Para el examen de no paramétrica, estudiar

→ Libro All of statistics capítulos 7 y 8.

El primer problema en estadística no paramétrica es el de estimar $X \sim F(x)$ usando la información de una m.a. $X_1, X_2, \ldots, X_n$ con este objetivo en mente, definimos la función de distribución acumulada empírica.

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x), \text{ para cada } x \in \mathbb{R}$$

en donde

$$\mathbb{I}(X_i \leq x) = \begin{cases} 1, & \text{si } X_i \leq x \\ 0, & \text{eoc.} \end{cases}$$

<u>Teorema</u>

Para cada $x \in \mathbb{R}$, se tiene que

1) $\mathbb{E}(\hat{F}_n(x)) = F(x)$

2) $Var(\hat{F}_n(x)) = F(x)(1 - F(x))$

3) $ECM_{F(x)}(\hat{F}_n) = Var(\hat{F}_n(x))$

4) $\hat{F}_n(x) \xrightarrow{P} F(x)$, en otras palabras $\hat{F}_n(x)$ es un

estimador consistente para $F(x)$ en todo punto $x \in \mathbb{R}$.

Dem

1) $Y_i = \mathbb{1}(X_i \leq x)$, para $i = 1, 2, \ldots, n$

$\Rightarrow Y_1, Y_2, \ldots, Y_n$ es una m.a. de $Y \sim \text{Bernoulli}(F(x))$

$P(Y=1) = P(X \leq x) = F(x)$

$\Rightarrow \overline{Y}_n = \frac{1}{n} \sum Y_i = \hat{F}_n(x)$

$\Rightarrow \mathbb{E}(\overline{Y}_n) = F(x)$ ✓

2) $Var(\overline{Y}_n) = \frac{1}{n^2} n \, F(x)(1-F(x))$

$= \frac{1}{n} F(x)(1-F(x))$

3) $ECM_{F(x)}(\hat{F}_n(x)) = Var(\hat{F}_n(x)) + \cancel{Sesgo}^2_{F(x)}(\hat{F}_n)$ $\leftarrow 0$

$= \frac{1}{n} F(x)(1-F(x))$

4) $\hat{F}_n(x) \xrightarrow{P} F(x)$

pero esto se sigue de que

$\lim_{n \to \infty} ECM_{F(x)}(\hat{F}_n(x)) = 0 \Rightarrow \hat{F}_n(x) \xrightarrow{P} F(x)$

$$\hat{F}_n(x) \xrightarrow{L^2} F(x)$$

Sean $X_1, X_2, \ldots, X_n$ una sucesión de v.a. y $X$ otra v.a.

$$\Rightarrow X_n \xrightarrow{L^2} X$$

$$\Leftrightarrow$$

$$\lim_{n\to\infty} \mathbb{E}\left((X_n - X)^2\right) = 0$$

Proposición

Sea $X_1, X_2, \ldots, X_n$ una sucesión de v.a. y $X$ otra v.a.

Si $X_n \xrightarrow{L_2} X \Rightarrow X_n \xrightarrow{P} X$

Para demostrar esto primero se demuestra la desigualdad de Markov

Si $X$ es una v.a. no negativa y $\varepsilon > 0$

$$\Rightarrow P(X \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

Dem

$$\mathbb{E}(X) = \int_0^\infty y\, f(x)\, dy$$

$$= \int_0^\varepsilon y f(y)\,dy + \int_\varepsilon^\infty y f(y)\,dy$$

$$\geq \int_\varepsilon^\infty y f(y)\,dy \geq \varepsilon \int_\varepsilon^\infty f(y)\,dy$$

$$= \varepsilon P(Y \geq \varepsilon)$$

$$\Rightarrow P(Y \geq \varepsilon) \leq \mathbb{E}(Y)/\varepsilon$$

Haciendo $Y = |X_n - X|$ se tiene el resultado

$\Rightarrow \hat{F}_n(x) \xrightarrow{P} F(x)$

Teorema (Glivenko-Cantelli)

Sea $X_1, X_2, \dots, X_n$ una m.a. de $X \sim F(x)$. Para $\varepsilon > 0$

$$\Rightarrow \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| \xrightarrow{P} 0$$

$\Rightarrow$ ya sabemos que $\hat{F}_n(x)$ es un buen estimador puntual de $F(x)$ para cualquier $x \in \mathbb{R}$

pero y si queremos un estimado por intervalo

Para $n$ suf. grande (para cada $x \in \mathbb{R}$)

$$\hat{F}_n(x) \overset{a}{\sim} N\left(F(x), \frac{F(x)(1-F(x))}{n}\right)$$

$$\Rightarrow \quad L_n = \hat{F}_n(x) - z_{\alpha/2}\sqrt{\frac{\hat{F}_n(x)(1-\hat{F}_n(x))}{n}}$$

$$U_n = \hat{F}_n(x) + z_{\alpha/2}\sqrt{\frac{\hat{F}_n(x)(1-\hat{F}_n(x))}{n}}$$

## Funcionales estadísticas

¿Por qué es relevante estimar $F(x)$?

Si $X$ es capaz de describir el fenómeno de interés

$\Rightarrow$ una cantidad de interés $\mu = \mathbb{E}(X)$

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{-\infty}^{\infty} x \, dF(x)$$

recordemos que $f(x) = \frac{d}{dx} F(x)$

$\Rightarrow$ ¡Claramente $\mu$ y muchos otros parámetros de interés son funciones de $F(x)$!

Un funcional estadístico $T(F)$ es cualquier función de $F$

$$\mu = \mathbb{E}(X) = \int x \, dF(x)$$

$$\sigma^2 = \mathrm{Var}(X) = \int (x-\mu)^2 \, dF(x)$$

La mediana

$$m = F^{-1}(1/2)$$

Definición (estimador plug-in)

El estimador plug-in de $\theta = T(F)$ se define como

$$\hat{\theta}_n = T(\hat{F}_n)$$

en otras palabras, en donde veamos $F$ en el funcional estadístico sustituimos por $\hat{F}_n$

Definición (Funcional estadístico lineal)

Si

$$T(F) = \int g(x) \, dF(x) \quad \text{para alguna función } g$$

$\Rightarrow$ $T$ es un funcional estadístico lineal

Teorema

El estimador plug-in en el caso de un funcional estadístico lineal

$$T(F) = \int g(x)\, dF(x)$$

$$\Rightarrow T(\hat{F}_n) = \int g(x)\, d\hat{F}_n(x)$$

$$= \frac{1}{n} \sum_{i=1}^{n} g(X_i)$$