

# Apuntes de Inferencia Estadística

Paco Mora

25 de octubre de 2022

Estos apuntes están hechos como un complemento a los apuntes de Lorenzo.

## Teorema de Wald

### Consistencia

Sea  $X \sim F(\cdot, \theta)$ ,  $\theta \in \Theta$ ,  $\hat{\theta}_n$  estimador de  $\theta \forall n \in \mathbb{N}$ , definimos las siguientes operaciones:

$$P_{\theta_0}(\hat{\theta}_n \in A) := P(\hat{\theta}_n \in A | \theta = \theta_0)$$

$$E_{\theta_0}[\hat{\theta}_n] = \int \cdots \int_{\psi} \hat{\theta}_n(x) L(x, \theta_0) dx$$

Se dice que  $\hat{\theta}_n$  es consistente para  $\theta \in \Theta$  si  $\hat{\theta}_n \xrightarrow{P_{\theta_0}} \theta_0$  (convergencia en probabilidad),  $\forall \theta_0 \in \Theta$

**Proposición 1.** Si  $\Theta$  es finito y  $\hat{\theta}_n$  es estimador de  $\theta$  entonces se da la consistencia del estadístico si y solo si:

$$\lim_{n \rightarrow \infty} P_{\theta_0}(\hat{\theta}_n = \theta_0) = 1, \forall \theta_0 \in \Theta$$

Es decir,

$$\hat{\theta}_n \xrightarrow{P_{\theta_0}} \theta_0 \forall \theta \in \Theta \iff \lim_{n \rightarrow \infty} P_{\theta_0}(\hat{\theta}_n = \theta_0) = 1$$

**Demostración**

$\Leftarrow$

Trivial.

$\Rightarrow$

Si  $\hat{\theta}_n$  es el EMV de  $\theta$ , tomará un valor aislado dentro de  $\Theta$ .

Si  $\hat{\theta}_n \xrightarrow{P_{\theta_0}} \theta_0$ , tomando  $\varepsilon$  suficientemente pequeño,  $\hat{\theta}_n$  solo puede tomar el valor de  $\theta_0$ , luego  $\lim_{n \rightarrow \infty} P_{\theta_0}(\hat{\theta}_n = \theta_0) = 1$

□

### Proposición 2. Desigualdad de Jensen

Sea  $X$  v.a. y  $g$  una función cóncava, entonces  $E[g(X)] < E(E[X])$  siempre que las esperanzas anteriores existan.

### Proposición 3. Ley fuerte de Kolmogorov

Sea  $\{X_n\}_{n=1}^{\infty}$  sucesión de v.v.a.a. independientes, idénticamente distribuidas y con media finita. Entonces:

$$\sum_{i=1}^n \frac{X_i}{n} \xrightarrow{c.s.} E[X]$$

**Lema 4.** Sean  $\{A_n\}$ ,  $\{B_n\}$  con  $\lim_{n \rightarrow \infty} \{P(A_n)\} = \lim_{n \rightarrow \infty} \{P(B_n)\} = 1$ , entonces:

$$\lim_{n \rightarrow \infty} P(A_n \cap B_n) = 1$$

**Demostración**

$$P(A_n \cap B_n) = P(A_n) + P(B_n) - P(A_n \cup B_n)$$

Pero tenemos:

$$1 \geq P(A_n \cup B_n) \geq P(A_n) \rightarrow 1$$

Luego  $\lim_{n \rightarrow \infty} P(A_n \cap B_n) = 1$  y tenemos que:

$$P(A_n \cap B_n) = P(A_n) + P(B_n) - P(A_n \cup B_n) = 1 + 1 - 1 = 1$$

□

**Teorema 5.** Sea  $X$  variable aleatoria con función de distribución  $F(\cdot, \theta)$  para  $\theta \in \Theta$ , siendo  $\Theta$  un conjunto infinito. Supongamos que se verifica:

- (A1) El soporte de  $F(\cdot, \theta)$  es común para todo  $\theta \in \Theta$ .
- (A2)  $E_{\theta_0} \left[ \log \frac{f}{X, \theta} \right]$  existe y es finita para todo  $\theta$ ,  $\theta_0 \in \Theta$

Si para todo  $n \in \mathbb{N}$  y  $(X_1, \dots, X_n)$  m.a.s de  $X$  existe el EMV( $\hat{\theta}_n$ ) de  $\theta$  y es único entonces  $\hat{\theta}_n$  es un estimador consistente del parámetro  $\theta$ .

**Demostración**

Utilizaremos la Desigualdad de Jensen, tomaremos  $g = \log$  y la v.a.  $\frac{f(X, \theta)}{f(X, \theta_0)}$ . La Desigualdad de Jensen entonces nos dice:

$$E_{\theta_0} \left[ \log \left( \frac{f(X, \theta)}{f(X, \theta_0)} \right) \right] < \log \left( E_{\theta_0} \left[ \frac{f(X, \theta)}{f(X, \theta_0)} \right] \right)$$

Pero notemos que:

$$E_{\theta_0} \left[ \frac{f(X, \theta)}{f(X, \theta_0)} \right] = \int \frac{f(x, \theta)}{f(x, \theta_0)} f(x, \theta_0) dx =_{(A1)} 1$$

Luego tenemos que:

$$E_{\theta_0} \left[ \log \left( \frac{f(X, \theta)}{f(X, \theta_0)} \right) \right] < 0$$

Usaremos ahora la ley fuerte de Kolmogorov a la sucesión  $\left\{ \log \left( \frac{f(X_n, \theta)}{f(X_n, \theta_0)} \right) \right\}_n$  y utilizando que la convergencia casi segura es más fuerte que la convergencia en probabilidad, tenemos :

$$\frac{\log \left( \frac{L(\mathbb{X}, \theta)}{L(\mathbb{X}, \theta_0)} \right)}{n} = \sum_{i=1}^n \frac{\log \left( \frac{f(x_i, \theta)}{f(x_i, \theta_0)} \right)}{n} \xrightarrow{P_{\theta_0}} E_{\theta_0} \left[ \log \left( \frac{f(X, \theta)}{f(X, \theta_0)} \right) \right] < 0$$

Como a partir de cierto  $n$  el logaritmo de la izquierda será negativo podemos tomar:

$$\lim_{n \rightarrow \infty} P_{\theta_0} \left( \log \frac{L(\mathbb{X}, \theta)}{L(\mathbb{X}, \theta_0)} < 0 \right) = \lim_{n \rightarrow \infty} P_{\theta_0} (L(\mathbb{X}, \theta) < L(\mathbb{X}, \theta_0)) = 1$$

Como el espacio paramétrico es finito, utilizaremos la equivalencia para la definición de consistencia y buscaremos el límite:

$$\lim_{n \rightarrow \infty} P_{\theta_0}(\hat{\theta}_n = \theta_0)$$

Estudiamos ahora el suceso  $\{\hat{\theta}_n(x) = \theta_0\} = \bigcap_{\theta \in \Theta, \theta \neq \theta_0} \{L(\mathbb{X}, \theta) < L(X, \theta_0)\}$

El suceso es intersección finita de sucesos cuya probabilidad tienen límite 1. Utilizando el lema llegamos a:

$$\lim_{n \rightarrow \infty} P(\hat{\theta}_n = \theta_0) = 1$$

□

**Ejemplo 1.**  $X \sim \mathcal{U}(\theta, \theta+1)$ ,  $\theta \in \mathbb{R}$ ,  $X_1, \dots, X_n$  m.a.s de  $X$ , ( $f(x, \theta) = 1$ , si  $x \in (\theta, \theta+1)$ ). Buscar el EMV de  $\theta$ .

$$L(\mathbb{X}, \theta) = 1 \text{ si } (x_i \in (\theta, \theta+1), i = 1, \dots, n) \iff (\theta < x_{1:n} \text{ y } x_{n:n} < \theta+1) \iff (x_{n:n} - 1 < \theta < x_{1:n})$$

En  $\alpha(x_{n:n} - 1) + (1 - \alpha)x_{1:n}$  se alcanza el máximo de  $L(\mathbb{X}, \theta) \forall \alpha \in (0, 1)$

## Método de la función pivote

**Proposición 1.** Sea  $X$  una variable aleatoria con función de distribución continua  $F(x, \theta)$ ,  $\theta \in \Theta$  y monótona y sea  $X = (X_1, \dots, X_n)$

Para demostrarlo necesitamos algunos resultados

**Lema 2.** Sea  $X$  v.a. con función de distribución  $F$  continua, entonces  $F(X) \sim \mathcal{U}(0, 1)$

**Lema 3.** Sea  $U \sim \mathcal{U}(0, 1)$ , entonces  $-\log U \sim \text{Exp}(1)$

Vamos ahora con la demostración de la proposición.

**Demostración**

Tenemos que  $\{F(X) \leq x\} \iff \{X \leq F^{-1}(x)\}$ , entonces:

$$P(F(X) \leq x) = P(X \leq F^{-1}(x)) = F(F^{-1}(x)) = x \forall x \in (0, 1)$$

□

**Demostración**

$T(\mathbb{X}, \theta) = - \sum_{j=1}^n \log F(X_j, \theta)$  trivialmente es monótona en  $\theta$  (al serlo  $F$ )

Vemos ahora que  $T(\mathbb{X}, \theta)$  no depende de  $\theta$  en su distribución.

$$T(\mathbb{X}, \theta) = \sum_{j=1}^n E_j$$

Donde  $E_j$  son  $\text{Exp}(1)$  indep. entre sí. Pero sabemos que la suma de exponenciales tiene una distribución Gamma:

$$T(\mathbb{X}, \theta) \sim \Gamma(1, n)$$

Luego no depende de  $\theta$ .

□

## Ejemplo del teorema de Newman-Pearson

**Ejemplo 2.** Sea  $X \sim N(\mu, \theta^2)$  con  $\theta^2$  conocida. Se considera una m.a.s simple de tamaño  $n$  de  $X$ . Obtener el test de máxima potencia y extensión de  $\alpha$  para el test (contraste):

$$H_0 : \mu = \mu_0$$

frente a:

$$H_1 : \mu = \mu_1$$

Aplicamos el teorema de Neyman-Pearson. Tenemos que el test UMP viene dado por:

$$S_1 = \{x \in \text{Sop}(X) : \frac{L(x, \mu_1)}{L(x, \mu_0)} \geq k\}$$

Donde  $k$  verifica que  $\alpha = P_{\theta_1}(X \in S)$

$$L(x, \mu) = \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2} \frac{\sum (x_i - \mu)^2}{\sigma^2}} = \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2} \frac{\sum x_i^2 - 2\mu \sum x_i + n\mu^2}{\sigma^2}} \rightarrow \frac{L(X, \mu_1)}{L(X, \mu_0)} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2} (2(\mu_0 - \mu_1) \sum x_i + n(\mu_1^2 - \mu_0^2))\right\}$$

$$S_1 = \{x \in \text{Sop}(X) : -\frac{1}{2\sigma^2} (2(\mu_0 - \mu_1) \sum x_i + n(\mu_1^2 - \mu_0^2)) \geq k' = \log k\} =$$

Suponiendo  $\mu_1 > \mu_0$

$$\{x \in \text{Sop}(X) : \sum x_i \leq \frac{-2\sigma^2 k' - n(\mu_1^2 - \mu_0^2)}{2(\mu_1 - \mu_0)}\} = \{x \in \text{Sop}(X) : \underbrace{\frac{\sum x_i}{n} \leq \frac{-2\sigma^2 k' - n(\mu_1^2 - \mu_0^2)}{2n(\mu_1 - \mu_0)}}_{=:k''}\}$$

Fijando  $\alpha$ , tendremos que buscar  $k''(\alpha)$ . Sabemos que  $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

$$\alpha = P_{\mu_0}(\bar{X} \geq k''(\alpha)) = P_{\mu_0} \left( \underbrace{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}}_{Z \sim N(0,1)} \geq \frac{k''(\alpha) - \mu_0}{\sigma/\sqrt{n}} \right) \rightarrow \frac{k''(\alpha) - \mu_0}{\sigma/\sqrt{n}} = Z_{1-\alpha}$$

$$\text{Luego } k''(\alpha) = \mu_0 + Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

La probabilidad de error de tipo 2 es:

$$\beta = P_{\mu_1}(\bar{X} < \mu_0 + Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}) = P\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} < \frac{\mu_0 - \mu_1 + Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}}{\sigma/\sqrt{n}}\right) = P\left(Z < \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + Z_{1-\alpha}\right)$$

Entonces:

$$Z_\beta = \frac{k''(\alpha) - \mu_1}{\frac{\sigma}{\sqrt{n}}}$$

$$\mu_0 + Z_{1-\alpha} \frac{\sigma}{\sqrt{n}} = \mu_1 + Z_\beta \frac{\sigma}{\sqrt{n}} \implies \frac{\mu_1 - \mu_0}{Z_{1-\alpha} - Z_\beta} = \frac{\sigma}{\sqrt{n}} \implies n = \left(\frac{\sigma(Z_{1-\alpha} - Z_\beta)}{\mu_1 - \mu_0}\right)^2$$

**Ejemplo 3.** Si  $X \sim N(\mu, \sigma^2 = 4)$  representa la duración en días de una determinada enfermedad y se considera la duración de una muestra de 9 enfermos resultando en días: 5,3,4,2,6,4,5,3,4. Aplicar los resultados del ejemplo anterior para decidir entre:

$$H_0 : \mu = 3$$

frente a:

$$H_1 : \mu = 4$$

con una extensión de  $\alpha = 0,05$

En primer lugar tenemos que nuestra región de rechazo es:

$$S_1 = \left\{x \in \mathbb{R}^9 : \bar{X} \geq 3 + Z_{0,95} \frac{2}{3}\right\} = \{x \in \mathbb{R}^9 : \bar{X} \geq 4,1\}$$

Además tenemos que en este caso  $\bar{X} = 4$ . Luego como  $4 \not\geq 4,1$  aceptamos  $H_0$ , es decir, que  $\mu = 3$ .

**Proposición 4.** Si  $A(\theta)$  es monótona creciente (decreciente).

a) En el caso  $\theta_0 < \theta$ :

$$\delta(x) = \begin{cases} 0 & T(x) < (>)c \\ 1 & T(x) \geq (\leq)c \end{cases}$$

b) En el caso  $\theta_0 > \theta$ :

$$\delta(x) = \begin{cases} 0 & T(x) > (<)c \\ 1 & T(x) \leq (\geq)c \end{cases}$$

**Demostración**

Si tenemos  $L(x, \theta) = \exp\{A(\theta)T(\theta) + B(\theta) + h(x)\}$

$$\frac{L(x, \theta_1)}{L(x, \theta_0)} = \exp\{(A(\theta_1)A(\theta_0))T + (B(\theta_1) - B(\theta_0))\}$$

Entonces si escribimos  $S_1$  como en el teorema de Neyman-Pearson y despejamos:

$$S_1 = \{x \in \text{Sop}(X) : T \geq (\leq) \frac{k' - (B(\theta_1) - B(\theta_0))}{A(\theta_1)A(\theta_0)}\}$$

La desigualdad depende del signo de  $A(\theta_1)A(\theta_0)$  (o sea, si es creciente o decreciente).

□

**Ejemplo 4.** Sea  $X \sim \text{Exp}(\theta)$ . Dada una m.a.s. de tamaño  $n$ , obtener el test de máxima potencia y extensión  $\alpha$  para el test (contraste)

$$H_0 : \theta = \theta_0$$

frente a:

$$H_1 : \theta = \theta_1$$

En este caso,  $A(\theta) = -\theta$  y  $T = \sum_{i=1}^n X_i$ . Luego  $A$  es decreciente. Supongamos  $\theta_0 < \theta_1$ , entonces:

$$S_1 = \{x \in R_+^n : \sum_{i=1}^n X_i \leq c\}$$

$$\alpha = P_{\theta=\theta_0}(X \in S_1) = P_{\theta_0} \left( \sum X_i \leq c \right)$$

En el tema 2 obtuvimos que  $\sum X_i \sim \Gamma()$ . Si llamamos a la función cuantil de  $\Gamma$ ,  $G$  entonces  $c = G_\alpha$