Para el examen de no paramétrica, estudiar
→ Libro All of statistics capítulos 7 y 8.

3. Estadística NO paramétrica

→ Es importante que estudien y entiendan el proceso de la inferencia estadística, es una de las preguntas del examen

1: FA → datos $x_1, \dots, x_n$
2: $\exists$ una v.a. X que pueda describir el comportamiento agregado de los datos que genera el FA.

en este caso $\underline{X} \sim F(x)$
ya que estamos en el caso no paramétrico

3: Modelo de recolección de datos $X_1, \dots, X_n$ una m.a. de $X \sim F(x)$

4: En estadística No paramétrica siempre vamos inferir alguna característica de X,

$$\mu = \mathbb{E}(X)$$

$$\sigma^2 = Var(X)$$

$$\gamma = \frac{\mathbb{E}(X^3)}{\left(\sqrt{Var(X)}\right)^3}$$

∴ etc.

desarrollaremos estrategias para hacer estimación puntual y por intervalos de μ, σ^2, γ, ..., etc.

para alcanzar este objetivo, primero necesitamos estimar $\underline{\underline{F(x)}}$

¿por qué? ¡$F(x)$ no nos interesa!

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

$$\text{pero} \quad \underline{\underline{f(x) = \frac{dF(x)}{dx}}}$$

notación más $\underline{\underline{\text{general}}}$

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

$$\underline{\underline{\mu = T_1(F)}}$$

$$\sigma^2 = T_2(F) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

$$= \int x^2 \, dF(x) - \mu^2$$

La idea que debe quedar clara es que si estimamos $F(x)$ eso nos permitirá estimar μ, σ^2, γ, etc.

Para estimar $F(x)$ usando la información de $X_1, X_2, \dots, X_n$ m.a. de $X \sim F(x)$ se construye la función de distribución acumulada empírica.

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

en donde

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

<u>Teorema</u>

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

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

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

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 $i = 1, 2, \dots, n$

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

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

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

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

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

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

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

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

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

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

$$\Longleftrightarrow$$

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

Proposición

Sea $X_1, X_2, \dots, X_n$ 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

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

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

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

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

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

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

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

Teorema (Glivenko-Cantelli)

Sea $X_1, X_2, \ldots, X_n$ 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 $x \in \mathbb{R}$

Funcionales estadísticos

¿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 $f(x) = \dfrac{d}{dx} F(x)$

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

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

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

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

La mediana

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

Definición (estimador plug-in)

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

... 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

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

$\Rightarrow$ T es un funcional estadístico lineal