6/02/2025

De los posibles problemas que ataca la inferencia estadística, vimos

i) Estimación puntual

- $\text{Sesgo}_\theta(\hat{\theta}_n) = \mathbb{E}(\hat{\theta}_n) - \theta$

- $\text{Var}(\hat{\theta}_n) = \mathbb{E}\left(\left(\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n)\right)^2\right)$

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

- $\text{ECM}_\theta(\hat{\theta}_n) = \mathbb{E}\left((\hat{\theta}_n - \theta)^2\right)$

$$= \text{Sesgo}^2_\theta(\hat{\theta}_n) + \text{Var}(\hat{\theta}_n)$$

- Consistencia, $\hat{\theta}_n \xrightarrow{P} \theta$

- Distribución exacta o asintótica

$$\hat{\theta}_n \sim \text{¿?} \quad \text{ó} \quad \hat{\theta}_n \xrightarrow{D} \underline{X}$$

El primer ejemplo que vimos es completamente NO paramétrico, sea $X_1, X_2, \dots, X_n$ una m.a. de $X \sim F_X(x)$ y queremos estimar $\mu = \mathbb{E}(X)$ y $\sigma^2 = \text{Var}(X) < c$

Si tomamos $\hat{\mu}_n = \bar{X}_n$

i) $\mathbb{E}(\hat{\mu}_n) = \mu$

ii) $Var(\hat{\mu}_n) = \sigma^2/n$

iii) $ECM_\mu(\hat{\mu}_n) = \sigma^2/n$

iv) $\hat{\mu}_n \xrightarrow{P} \mu$ LDGN

v) Para n suf grande $\hat{\mu}_n \sim N(\mu, \sigma^2/n)$

Segundo ejemplo

Sea $X_1, X_2, \ldots, X_n$ una m.a. de $X \sim Bernoulli(\theta)$

y digamos que de alguna manera tomamos $\hat{\theta}_n = \bar{X}_n$

i) $\mathbb{E}(\hat{\theta}_n) = \mathbb{E}(X) = \theta$

ii) $Var(\hat{\theta}_n) = \frac{1}{n} Var(X) = \frac{\theta(1-\theta)}{n}$

iii) $ECM_\theta(\hat{\theta}_n) = \frac{\theta(1-\theta)}{n}$

iv) $\hat{\theta}_n \xrightarrow{P} \theta$ ¿?

LDGN $\underline{\mathbb{E}(X) = \theta}$ ✓

v) $\dfrac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow{D} Z \sim N(0,1)$

$$se(\hat{\theta}_n) = \sqrt{\frac{\theta(1-\theta)}{n}}$$

Sólo es un esbozo de estimación puntual, más adelante regresaremos a este punto con mas detalle.

ii) Estimación por intervalo

Sea $X_1, \ldots, X_n$ una m.a. de $X \sim f(x|\theta)$ y el objetivo es estimar θ, pero aquí se obtienen

$$\left.\begin{array}{l} L_n = L(X_1, X_2, \ldots, X_n) \\ U_n = U(X_1, X_2, \ldots, X_n) \end{array}\right\}$$ funciones aleatorias de la m.a. sin parámetros desconocidos

tales que

$$P(L_n \leq \theta \leq U_n) = 1 - \alpha,$$

en este caso (L_n, U_n) es un intervalo que cubre a θ con un $(1-\alpha) \times 100\%$ de confianza

Primero, ¿por qué se usaría una estimación por intervalo en lugar de una estimación puntual?

Supongamos que X es una v.a. continua

$\Rightarrow \quad \hat{\Theta}_n = h(X_1, X_2, \ldots, X_n)$ es una v.a. continua

casi seguramente

a) $P(\hat{\Theta}_n = \Theta) = 0$, sin embargo

$$P(L_n \leq \Theta \leq U_n) = 1 - \alpha$$, podemos tomar $\alpha = 0.05$

$= 0.95$

b) Sin obtener el ECM o distribución, la longitud del intervalo nos dice que tan buena es nuestra estimación: intervalos mas delgados nos dan mayor información

Observaciones importantes

1) ¿Qué es eso de la confianza?

¿podemos hablar de probabilidades?

¿Cómo interpretar la confianza?