Inferencia estadística

1. Fenómeno aleatorio o característica de interés de la población $\rightsquigarrow \underline{X}$

2. $\underline{X} \sim \begin{cases} f(x|\theta) \\ \\ F(x) \end{cases}$

3. La información que tenemos $x_1, x_2, \ldots, x_n$ es una realización de una m.a. $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ de $\underline{X}$

4. Usamos esta información para estimar el parámetro de interés de $\underline{X}$

$$\mu = \mathbb{E}(\underline{X}) \quad , \quad \sigma^2 = Var(\underline{X})$$
$$\theta \quad , \quad f(x|\theta), \quad F(x) \quad , \quad P(\underline{X} > a)$$

Existen 3 maneras de atacar el punto 4

i) Estimación puntual
ii) Estimación por intervalo
iii) Prueba de hipótesis

i) Estimación puntual. Si el objetivo es estimar θ

$\Rightarrow$ encontrar un único mejor estimación (estimador) $\hat{\theta}_n$ para θ

$$\hat{\theta}_n = \begin{cases} h(X_1, \dots, X_n) & \leftarrow \text{estimador} \quad \underline{\text{v.a.}} \\ h(x_1, \dots, x_n) & \leftarrow \text{estimación} \quad \underline{\text{valor en } \mathbb{R}} \end{cases}$$

Como tenemos que $X_1, \dots, X_n$ es una m.a. de $f(x|\theta)$

$\Rightarrow$ el estimador $\hat{\theta}_n = h(X_1, \dots, X_n)$ es una v.a.

por lo que podemos estudiar sus propiedades y determinar si es un buen estimador de θ

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

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

3) Error estándar $\text{se}(\hat{\theta}_n) = \sqrt{\text{Var}(\hat{\theta}_n)}$

4) Consistencia $\hat{\theta}_n \xrightarrow{P} \theta$

5) Distribución de muestreo de $\hat{\theta}_n \sim$ ¿? (y/o su distribución asintótica)

Normalidad asintótica

Sea $X_1, X_2, \ldots, X_n$ una m.a. de $f(x|\theta)$ y $\hat{\theta}_n$ un estimador puntual de θ, si

$$\frac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow{D} N(0,1)$$

$\Rightarrow$ para n suf. grande

$$\hat{\theta}_n \sim N(\theta, se^2(\hat{\theta}_n))$$

y por lo tanto diremos que $\hat{\theta}_n$ converge asintóticamente a una distribución normal.

Convergencia en distribución

$$\hat{\theta}_n \xrightarrow{D} X \iff \lim_{n \to \infty} F_{\hat{\theta}_n}(x) = F_X(x)$$

ii) Estimación por intervalo

En lugar de buscar una única mejor estimación el objetivo es encontrar un intervalo (L_n, U_n)

de manera que

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

tales que

$$P(L_n \leq \theta \leq U_n) = 1-\alpha \quad \cdots (*)$$

para todo $\theta \in \mathbb{R}$

¡Importante! $X_1, X_2, \ldots, X_n \Rightarrow (L_n, U_n)$ es un intervalo aleatorio y tiene sentido hablar de probabilidades

$x_1, x_2, \ldots, x_n \Rightarrow (L_n, U_n)$ son valores fijos una vez que sustituimos los valores observados

Cuando sustituimos los valores observados <u>NO</u> tiene sentido hablar de probabilidades

$\Rightarrow$ se crea el termino de "<u>confianza</u>" para los 2 casos

Así en ambos casos a (L_n, U_n) que cumple con $(*)$ se le llama intervalo del $(1-\alpha)\times 100\%$ de confianza para θ.

Obs. 1

El intervalo (L_n, U_n) cuando manejamos v.a. es un intervalo aleatorio ("se mueve"), mientras que θ siempre es desconocido pero fijo

$\Rightarrow$ diremos que (L_n, U_n) cubre, contiene o atrapa a θ ← indican que el intervalo se mueve ✓

jamás diremos θ cae en el intervalo ← indica que θ se mueve ✗

Obs 2. ¿Por qué tenemos que hablar de confianza?

Por una parte cuando tenemos v.a. podemos hablar de probabilidades. Como saben existe una teoría formal de la probabilidad. Sin embargo, al sustituir los valores observados en la muestra ya NO podemos hablar de probabilidades

$\Rightarrow$ para que no existan contradicciones se crea el término "confianza" y se usa en ambos casos.

Obs 3 ¿Cómo interpretamos "un intervalo del $(1-\alpha)\times 100\%$ de confianza para θ"
- piensen en una simulación -

$X_1, \dots, X_n$ es una m.a. de $f(x|\theta)$

$1 = x_1^1, x_2^1, \dots, x_n^1 \longrightarrow (L_n^1, U_n^1)$
$2 = x_1^2, x_2^2, \dots, x_n^2 \longrightarrow (L_n^2, U_n^2)$
$\vdots$
$m = x_1^m, x_2^m, \dots, x_n^m \longrightarrow (L_n^m, U_n^m)$

aproximadamente $(1-\alpha)\times 100\%$ de los $\underline{\underline{m}}$ intervalos contendrían a θ.

Obs 3. ¿Por qué cambia un estimador puntual por un estimador por intervalo?

Si $X_1, \dots, X_n$ m.a. de $X \sim f(x|\theta)$ y X es una v.a. continua.

$\Rightarrow \hat{\theta}_n = h(X_1, X_2, \dots, X_n)$ también es una v.a. continua

$P(\hat{\theta}_n = \theta) = ¿? \; 0$

En cambio, luego toda la estructura de probabilidad que tenemos

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

generalmente $\alpha = 0.05$

$= 0.95$