UNIVERSIDAD NACIONAL DE COLOMBIA

Estadística I

Estimación Puntual

Estas funciones (estimadores o estadísticos como la media muestral, la proporción muestral, etc.) nos proporcionan **valores puntuales** que se espera estén muy cerca del valor real del parámetro. **¿Qué propiedades deben cumplir para que sean considerados "buenos" estimadores?**

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

Muestral	Poblacional
$\bar{X}$	$\to \mu$
S^2	$\to \sigma^2$
$\hat{p}$	$\to p$
$\bar{X}_1 - \bar{X}_2$	$\to \mu_1 - \mu_2$
$\hat{p}_1 - \hat{p}_2$	$\to p_1 - p_2$

Estimación Puntual

$\bar{x} \rightarrow \leftarrow \hat{\theta} \rightarrow$ estimador o aprox

$\mu \rightarrow \theta \rightarrow$ parámetro desconocido.

Estimador:

Un estimador es una regla que establece cómo usar la información obtenida de una muestra aleatoria, para obtener una aproximación o estimación de un parámetro de interés θ, en general se denotará como $\hat{\theta}$. Si $\hat{\theta}$ es un estimador de un parámetro θ, a medida que el tamaño de muestra (n) aumenta, se espera que los valores de $\hat{\theta}$ estén muy cerca de θ, así:

$$\hat{\theta} = \theta + Error$$

Propiedades deseables de un estimador puntual

Insesgamiento:

Se dice que un estimador puntual $\hat{\theta}$ de un parámetro poblacional θ es **insesgado** si

$$\boxed{E\left[\hat{\theta}\right] = \theta}$$

En caso contrario, se dice que el estimador es sesgado y el sesgo está dado por:

$$\boxed{\beta = E\left[\hat{\theta}\right] - \theta}$$

$$\bar{X} \longrightarrow E[\bar{X}] = \mu$$

$$\beta(\bar{x}) = E[\bar{x}] - \mu$$
$$= \mu - \mu = 0.$$

Propiedades deseables de un estimador puntual

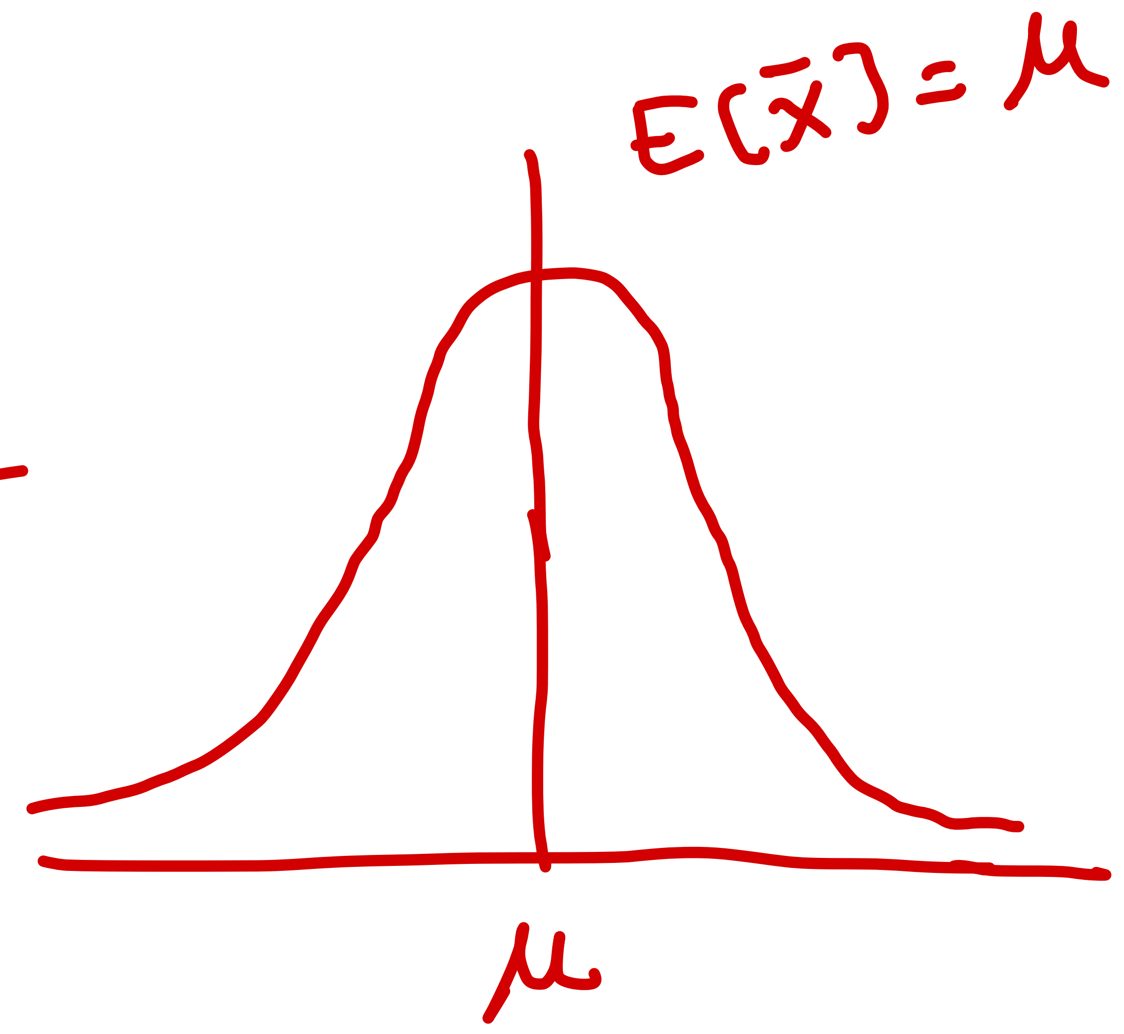
Ejemplo:

Sea $X \sim B(n, P)$ con P desconocido. Muestre que $\hat{P} = \boxed{\dfrac{X}{n}}$ es un estimador insesgado de P.

X: # éxitos en n ensagos

Veamos,

$$E[\hat{p}] = E\left[\frac{X}{n}\right] = \frac{1}{n} E[\bar{X}] = \frac{1}{\cancel{n}}[\cancel{n}p] = p \checkmark$$



Propiedades deseables de un estimador puntual

Propiedades deseables de un estimador puntual

Ejemplo:

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con media $\boxed{\theta}$ y varianza σ^2, ambas desconocidas. Se proponen los siguientes estimadores para θ:

$$\hat{\theta}_1 = \frac{4X_1 + 6X_2 + RX_n}{20}; \quad \hat{\theta}_2 = \frac{5X_1 + 3X_2 + FX_n}{15}$$

a) ¿Cuál es el valor de R y de F para que los 2 estimadores sean insesgados?

b) Reemplace el valor de R y de F respectivamente. ¿Cuál de los 2 estimadores es mejor para θ?

$$\hat{\theta}_1 = \frac{4X_1 + 6X_2 + 10X_n}{20} \quad y \quad \hat{\theta}_2 = \frac{5X_1 + 3X_2 + 7X_n}{15}$$

Propiedades deseables de un estimador puntual

$$\theta = E[\hat{\theta}_1] = E\left[\frac{4X_1+6X_2+RX_n}{20}\right]$$

$$\theta = \frac{1}{20}E[4X_1+6X_2+RX_n]$$

$$\theta = \frac{1}{20}\left[E[4X_1]+E[6X_2]+E[RX_n]\right]$$

$$\theta = \frac{1}{20}\left[4E[X_1]+6E[X_2]+RE[X_n]\right]$$

$$\theta = \frac{1}{20}[4\theta+6\theta+R\theta]$$

$$\theta = \frac{1}{20}(10\theta+R\theta)$$

$$\cancel{\theta} = \frac{\cancel{\theta}}{20}[10+R]$$

$$20 = 10+R \Rightarrow R = 20-10$$

$$\boxed{R=10}$$

$$E[\hat{\theta}_2] = E\left[\frac{5X_1+3X_2+FX_n}{15}\right]$$

$$\theta = \frac{1}{15}E[5X_1+3X_2+FX_n]$$

$$\theta = \frac{1}{15}\left[5E(X_1)+3E(X_2)+FE(X_n)\right]$$

$$\theta = \frac{1}{15}[5\theta+3\theta+F\theta]$$

$$\theta = \frac{1}{15}[8\theta+F\theta]$$

$$\cancel{\theta} = \frac{\cancel{\theta}}{15}[8+F]$$

$$15 = 8+F \Rightarrow \boxed{F=7}$$

Propiedades deseables de un estimador puntual

Estimador insesgado de varianza mínima:

Entre dos estimadores insesgados para θ, se prefiere aquel con menor varianza.

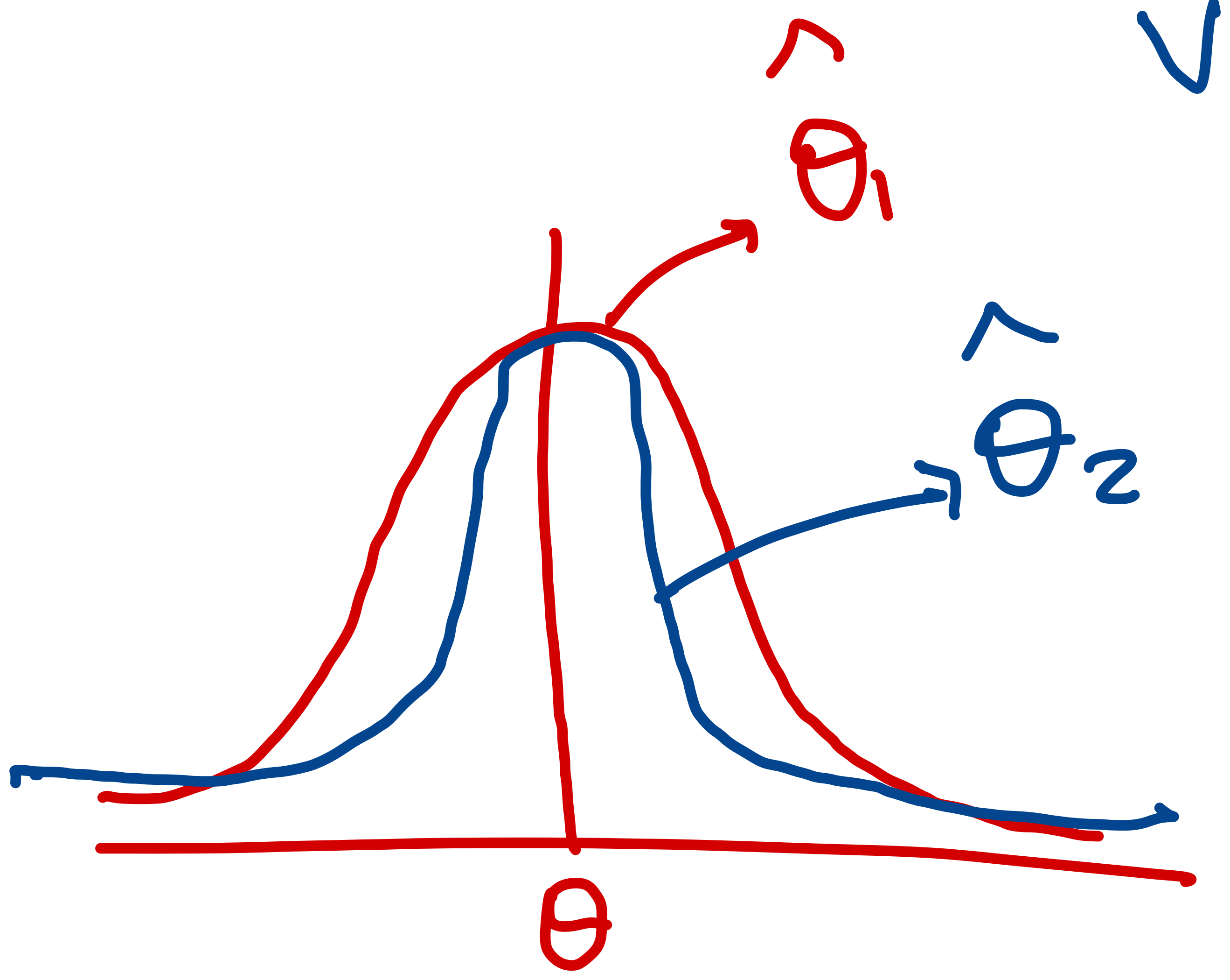
Definición: Error Estándar

Si $\hat{\theta}$ es un estimador de θ, el *error estándar* de $\hat{\theta}$ será su desviación estándar.

$$ee[\hat{\theta}] = \sqrt{Var(\hat{\theta})}$$

$V[\hat{\theta}_2] < V[\hat{\theta}_1]$

$\hat{\theta}_2$ es mejor que $\hat{\theta}_1$



UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades deseables de un estimador puntual

Ejemplo:

Sea una muestra aleatoria de una población con media θ y varianza σ^2, ambas desconocidas. Se proponen los siguientes estimadores para θ:

$$\hat{\theta}_1 = \frac{4X_1 + 6X_2 + RX_n}{20}; \quad \hat{\theta}_2 = \frac{5X_1 + 3X_2 + FX_n}{15}$$

a) ¿Cuál es el valor de R y de F para que los 2 estimadores sean insesgados?
b) Reemplace el valor de R y de F respectivamente. ¿Cuál de los 2 estimadores es mejor para θ?

$R = 10$ y $F = 7$

Propiedades deseables de un estimador puntual

Veamos cuál es mejor

$$V[\hat{\Theta}_1] = Var\left[\frac{4X_1+6X_2+10X_n}{20}\right]$$

$$= \frac{1}{20^2} Var[4X_1+6X_2+10X_n]$$

$$= \frac{1}{400}\left[Var[4X_1]+Var[6X_2]+Var[10X_n]\right]$$

$$= \frac{1}{400}\left(16Var(X_1)+36Var[X_2]+100Var[X_n]\right)$$

$$= \frac{1}{400}\left[16\sigma^2+36\sigma^2+100\sigma^2\right]$$

$$= \frac{152\sigma^2}{400} = \underline{0.38}\,\sigma^2$$

$$Var(\hat{\Theta}_2) = Var\left[\frac{5X_1+3X_2+7X_n}{15}\right]$$

$$= \frac{1}{15^2}\left[25\sigma^2+9\sigma^2+49\sigma^2\right]$$

$$= \frac{83\sigma^2}{225} = \underline{0.3688\,\sigma^2}$$

$\hat{\Theta}_2$ es el mejor para estimar Θ.

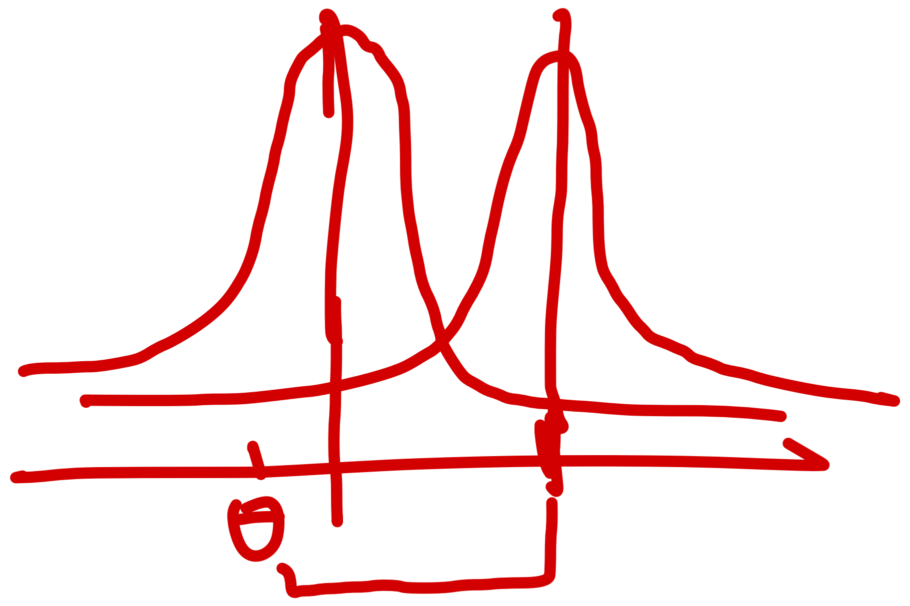
UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades deseables de un estimador puntual

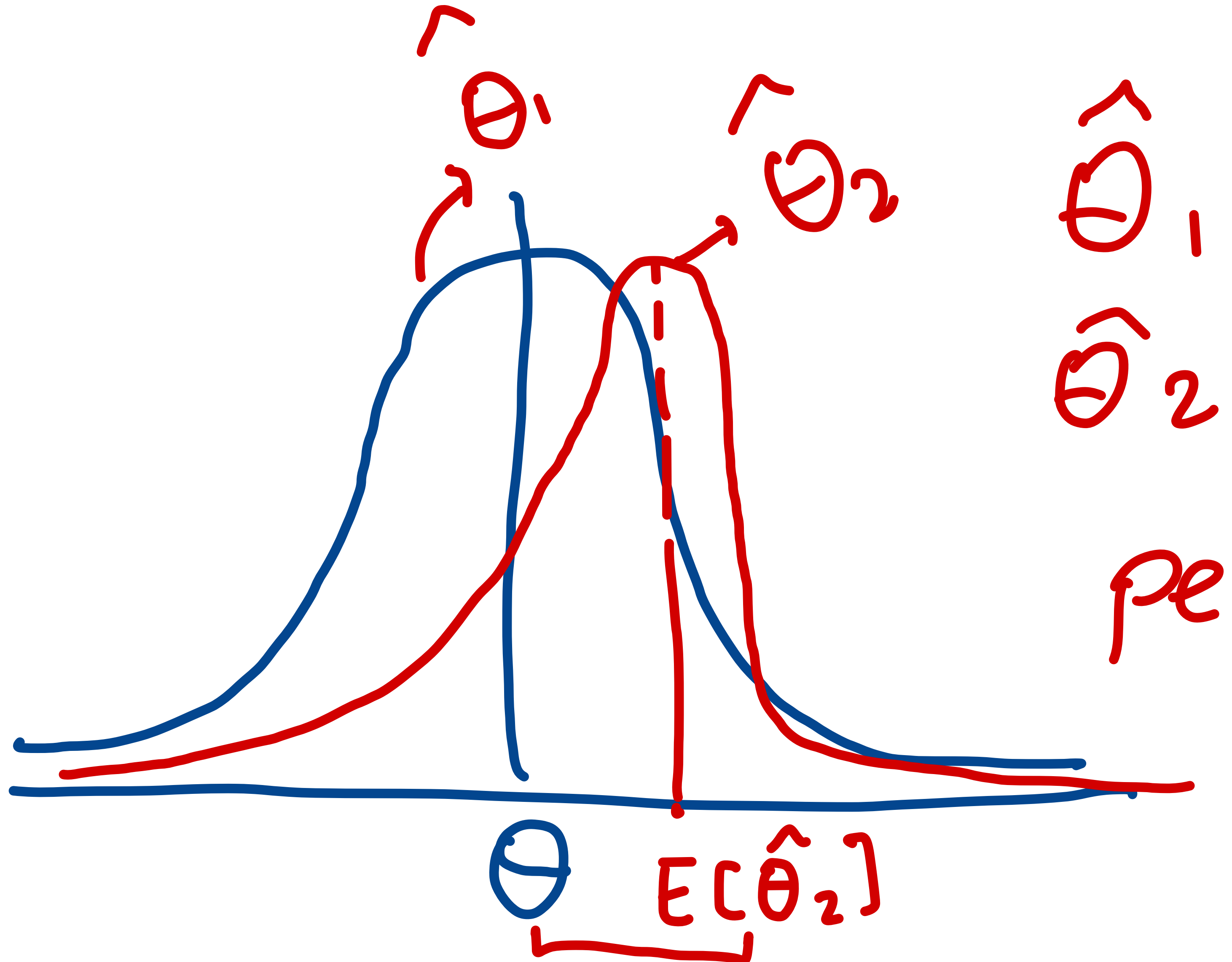
Error Cuadrático Medio (ECM):

Sea $\hat{\theta}$ un estimador de un parámetro poblacional θ, el error cuadrático medio de $\hat{\theta}$ se define como:

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



Nota: entre varios estimadores de un parámetro, se prefiere el que tenga el menor error cuadrático medio.



$\hat{\theta}_1 \rightarrow$ es insesgado

$\hat{\theta}_2 \rightarrow$ es sesgado

pero $V[\hat{\theta}_2] < Var[\hat{\theta}_1]$

¿cuál es el mejor?

Propiedades deseables de un estimador puntual

Ejemplo:

Sea $X_1, X_2, \ldots, X_n$ una m.a de una distribución Poisson con parámetro λ desconocido, considere los siguientes dos estimadores de λ:

$$\hat{\lambda}_1 = \frac{1}{n-1}\sum_{i=1}^{n} X_i \quad y \quad \hat{\lambda}_2 = \frac{1}{n}\left(\sum_{i=1}^{n} X_i - 1\right)$$

a) Determine si ambos estimadores son insesgados para λ.

b) Si $n = 25$ y $\lambda = 4$ ¿cuál de los dos estimadores es mejor para estimar λ?

Propiedades deseables de un estimador puntual

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

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} E[X_i] = \frac{1}{n-1}\sum_{i=1}^{n}\lambda = \frac{n\lambda}{n-1} \neq \lambda$$

Sesgo:

$$B(\hat{\lambda}_1) = E[\hat{\lambda}_1] - \lambda = \frac{n\lambda}{n-1} - \lambda$$

$$= \frac{n\lambda - \lambda(n-1)}{n-1} = \frac{\cancel{n\lambda} - \cancel{n\lambda} + \lambda}{n-1}$$

$$= \frac{\lambda}{n-1} \checkmark$$

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

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} X_i - 1\right)$$

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

$$= \frac{1}{n}\left[\sum_{i=1}^{n}\lambda - 1\right]$$

$$= \frac{1}{n}[n\lambda - 1] = \lambda - \frac{1}{n}$$

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

$$= \cancel{\lambda} - \frac{1}{n} - \cancel{\lambda}$$

$$= -\frac{1}{n} \checkmark$$

Propiedades deseables de un estimador puntual

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

$$= \frac{1}{(n-1)^2} Var\left[\sum_{i=1}^{n} x_i\right]$$

$$= \frac{1}{(n-1)^2}\sum_{i=1}^{n} Var(x_i) = \frac{1}{(n-1)^2}[n\lambda]$$

$$= \frac{n\lambda}{(n-1)^2}$$

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

$$= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} x_i - 1\right)$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(x_i)\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\lambda = \frac{n\lambda}{n^2} = \frac{\lambda}{n}$$

Propiedades deseables de un estimador puntual

$$ECM(\hat{\lambda}_1) = Var(\hat{\lambda}_1) + [\beta(\hat{\lambda}_1)]^2$$

$$= \frac{n\lambda}{(n-1)^2} + \left[\frac{\lambda}{(n-1)}\right]^2$$

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

$$= \frac{25(4) + 4^2}{(24)^2} = 0.2013$$

$$ECM(\hat{\lambda}_2) = Var(\hat{\lambda}_2) + [\beta(\lambda_2)]^2$$

$$= \frac{\lambda}{n} + \left(-\frac{1}{n}\right)^2$$

$$= \frac{\lambda}{n} + \frac{1}{n^2}$$

$$= \frac{4}{25} + \frac{1}{25^2} = 0.1616$$

Propiedades deseables de un estimador puntual

Ejemplo:

Sea $X_1, X_2, \ldots, X_n$ una m.a de una población con función de densidad de probabilidad:

$$f(x) = \frac{2x}{\theta^2}, \quad 0 < x < \theta; \ \theta > 0$$

$$Var[X] = E[X^2) - (E(X))^2.$$

$$Var[X] = \frac{1}{2}\theta^2 - \left(\frac{2}{3}\theta\right)^2$$

$$= \frac{1}{2}\theta^2 - \frac{4}{9}\theta^2 = \boxed{\frac{1}{18}\theta^2}$$

Sean

$$\hat{\theta}_1 = \frac{2X_1 - X_2 + X_3}{3} \quad y \quad \hat{\theta}_2 = \frac{2X_1 - X_2 + X_3}{2}$$

¿Cuál de los dos estimadores es mejor y por qué?

$$E[X] = \int_0^\theta x f(x) dx = \int_0^\theta x \cdot \frac{2x}{\theta^2} dx = \int_0^\theta \frac{2x^2}{\theta^2} = \frac{2}{\theta^2}\int_0^\theta x^2 dx = \frac{2}{\theta^2}\frac{x^3}{3}\Big|_0^\theta = \frac{2}{\cancel{\theta^2}}\frac{\theta^{\cancel{3}}}{3}$$

$$= \frac{2}{3}\theta$$

$$E[x^2] = \int_0^\theta x^2 \frac{2x}{\theta^2} dx = \frac{2}{\theta^2}\int_0^\theta x^3 dx = \frac{2}{\theta^2}\frac{x^4}{4}\Big|_0^\theta = \frac{1}{2}\theta^2$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades deseables de un estimador puntual

$$E[\hat{\theta}_1] = E\left[\frac{2X_1 - X_2 + X_3}{3}\right]$$

$$= \frac{1}{3}\left[E[2X_1] - E(X_2) + E(X_3)\right]$$

$$= \frac{1}{3}\left[2E[X_1] - E[X_2] + E(X_3)\right]$$

$$= \frac{1}{3}\left[2\,\frac{2}{3}\theta - \cancel{\frac{2}{3}\theta} + \cancel{\frac{2}{3}\theta}\right]$$

$$= \frac{4}{9}\theta \neq \theta \quad \text{sesgado}$$

$$\beta(\hat{\theta}_1) = E[\hat{\theta}_1] - \theta$$

$$= \frac{4}{4}\theta - \theta \quad = -\frac{5}{9}\theta$$

$$E[\hat{\theta}_2] = \frac{2}{3}\theta \neq \theta \quad \text{sesgado.}$$

$$\beta(\hat{\theta}_2] = -\frac{1}{3}\theta.$$

Propiedades deseables de un estimador puntual

$$Var[\hat{\theta}_1] = Var\left[\frac{2X_1 - X_2 + X_3}{3}\right]$$

$$= \frac{1}{9}\left[Var[2X_1] + var(X_2) + Var(X_3)\right]$$

$$= \frac{1}{9}\left[4\,var(X_1) + Var(X_2) + Var(X_3)\right]$$

$$= \frac{1}{9}\left[4\left(\tfrac{1}{18}\theta^2\right) + \left(\tfrac{1}{18}\theta^2\right) + \left(\tfrac{1}{18}\theta^2\right)\right]$$

$$= \frac{1}{27}\theta^2$$

$$Var[\hat{\theta}_2] = \frac{1}{12}\theta^2$$

Propiedades deseables de un estimador puntual

$$ECM(\hat{\theta}_1) = Var[\hat{\theta}_1] + (\beta(\hat{\theta}_1))^2$$

$$= \frac{1}{27}\theta^2 + \left(-\frac{5}{9}\theta\right)^2$$

$$= \frac{1}{27}\theta^2 + \frac{25}{81}\theta^2 =$$

$$= 0.3457\theta^2.$$

$\rightarrow$ lo mismo para $\hat{\theta}_2$

Gracias

Clase preparada por: Verónica Guarín