9.9 La duración, X, en años de ciertos discos rígidos tiene la distribución Pareto con densidad

$$f_\theta(x) = \theta x^{-(\theta+1)} \mathbf{1}\{x > 1\}, \qquad \theta > 1.$$

(a) Usar el *criterio de factorización de Neyman-Fisher* para hallar un estadístico suficiente para θ, basado en una muestra aleatoria de la duración de n discos.

(b) Mostrar que las distribuciones f_θ, $\theta > 1$, pertenecen a una familia exponencial y usar esa propiedad para hallar un estadístico suficiente para θ. ¿Cuál es su distribución? :*notar que* $\log X \sim$Exponencial(θ).

(c) Hallar el estimador de máxima verosimilitud de θ basado en una muestra aleatoria de la duración de n discos rígidos. Mostrar que se trata de un estimador asintóticamente insesgado y cuya varianza tiende a 0 cuando $n \to \infty$.

(d) Hallar la distribución asintótica del estimador de máxima verosimilitud de θ. : *es fácil ver que* $I(\theta) = \theta^{-2}$.

(a)

$$f_\theta(x) = \theta\, x^{-(\theta+1)} \mathbb{1}\{x>1\}, \quad \theta > 1$$

Muestra: $(X_1, \dots X_n) \overset{iid}{\sim} \mathrm{Par}(1,\theta)$

$$f_\theta(\underline{x}) = \prod_{i=1}^n f_\theta(x) = \theta^n \prod_{i=1}^n x_i^{-(\theta+1)} \mathbb{1}\{x_i>1\} = \theta^n \left(\prod_{i=1}^n x_i\right)^{-(\theta+1)} \cdot \prod_{i=1}^n \mathbb{1}\{x_i>1\}$$

$$= \underbrace{\theta^n \left(\prod_{i=1}^n x_i\right)^{-\theta}}_{g(r(\underline{x}),\theta)} \underbrace{\left(\prod_{i=1}^n x_i\right)^{-1} \prod_{i=1}^n \mathbb{1}\{x_i>1\}}_{h(x)}$$

Por teorema de factorización

$$\Longrightarrow \boxed{T = \prod_{i=1}^n X_i}$$

(6)

$$f(\theta) = \theta\, x^{-(\theta+1)}\, \mathbb{1}\{x>1\}$$

$$= \underbrace{\theta}_{A(\theta)}\, e^{-(\theta+1)\ln x}\, \underbrace{\mathbb{1}\{x>1\}}_{h(x)}$$

$C(\theta) = -\theta - 1$

$r(x) = \ln x$

$\Longrightarrow$ $\boxed{\text{Pertenecen a una familia exponencial}}$

$\Longrightarrow$ $\boxed{T = \ln X \text{ es un estadístico suficiente para } \theta}$

¿Distribución de T?

$$F_T(t) = \mathbb{P}(T \le t) = \mathbb{P}(\ln X \le t) = \mathbb{P}(X \le e^t)$$

$$= \int_1^{e^t} \theta x^{-\theta-1}\, dx = \cancel{\theta}\, \frac{x^{-\theta}}{-\cancel{\theta}}\Bigg|_1^{e^t} = \underbrace{1^{-\theta}}_{=1} - e^{-t\theta}\, \mathbb{1}\{t\ge 0\}$$

$$f_T(t) = \frac{dF_T}{dt} = t\, e^{-t\theta}\, \mathbb{1}\{t \ge 0\}$$

$$\Longrightarrow \boxed{T_i \sim Exp(\theta)}$$

← lo pongo el i porque es es estadístico cuando solo tengo 1 muestra

¿Estadístico suficiente para θ basado en una muestra de n discos?

$$\Longrightarrow \boxed{T = \sum_{i=1}^{n} \ln X_i} \sim \Gamma(n, \theta)$$

↑ suma de exponenciales independientes

(c)

Los $\prod$ se van porque los x valen un número

$$L(\theta) = F_\theta(\underline{x}) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{-\theta} \left(\prod_{i=1}^{n} x_i\right)^{-1}$$

$$\ln(L(\theta)) = n\ln(\theta) - \theta \ln\left(\prod_{i=1}^{n} x_i\right) - \ln\left(\prod_{i=1}^{n} x_i\right)$$

$$\frac{d \ln(L(\theta))}{d\theta} = \frac{n}{\theta} - \underbrace{\ln\left(\prod_{i=1}^{n} x_i\right)}_{= \sum_{i=1}^{n} \ln(x_i)} = 0$$

$$\Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

$$\Rightarrow \boxed{\hat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} \ln(X_i)}}$$

$$E\left[\hat{\theta}_{MV}\right] = E\left[\frac{n}{\underbrace{\sum_{i=1}^{n} \ln(X_i)}_{=Y}}\right] = E\left[\frac{n}{Y}\right] = \int_{-\infty}^{\infty} \frac{n}{y} f_Y(y)\, dy$$

$Y \sim \Gamma(n, \theta)$

$$= n \int_{-\infty}^{\infty} \frac{1}{y} \frac{\theta^n}{(n-1)!} y^{n-1} e^{-\theta y} dy = \frac{n}{n-1} \theta \underbrace{\int_{-\infty}^{\infty} \frac{\theta^{n-1}}{(n-2)!} y^{n-2} e^{-\theta y} dy}_{=1}$$

$$\rightarrow E\left[\hat{\theta}_{MV}\right] = \frac{n}{n-1} \theta \neq \theta \Rightarrow \boxed{\text{no es insesgado}}$$

$$\lim_{n\to\infty} \frac{n}{n-1}\theta = \theta \Rightarrow \boxed{\text{Es un estimador sesgado asintóticamente}}$$

$$Var(\hat{\theta}_{MV}) = E[\hat{\theta}_{MV}^2] - E[\hat{\theta}_{MV}]^2$$

$$E[\hat{\theta}_{MV}^2] = \int_{-\infty}^{\infty} \frac{n^2}{y^2} \frac{\theta^n}{(n-1)!} y^{n-1} e^{-\theta y} dy$$

$$= \frac{n^2\theta^2}{(n-2)(n-1)} \underbrace{\int_{-\infty}^{\infty} \frac{\theta^{n-2}}{(n-3)!} y^{n-3} e^{-\theta y} dy}_{=1}$$

$$Var(\hat{\theta}_{MV}) = \frac{n^2\theta^2}{(n-2)(n-1)} - \left(\frac{n}{n-1}\theta\right)^2$$

Cuando $n\to\infty$ $= \theta^2\left(\underbrace{\frac{n^2}{(n-2)(n-1)}}_{\to 1} - \underbrace{\frac{n^2}{(n-1)^2}}_{\to 1}\right) \longrightarrow 0$

(d)

$$I(\theta) = -E\left[\frac{d^2}{d\theta^2} \ln(f_\theta(X))\right]$$

$$f_\theta(x) = \theta\, x^{-(\theta+1)}\, \mathbb{1}\{x>1\}, \quad \theta>1$$

$$\ln f_\theta = \ln\theta - (\theta+1)\ln x$$

$$\frac{d \ln f_\theta}{d\theta} = \frac{1}{\theta} - \ln x \quad , \quad \frac{d^2 \ln f_\theta}{d\theta^2} = -\frac{1}{\theta^2}$$

$$I(\theta) = -E\left[-\frac{1}{\theta^2}\right] = \frac{1}{\theta^2} \quad , \text{con } q(\theta)=\theta,\ q'(\theta)=1$$

$$\Rightarrow \boxed{\sqrt{n}\sqrt{\frac{1}{\theta^2}}\,(\hat{\theta}_n-\theta) \overset{(a)}{\sim} \mathcal{N}(0,1)}$$

$$(\hat{\theta}_n-\theta) \overset{(a)}{\sim} \mathcal{N}\left(0, \frac{\theta^2}{n}\right)$$