hat{\mu}_3)} = \frac{\frac{\sigma^2 n}{8(n-2)}}{\left(\frac{\sigma^2}{n}\right)} = \frac{n^2}{8(n-2)}$$

$$n=2 : \frac{4}{0} \quad / \quad n=2.5 : \frac{4.25}{4} > 0 \quad / \quad n=3 : \frac{9}{8} > 1$$

si $n > 2 \Rightarrow Ef > 1 \Rightarrow \hat{\mu}_3 \text{ es mejor}$

3.1.11

3.1.11. Si \bar{X}_1 y \bar{X}_2 son las medias de muestras aleatorias independientes de tamaño n_1 y n_2 tomadas de una población normal con media μ y varianza σ^2 :

- Demuestre que $\hat{\mu} = W\bar{X}_1 + (1-W)\bar{X}_2$ es un estimador insesgado de μ para cualquier W tal que $0 \leq W \leq 1$.
- ¿Cuál es el valor que debe tomar W para que la varianza del estimador $\hat{\mu}$ sea mínima?

a) P.d. $B(\hat{\mu}) = 0$ para $w \in [0,1]$

$$\hat{\mu} = w\bar{X}_1 + (1-w)\bar{X}_2 \quad M_1 = \bar{X}_1, \quad M_2 = \bar{X}_2$$

$$\begin{aligned} B(\hat{\mu}) &= E(\hat{\mu}) - \mu \\ &= E(w\bar{X}_1 + (1-w)\bar{X}_2) - \mu \\ &= wE(\bar{X}_1) + (1-w)E(\bar{X}_2) - \mu \end{aligned}$$

Donde: $E(\bar{X}_1) = E(\bar{X}_2) = \mu$

$$\Rightarrow w\mu + (1-w)\mu - \mu = 0 \quad \text{Q.E.D}$$

b) $\frac{d \text{Var}(\hat{\mu})}{d w} = 0$

$$\begin{aligned} \Rightarrow \text{Var}(\hat{\mu}) &= w^2 \text{Var}(\hat{\mu}_1) + (1-w)^2 \text{Var}(\hat{\mu}_2) \\ &= w^2 \frac{\sigma^2}{n_1} + (1-w)^2 \frac{\sigma^2}{n_2} \\ &= w^2 \left(\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} \right) + \frac{\sigma^2}{n_2} - 2w \frac{\sigma^2}{n_2} \end{aligned}$$

$$\Rightarrow \frac{d \text{Var}(\hat{\mu})}{d w} = 2w \left(\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} \right) - 2 \frac{\sigma^2}{n_2}$$

$$w = \frac{\frac{\sigma^2}{n_2}}{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} = \frac{\frac{\sigma^2}{n_2}}{\frac{n_2 \sigma^2 + n_1 \sigma^2}{n_1 n_2}} = \frac{\sigma^2 n_1 n_2}{\sigma^2 (n_1 + n_2) n_2} = \frac{n_1}{n_1 + n_2}$$

¿Es mínimo?

$$\Rightarrow \frac{d^2 \text{Var}(\hat{\mu})}{d w^2} = \left(\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} \right) 2 > 0 \Rightarrow \text{es mínimo}$$

$$\therefore w = \frac{n_1}{n_1 + n_2}$$

3.1.15

3.1.15. Sea Y una variable aleatoria de Poisson con media λ . Se obtiene una muestra aleatoria de tamaño n y se desea estimar λ . ¿Es $\hat{\lambda} = \sum_{i=1}^n Y_i(Y_i - 1)$ estimador insesgado de λ ?

$$Y \sim P_0(\lambda)$$

$$\text{Sea } \hat{\lambda} = \sum_{i=1}^n Y_i(Y_i - 1)$$

$$\begin{aligned} B(\hat{\lambda}) &= E(\hat{\lambda}) - \lambda = E\left[\sum_{i=1}^n Y_i(Y_i - 1)\right] - \lambda \\ &= E\left[\sum_{i=1}^n Y_i^2\right] - E\left[\sum_{i=1}^n Y_i\right] - \lambda \\ &= n E[Y^2] - n E[Y] - \lambda \\ &= n[\lambda^2 + \lambda - \lambda] - \lambda \\ &= n\lambda^2 - \lambda \neq 0 \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= E[Y^2] - E^2[Y] \\ \lambda &= E[Y^2] - \lambda^2 \\ E[Y^2] &= \lambda^2 + \lambda \end{aligned}$$

∴ $\hat{\lambda}$ no es un estimador insesgado de λ

3.1.18

3.1.18 adicionalmente calcula el EMV de θ , la CICR(θ) y la distribución asintótica de θ_{EMV}

3.1.18. El ingreso mensual percibido por los empleados de nivel ejecutivo de una empresa tiene la distribución:

$$f(x; \theta) = 3\theta^3/x^4 \quad \text{para } x > \theta.$$

Para la contratación de nuevos solicitantes la empresa desea estimar el ingreso mínimo θ , por lo que selecciona una muestra aleatoria de tres empleados de ese nivel y registra su ingreso X_1, X_2 , y X_3 .

Se propone a $\hat{\theta} = b\bar{X}$. Determine el valor de la constante b para que $\hat{\theta}$ sea un estimador insesgado.

$$\begin{aligned} B(\hat{\theta}) &= E(\hat{\theta}) - \theta = E[b\bar{X}] - \theta = bE[\bar{X}] - \theta \\ &= bE\left[\frac{X_1 + X_2 + X_3}{3}\right] - \theta = \frac{b}{3}E[3X] - \theta = bE[X] - \theta \\ &\quad \text{ident. distrib.} \end{aligned}$$

↓
falta!

$$f(x; \theta) = \frac{3\theta^3}{x^4}$$

$$n = 3$$

$$\hat{\theta} = b\bar{X}$$

$$b = ? \text{ para que } B(\hat{\theta}) = 0$$

$$E[X] = \int_0^\infty x \frac{3\theta^3}{x^4} dx = \int_0^\infty \frac{3\theta^3}{x^3} dx = 3 \left[\frac{\theta^3}{-2x^2} \right]_0^\infty = \frac{3}{2} \left[\theta^3 - \frac{\theta^3}{0} \right] = \frac{3\theta^2}{2}$$

$$\Rightarrow = b \left[\frac{3}{2} \theta \right] - \theta = 0$$

$$\theta \left[\frac{3}{2} b - 1 \right] = 0$$

$$b = \frac{1(2)}{3} = \frac{2}{3}$$

$$\Rightarrow b = \frac{2}{3}$$

$$\hat{\theta}_{EMV} = ?$$

Pasos:

1) Fn. verosimilitud.

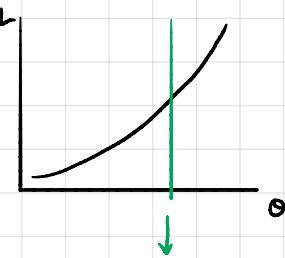
$$f(x; \theta) = \frac{3\theta^3}{x^4} \longrightarrow L(\theta, x) = \prod_{i=1}^n \frac{3\theta^3}{x_i^4} = \frac{27\theta^9}{\pi x^4} = \frac{27\theta^9}{x^{12}}$$

$$\ln L(\theta, x) = 9 \ln[27\theta] - 12 \ln[x]$$

2) Optimizar:

$$\frac{d \ln L(\theta, x)}{d\theta} = \frac{9 \cdot 27}{27\theta} = \frac{9}{\theta} = 0 \quad \nabla$$

* gráficamente:



$\frac{3\theta^3}{x^4}$ es máximo cuando $x \rightarrow 0$

$$\hat{\theta}_{EMV} = \min \{x_1, x_2, x_3\}$$

CICR(θ) = ?

$$CICR(\theta) = \frac{1}{n I(\theta)}, \quad I(\theta) = E \left[\left(\frac{d \ln f(x; \theta)}{d\theta} \right)^2 \right]$$

falta:

$$\Rightarrow \ln f(x; \theta) = \ln \left(\frac{3\theta^3}{x^4} \right) = \ln(3) + 3 \ln(3\theta) - 4 \ln(x)$$

$$\frac{d \ln f(x; \theta)}{d\theta} = \frac{3}{\theta}$$

$$I(\theta) = E \left[\left(\frac{3}{\theta} \right)^2 \right] = E \left[\frac{9}{\theta^2} \right] = \frac{9}{\theta^2}$$

$$\Rightarrow CICR(\theta) = \frac{1}{n I(\theta)} = \frac{1}{3\left(\frac{\theta}{\theta^2}\right)} = \frac{\theta^2}{27} \longrightarrow CICR(\theta) = \frac{\theta^2}{27}$$

• Distribución asintótica de $\hat{\theta}_{EMV}$: $\hat{\theta}_{EMV} \sim AN(\theta, \frac{\theta^2}{27})$

3.1.19

3.1.19. Considere la población constituida por las industrias medianas en México. Es de interés estimar la proporción de éstas industrias que han adquirido algún financiamiento para seguir trabajando. Para ésto se define una variable aleatoria Y tal que toma el valor de uno si la industria adquirió algún financiamiento, y el valor de cero en caso contrario. Si θ es la proporción que se desea estimar, entonces Y se distribuye como una Bernoulli con parámetro θ .

Se ha decidido seleccionar una muestra aleatoria de n industrias. Considere Y_1, Y_2, \dots, Y_n en dicha muestra.

a) Se propone que un estimador de θ sea \bar{Y} .

¿Es este estimador insesgado?

b) Como la varianza del estimador $\hat{\theta} = \bar{Y}$ depende de θ , y θ no se conoce, se desea también estimar dicha varianza. Se propone:

$$\hat{V}(\bar{Y}) = \frac{\bar{Y}(1-\bar{Y})}{n}.$$

Demuestre que este estimador es sesgado. ¿Cuál es el sesgo?

a) $\hat{\theta} = \bar{Y}$ es insesgado?

$$\begin{aligned} B(\hat{\theta}) &= E[\bar{Y}] - \theta = E\left[\frac{Y_1, Y_2, \dots, Y_n}{n}\right] - \theta = E\left[\frac{\cancel{1}Y}{\cancel{n}}\right] - \theta = E[Y] - \theta \\ &= \theta - \theta = 0 \end{aligned}$$

↑
Ident dist.
Bernoulli
 $E[Y] = \theta$

∴ Es insesgado

$$b) \hat{V}(\hat{\theta}) = \frac{\bar{Y}(1-\bar{Y})}{n}, \quad \text{Var}(\bar{Y}) = \frac{1-\theta}{n}$$

$$\begin{aligned} B[\hat{V}(\bar{Y})] &= E[\hat{V}(\bar{Y})] - \text{Var}(\bar{Y}) \\ &= E\left[\frac{\bar{Y}(1-\bar{Y})}{n}\right] - \frac{1-\theta}{n} = \frac{1}{n} E(\bar{Y}) E(1-\bar{Y}) - \frac{1-\theta}{n} \\ &= \frac{\theta(1-\theta)}{n} - \frac{1-\theta}{n} = \frac{\theta - \theta^2 - 1 + \theta}{n} = \frac{\theta(2-\theta) - 1}{n} \end{aligned}$$

∴ Sesgado $B(\bar{Y}) = \frac{\theta(2-\theta) - 1}{n}$

3.1.21

- 3.1.21. Se tiene una muestra aleatoria X_1, X_2, \dots, X_n de una población Poisson con parámetro λ . Calcule el sesgo de los siguientes estimadores:

$$\hat{\lambda}_1 = \frac{2 \sum_{i=1}^n i X_i}{n(n+1)}$$

$$\hat{\lambda}_2 = (X_1 + X_n) / 2n$$

$$X \sim P_0(\lambda)$$

$$B(\hat{\lambda}_1) = E\left[\frac{2 \sum_{i=1}^n i X_i}{n(n+1)}\right] - \lambda$$

$$E[\hat{\lambda}_1] = \frac{2}{n(n+1)} E[1X_1 + 2X_2 + 3X_3 + \dots + nX_n] = \frac{2}{n(n+1)} \cdot E[X_1] + 2E[X_2] + \dots + nE[X_n]$$

$$= \frac{2}{n(n+1)} \cdot (\lambda + 2\lambda + \dots + n\lambda) = \frac{2}{n(n+1)} \cdot \lambda(1+2+\dots+n)$$

sum
gauss.

$$= \frac{2}{n(n+1)} \cdot \lambda \frac{n(n+1)}{2} = \lambda$$

$$\Rightarrow B(\hat{\lambda}_1) = \lambda - \lambda = 0$$

∴ $\hat{\lambda}_1$ es insesgado

$$B(\hat{\lambda}_2) = E(\hat{\lambda}_2) - \lambda$$

$$\hookrightarrow E(\hat{\lambda}_2) = E\left[\frac{X_1 + X_n}{2n}\right] = E\left[\frac{X + X}{2n}\right] = E\left[\frac{2X}{2n}\right] = E\left[\frac{X}{n}\right] = \frac{\lambda}{n}$$

$$\Rightarrow B(\hat{\lambda}_2) = \frac{\lambda}{n} - \lambda = \frac{\lambda - n\lambda}{n} = \lambda \frac{(1-n)}{n}$$

∴ $\hat{\lambda}_2$ es sesgado

4.1.2

4.1.2. Diga si cada una de las siguientes aseveraciones es verdadera o falsa. Justifique su respuesta:

- a) A mayor tamaño de muestra, menor longitud del intervalo de confianza.
- b) Un intervalo de confianza del $(1-\alpha) 100\%$ para un parámetro está contenido en el correspondiente intervalo de confianza al $(1-\alpha') 100\%$ con $\alpha > \alpha'$.
- c) El intervalo es una gama de valores que se usan para estimar la forma de la distribución de una población.
- d) La probabilidad de que un intervalo de confianza contenga al verdadero valor del parámetro recibe el nombre de nivel de confianza.
- e) No es recomendable emplear altos niveles de confianza, pues producen intervalos de confianza más amplios.

- a) V. Ante $\uparrow n \Rightarrow$ nos acercamos + a los valores reales
- b) F. $\alpha < \alpha'$
- c) F. se usa para estimar valores reales (de estimadores)
- d) V. Ya no usamos el término "probabilidad" porque no depende de v.a.
- e) V. $\uparrow I.C \Rightarrow \uparrow$ longitud del intervalo.