Martes 10/Mayo/2016.

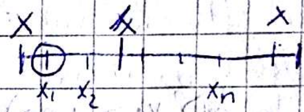
Ayudantía
Intro a Estadística

1) $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exponencial}(\lambda)$

$$F_{X_i}(x_i) = 1 - e^{-\lambda x_i} \qquad i \in \{1, \ldots, n\}$$

¿ $f_{X_{(1)}}(x)$?

$x_i \geq \min x_i \geq x$



$$F_{X_{(1)}}(x) = \mathbb{P}(\min\{X_1, \ldots, X_n\} \leq x)$$
$$= 1 - \mathbb{P}(\min\{X_1, \ldots, X_n\} > x)$$
$$= 1 - \mathbb{P}(X_1 > x, X_2 > x, \ldots, X_n > x)$$

por independencia $= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x)$

$$= 1 - \prod_{i=1}^{n} (1 - F_{X_i}(x))$$
$$= 1 - (1 - F_{X_i}(x))^n$$

$$F_{X_{(1)}}(x) = 1 - (e^{-\lambda x})^n$$
$$= 1 - e^{-\lambda n x}$$

~~$F_{X_{(1)}} = \frac{d}{dx} f_{X_{(1)}}(x)$~~ $\Rightarrow f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) = \lambda n \cdot e^{-\lambda n x}$

$$\Rightarrow X_{(1)} \sim \text{Exponencial}(\lambda n).$$

2) X = "nº de clientes que visitan en un día en un local"

$X \sim \text{Poisson}(\lambda = 100)$

Y = "nº de clientes que compro en un día"

$Y | X = x \sim \text{Binomial}(x, 1/5)$

$\Rightarrow Y \sim \text{Poisson}(\lambda = 20)$

a) ¿$\mathbb{P}(\bar{Y} < 19)$?

Por Teorema central del límite, sabemos que:

$\bar{Y} \overset{\cdot}{\sim} \text{Normal}\left(\mathbb{E}(Y), \frac{Var(Y)}{n}\right)$ donde $n = 80$, $\mathbb{E}(Y) = 20 = Var(Y)$

Luego, $\bar{Y} \dot{\sim} \text{Normal}\left(20, \sigma^2 = \frac{20}{80} = \frac{1}{4}\right)$

$\sigma = 0.5$

$$\mathbb{P}(\bar{Y} < 19) = \mathbb{P}\left(Z < \frac{19-20}{1/2}\right)$$
$$= \mathbb{P}(Z < -2)$$
$$= \Phi(-2)$$
$$= 1 - \Phi(2) = 1 - 0.9972 = 0.0028$$

b) $\mathbb{P}\left(\sum_{i=1}^{40} X_i > 100\right)$

Por Teorema Central del Límite

$$\sum_{i=1}^{40} X_i \dot{\sim} \text{Normal}(n \cdot \mathbb{E}(X), n \cdot Var(X))$$

Luego $\sum_{i=1}^{40} X_i \dot{\sim} \text{Normal}(\mu = 40 \cdot 100 = 4.000, \sigma^2 = 40 \cdot 100 = 4.000)$

$\sigma = \sqrt{4000}$

$$\mathbb{P}\left(\sum_{x=1}^{40} X_i > 100\right) = 1 - \mathbb{P}\left(\sum_{i=1}^{40} X_i \leq 100\right)$$
$$= 1 - \mathbb{P}\left(Z \leq \frac{-3900}{\sqrt{4000}}\right)$$
$$= 1 - \mathbb{P}(Z \leq -61)$$
$$= 1 - \Phi(-61) \cong 1 - 0 = 1$$

③ Sea una muestra aleatoria $X_1, \dots, X_n$ con valores en $\mathcal{X}$ y

$$\mathbb{P}(X = x) = p_\theta(x) = \mathbb{P}(X = x \mid \theta)$$
$$\mathbb{P}(T(\underbrace{X}_{}) = t) = q_\theta(t)$$

• **Def:** Sea $\overbrace{X_1, \dots, X_n}^{\underset{\sim}{X}}$ muestra aleatoria con f.d.p $p_\theta(x)$, y sea $T(X)$ un estadístico con función de prob. $q_\theta(t)$

$T(X)$ es un estadístico suficiente $\Leftrightarrow$

$$\mathbb{P}_\theta(\underset{\sim}{X} = \underset{\sim}{x} \mid T(\underset{\sim}{X}) = t(\underset{\sim}{x}))$$ ~~no de~~ es constante.

(es decir, no depende de θ).

Para probar esto calcularemos

$$\mathbb{P}_\theta(\underset{\sim}{X} = \underset{\sim}{x} \mid T(\underset{\sim}{X}) = T(\underset{\sim}{x}))$$

$$= \frac{\mathbb{P}_\theta(\underset{\sim}{X} = \underset{\sim}{x}, T(\underset{\sim}{X}) = T(\underset{\sim}{x}))}{\mathbb{P}_\theta(T(\underset{\sim}{X}) = T(\underset{\sim}{x}))}$$ notar $\{\underset{\sim}{X} = x\} \subseteq \{T(\underset{\sim}{X}) = T(\underset{\sim}{x})\}$ $\rightarrow$ (más específico).

$$= \frac{\mathbb{P}_\theta(X = x)}{\mathbb{P}_\theta(T(\underset{\sim}{X}) = T(\underset{\sim}{x}))} = \frac{p_\theta(\underset{\sim}{x})}{q_\theta(T(\underset{\sim}{x}))} = (*)$$

Luego T es ~~esp~~ ~~es~~ un estadístico suficiente $\Leftrightarrow$

$$\frac{p_\theta(x)}{q_\theta(T(\underset{\sim}{x}))} \quad \forall x \in \mathcal{X} \text{ es cte en función de } \theta.$$

④ Sea $X_1, X_2, X_3 \overset{iid}{\sim} \text{Bernoulli}(\theta) \quad \theta \in [0,1]$

$$X_i \begin{cases} \theta & x = 1 \\ 1-\theta & x = 0 \end{cases}$$

$$T(\underset{\sim}{x}) = X_1 + 2X_2 + X_3$$

Basta encontrar un contraejemplo para que no sea estadístico suficiente.

$\underset{\sim}{X} = 1, 0, 1 \rightarrow T(\underset{\sim}{X}) = 2$

$$\mathbb{P}_\theta(X_i = (1,0,1) \mid T(\underset{\sim}{X}) = 2) = \frac{\mathbb{P}_\theta(X_i = (1,0,1), T(\underset{\sim}{X}) = 2)}{\mathbb{P}_\theta(T(\underset{\sim}{X}) = 2)} = \frac{\mathbb{P}_\theta(\underset{\sim}{X} = (1,0,1))}{}$$

$$= \frac{\mathbb{P}_\theta(\underset{\sim}{X} = (1,0,1))}{\mathbb{P}_\theta(\underset{\sim}{X} = (1,0,1)) + \mathbb{P}_\theta(\underset{\sim}{X} = (0,1,0))}$$

$$= \frac{\theta \cdot (1-\theta) \cdot \cancel{\theta}}{\theta \cdot (1-\theta)\cancel{\theta} + (1-\theta)^2 \cdot \cancel{\theta}} = \frac{\theta(\cancel{1-\theta})}{(\cancel{1-\theta})(\cancel{\theta}+1-\cancel{\theta})} = \frac{\theta}{1} = \theta$$

Luego, no es estadístico suficiente.

4) Sea $X_1, X_2, \dots, X_n$ con $f_\theta(x) = \theta \cdot x^{\theta-1}$ $\quad 0 < x < 1, \ \theta > 0.$

1° Paso: Ver ~~veo~~ verosimilitud.

$$L_\theta(x_1, \dots, x_n) = \prod_{i=1}^{n} f_{x_i}(x)$$

$$= \prod_{i=1}^{n} \{\theta \cdot x_i^{\theta-1}\}$$

$$= \theta^n \cdot \prod_{i=1}^{n} x_i^{\theta-1}$$

2° Paso: Ver $L_\theta(x_1, \dots, x_n) \approx G(T(x), \theta) \cdot h(x)$.

$$G(T(x), \theta) = \underbrace{\prod_{i=1}^{n} x_i^{\theta-1}} = \underbrace{\left(\prod_{i=1}^{n} x_i\right)^{\theta-1}}_{G(T(X), \theta)}$$