Guía 9: Estadística

Supongamos que tenemos una muestra de tamaño n:

$$X_1, X_2, \dots, X_n \quad \text{iid}$$

de una v.a. $X \sim F_{\underline{\theta}}$, donde F_{θ} pertenece a una familia paramétrica $\mathcal{F}$

$\mathcal{F} = \{F_{\underline{\theta}} : \underline{\theta} \in \Theta\}$ con $\Theta \subseteq \mathbb{R}^n$ siendo el conjunto paramétrico

Ej: $\mathcal{F} = \{N(\mu, \sigma^2) : \underbrace{\mu = \mathbb{R}, \sigma > 0}\}$

$\underline{\theta} = (\mu, \sigma) \in (\mathbb{R}, \mathbb{R}^+) = \Theta$

Conocemos la "forma" de la función de distribución $F_{\underline{\theta}}$, pero no el parámetro $\underline{\theta}$

El parámetro $\underline{\theta}$ es determinístico y desconocido

Ej: Tenemos una moneda que sospechamos esté cargada

$$X = \begin{cases} 0, \text{ sale ceca} \\ 1, \text{ sale cara} \end{cases}$$

$X \sim \text{Ber}(p)$, $p \in [0,1] = \Theta$

$\hookrightarrow$ Parámetro desconocido

$F = \{ Ber(p) : p \in [0,1] \}$

Tiramos la moneda n veces y obtenemos una muestra de tamaño n:

$X_1, X_2, \ldots, X_n$

$$p_{X,p}(x) = (1-p)^x \, p^{1-x}, \quad x = 0,1 \quad \text{con } p \in \Theta = [0,1]$$

Analizamos $p_{\underline{X},p}(\underline{x})$ para "hacer cuentas"

$\underline{x} = [x_1 \; x_2 \ldots x_n]^T$

$\underline{X} = [X_1 \; X_2 \; \ldots \; X_n]^T$

$(X_1 \; X_2 \ldots X_n)$

Llamamos estadístico (o estadística) a cualquier función de la muestra:
$r(X)$

Ej.: $\underline{X} = [X_1 \, X_2 \, X_3]$, $r(\underline{X}) = X_1 + 3X_2 + \sinh(X_3)$

Estadístico suficiente

Def: Dada una muestra de tamaño n, $\underline{X} = [X_1 \; X_2 \cdots X_n]^T$ de una

va con distribución perteneciente a una familia F con un conjunto paramétrico $\Theta \subseteq \mathbb{R}^n$, decimos que el estadístico $T = r(\underline{X})$ es suficiente para $\underline{\theta}$ si

$$f_{\underline{X},\underline{\theta}|T=t} = f_{\underline{X}|T=t}, \quad \forall \underline{\theta} \in \Theta, \forall t \in \mathbb{R}$$

i.e. (es decir) la distribución condicional es ind. del parámetro

9.1 Sea X_1, X_2, X_3 una muestra aleatoria de la distribución Bernoulli(p).

(a) Verificar que $T = X_1 + X_2 + X_3$ es un estadístico suficiente para p.

(b) ¿$T = X_1 + 2X_2 + X_3$ es un estadístico suficiente para p?. *:no es difícil analizar los 8 casos.*

$X \sim Ber(p)$, $\underline{X} = [X_1 \; X_2 \; X_3]$

a) Verificar que $T = X_1 + X_2 + X_3$ es suficiente para p

$r(\underline{x}) = x_1 + x_2 + x_3$; $T = r(\underline{X})$

Queremos ver que $p_{\underline{X},p|T=t} = p_{\underline{X}|T=t}$

$$p_{X,p}(x) = (1-p)^{1-x} p^{x}, \quad x = 0,1$$

$$p_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} p_{X_i,p}(x_i) = (1-p)^{n-\sum_{i=1}^{n} x_i} \, p^{\sum_{i=1}^{n} x_i}, \quad x_i = 0,1, \quad i = 1,2,\dots,n$$

$$p_{\underline{X},p|T=t}(\underline{x}) = \mathbb{P}(\underline{X} = \underline{x} \mid T = t), \quad \underline{t = 0,1,2,\dots,n}$$

Porque $T = r(\underline{X}) = \sum_{i=1}^{n} X_i$

$$= \mathbb{P}\left(\underline{X} = \underline{x} \mid \sum_{i=1}^{n} X_i = t\right)$$

$$= \frac{\mathbb{P}\left(\underline{X} = \underline{x},\ \sum_{i=1}^{n} X_i = t\right)}{\mathbb{P}\left(\sum_{i=1}^{n} X_i = t\right)} \longrightarrow \binom{n}{t} p^t (1-p)^{n-t}$$

porque $\sum_{i=1}^{n} X_i \sim Bin(n,p)$

$t = 0, 1, \dots, n$

$$\mathbb{P}\left(\underline{X} = \underline{x},\ \sum_{i=1}^{n} X_i = t\right) = \sum_{\substack{\forall \underline{y} / \underline{y} = \underline{x} \\ y_1 + y_2 + y_3 = t}} p_{\underline{X},p}(\underline{y})$$

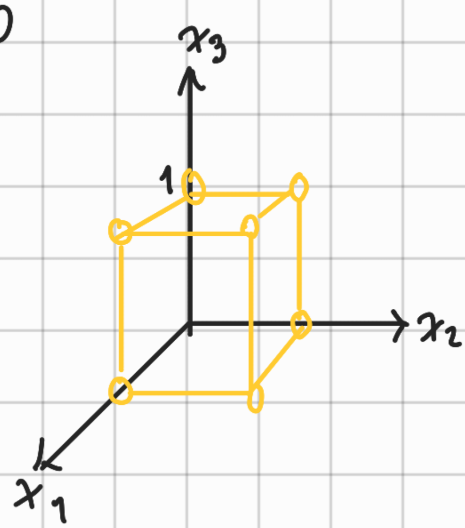
Particularizamos $n = 3$.

$$\mathbb{P}\left(\underline{X} = \underline{x},\ X_1 + X_2 + X_3 = t\right) = \sum_{\substack{\forall \underline{y} / \underline{y} = \underline{x} \\ y_1 + y_2 + y_3 = t}} p_{\underline{X},p}(\underline{y})$$

Caso $T=0$

$$\mathbb{P}(\underline{X}=\underline{x}\,,\,X_1+X_2+X_3=0)=\begin{cases} p_{\underline{X},p}(0,0,0)\,, & x_1=0\,,\,x_2=0,\,x_3=0 \\ 0\,, & \forall \text{ altro caso} \end{cases}$$

$$=\begin{cases} (1-p)^{3-0}\,p^{0}\,, & x_1=0,\,x_2=0,\,x_3=0 \\ 0 & \end{cases}$$



Caso $T=1$

$$\mathbb{P}(\underline{X}=\underline{x}\,,\,X_1+X_2+X_3=1)=\begin{cases} p_{\underline{X},p}(0,0,1)\,, & \underline{x}=(0,0,1) \\ p_{\underline{X},p}(0,1,0)\,, & \underline{x}=(0,1,0) \\ p_{\underline{X},p}(1,0,0)\,, & \underline{x}=(1,0,0) \\ 0\,, & \forall \text{ altro } \underline{x} \end{cases}$$

$$(1-p)^{3-1}p^{1}=(1-p)^{2}p$$

$$= \begin{cases} (1-p)^2 p \ , & x = (0,0,1), (0,1,0), (1,0,0) \\ 0, & \forall \text{ otro } \underline{x} \end{cases}$$

$p_{\underline{X},p}(1,1,0) = p_{\underline{X},p}(1,0,1) = p_{\underline{X},p}(0,1,1)$

Caso T=2

$$\mathbb{P}(\underline{X} = \underline{x}, \ X_1+X_2+X_3 = 2) = \begin{cases} (1-p)p^2 \ , & \underline{x} = (1,1,0), (1,0,1), (0,1,1) \\ 0, & \forall \text{ otro } \underline{x} \end{cases}$$

Caso T=3

$$\mathbb{P}(\underline{X} = \underline{x}, \ X_1+X_2+X_3 = 3) = \begin{cases} p^3, & x = (1,1,1) \\ 0, & \forall \text{ otros } \underline{x} \end{cases}$$

Volvemos a $p_{\underline{X},p} | T=t$

$$p_{\underline{X},p \ | T=0}(\underline{x}) = \begin{cases} \dfrac{(1-p)^3}{\binom{3}{0} p^0 (1-p)^{3-0}} \ , & \underline{x} = (0,0,0) \\ 0 \ , & \forall \text{ otro } x \end{cases}$$

$$= \begin{cases} 1 \ , & x = (0,0,0) \\ 0, & \forall \text{ otro } \underline{x} \end{cases}$$

no depende de p ✓

$$p_{\underline{X},p \mid T=1}(\underline{x}) = \begin{cases} \dfrac{(1-p)^2 p}{\binom{3}{1} p^1 (1-p)^2}, & \underline{x} = (0,0,1), (0,1,0), (1,0,0) \\ 0 & \forall \text{ otro } \underline{x} \end{cases}$$

$$= \begin{cases} 1/3, & \underline{x} = (0,0,1), (0,1,0), (1,0,0) \\ 0 & \forall \text{ otros } \underline{x} \end{cases}$$

no depende de p ✓

$$p_{\underline{X},p \mid T=2}(\underline{x}) = \begin{cases} \dfrac{(1-p)^1 p^2}{\binom{3}{2} p^2 (1-p)^1}, & \underline{x} = (1,1,0), (1,0,1), (0,1,1) \\ 0, & \forall \text{ otro } x \end{cases}$$

$$= \begin{cases} 1/3, & \underline{x} = (1,1,0), (1,0,1), (0,1,1) \\ 0, & \forall \text{ otro } \underline{x} \end{cases}$$

✓

$$p_{\underline{X},p \mid T=3}(\underline{x}) = \begin{cases} \dfrac{p^3}{\binom{3}{3} p^3 (1-p)^{3-3}} & \underline{x} = (1,1,1) \\ 0 & , \forall \text{ otro } \underline{x} \end{cases}$$

$$= \begin{cases} 1, & \underline{x} = (1,1,1) \\ 0, & \forall \text{ otro } \underline{x} \end{cases}$$

En general

$$p_{\underline{X},p \mid T=k}(\underline{x}) = \begin{cases} \dfrac{1}{\binom{n}{k}} \\ 0 \end{cases}$$

Para todo t, nos dio independiente de p por lo tanto el $T = \sum_{i=1}^{n} X_i$ es estadístico suficiente para el parámetro p, para la familia de $Ber(p)$