/ /

Lista 5

① Uma amostra aleatoria, $X_1, \dots, X_n$ tem distribuição de Pareto com pdf

$$f(x|\theta, \nu) = \frac{\theta \nu^{\theta}}{x^{\theta+1}} \mathbb{1}(\nu \le x < \infty), \quad \theta, \nu > 0$$

ⓐ Encontre o MLE de θ e ν.

ⓑ Mostre que o LRT de

$$H_0: \theta = 1 \text{ vs. } H_1: \theta \neq 1, \quad \nu \text{ desconhecido}$$

tem região critica da forma $\{x : T(x) \le c_1 \text{ ou } T(x) \ge c_2\}$ onde $0 < c_1 < c_2$ e

$$T(x) = \log\left[\frac{\prod X_i}{(\min X_i)^n}\right]$$

ⓒ Mostre que, sobre H_0, $2T$ tem uma distribuição χ^2, e encontre o número de graus de liberdade

ⓐ O log da verossimilhança é

$$\log L(\theta, \nu | X) = n\log\theta + n\theta\log\nu - (\theta+1)\log\left(\prod X_i\right), \quad \nu \le X_{(1)},$$

onde $X_{(1)} = \min_i X_i$. Para algum valor de θ, é uma função crescente de ν para $\nu < X_{(1)}$. Então ambos ~~restrições~~ os MLEs restrito e irrestrito de ν são $\hat{\nu} = X_{(1)}$. Para encontrar o MLE de θ, seja

$$\frac{\partial}{\partial\theta} \log L(\theta, X_{(1)} | X) = \frac{n}{\theta} + n\log X_{(1)} - \log\left(\prod X_i\right) = 0,$$

e resolvendo em θ

$$\hat{\theta} = \frac{n}{\log\left(\prod X_i / X_{(1)}^n\right)} = \frac{n}{T}.$$

como $\frac{\partial^2}{\partial\theta^2} \log L(\theta, x_{(1)}|x) = -n/\theta^2$, $\forall\theta$. Então $\hat{\theta}$ é o máximo.

(b) Sobre H_0, o MLE de θ é $\hat{\theta}_0 = 1$, e o MLE de ν continua sendo $\hat{\nu}_{(1)} = X_{(1)}$. Então a razão da verossimilhança

$$\lambda(x) = \frac{x_{(1)}^n / \left(\prod x_i\right)^2}{(n/T)^n x_{(1)}^{n^2/T} / \left(\prod x_i\right)^{\frac{n}{T}+1}}$$

$$= \left(\frac{T}{n}\right)^n \frac{e^{-T}}{(e^{-T})^{n/T}} = \left(\frac{T}{n}\right)^n e^{-T+n}.$$

E $\frac{\partial}{\partial T} \log\lambda(x) = \left(\frac{n}{T}\right) - 1$. Portanto, $\lambda(x)$ é crescente se $T \leq n$ e decrescente se $T \geq n$. Daí, $T \leq c$ é equivalente a $T \leq c_1$ ou $T \geq c_2$, para constantes c_1 e c_2 apropriamente escolhidas.

(c) Seja as seguintes transformações. Primeiro, seja $Y_i = \log X_i$. Depois seja $Z_1 = \min Y_i$ e $Z_2, \ldots, Z_n$ iguais as demais v.a. $Y_2, \ldots, Y_n$ respectivamente. Finalmente, seja $W_1 = Z_1$ e $W_i = Z_i - Z_1$ para $i = 2, \ldots, n$. Então temos que W_i's são independentes com $W_1 \sim f_W(w) = n\nu^n e^{-nw}$, $w > \log\nu$, e $W_i \sim Exp(1)$, $i = 2, \ldots, n$. Agora $T = \sum W_i \sim gamma(n-1, 1)$ e, daí,

$$2T \sim gamma(n-1, 2) = \chi^2_{2(n-1)}.$$

Digitalizado com CamScanner

/ /

② Seja $X_1, X_2 \sim Unif(\theta, \theta+1)$. Para testar $H_0: \theta = 0$ vs $H_1: \theta > 0$ temos 2 testes com diferentes:

$$\phi_1(X_1): \text{Rejeita } H_0 \text{ se } X_1 > 0.95$$
$$\phi_2(X_1, X_2): \text{Rejeita } H_1 \text{ se } X_1 + X_2 > C$$

ⓐ Encontre C tal que ϕ_2 e ϕ_1 tenham mesmo tamanho.

ⓑ Calcule a função poder de cada teste. Desenhe o gráfico da função poder.

ⓒ Prove ou desprove: ϕ_2 é um teste mais poderoso que ϕ_1.

ⓓ Mostre como conseguir um teste com mesmo tamanho porém mais poderoso que ϕ_2.

ⓐ O tamanho de ϕ_1 é $\alpha_1 = P(X_1 > 0.95 | H_0) = 0.05$. O tamanho de ϕ_2 é $\alpha_2 = P(X_1 + X_2 > C | H_0)$. Se $1 \leq C \leq 2$

$$\alpha_2 = P(X_1 + X_2 > C | H_0) = \int_{1-C}^{1} \int_{C-x_1}^{1} 1 \, dx_2 dx_1 = \frac{(2-C)^2}{2}.$$

Tomando esta expressão para ser igual a α_1 e resolvendo em C temos $C = 2 - \sqrt{2\alpha_1}$, para $\alpha_1 = 0.05$ temos ~~$C = 2 - \sqrt{\ }$~~ $C = 2 - \sqrt{0.1} \approx 1.68$.

ⓑ Para este teste temos como função poder

$$\beta_1(\theta) = P_\theta(X_1 > 0.95) = \begin{cases} 0 & \theta \leq -0.05 \\ \theta + 0.05 & -0.05 < \theta \leq 0.95 \\ 1 & 0.95 < \theta \end{cases}$$

Para o segundo teste seja $Y = X_1 + X_2$, que tem distribuição

$$f_Y(y|\theta) = \begin{cases} y - 2\theta & 2\theta \le y < 2\theta + 1 \\ 2\theta + 2 - y & 2\theta + 1 \le y < 2\theta + 2 \\ 0 & c.c. \end{cases}$$

obtemos a função poder para o segundo teste com

$$\beta_2(\theta) = P_\theta(Y > C) = \begin{cases} 0 \quad \cancel{2\theta \le y < 2\theta + 1} & \theta < (C/2) - 1 \\ (2\theta + 2 - C)^2/2 & (C/2) - 1 < \theta \le (C-1)/2 \\ 1 - (C - 2\theta)^2/2 & (C-1)/2 < \theta \le C/2 \\ 1 & C/2 < \theta. \end{cases}$$

(c) É claro que ϕ_1 é mais poderoso para θ perto de 0, mas ϕ_2 é mais poderoso para valores grandes de θ. Ou seja, ϕ_2 não é uniformemente mais poderoso que ϕ_1.

(d) Se qualquer $X_1 \ge 1$ ou $X_2 \ge 1$, devemos rejeitar H_0, pois se $\theta = 0$, $P(X_i < 1) = 1$. Daí, considere a região de rejeição

$$\{(X_1, X_2) : X_1 + X_2 > C\} \cup \{(X_1, X_2) : X_1 > 1\} \cup \{(X_1, X_2) : X_2 > 1\}$$

O primeiro conjunto é a região de rejeição de ϕ_2. Esta região de rejeição tem mesmo tamanho que ϕ_2 pois os outros 2 conjuntos tem probabilidade 0 com $\theta = 0$. Mas para $0 < \theta < C - 1$, a função poder deste teste é maior que β_2. E para $C - 1 \le \theta$ tem mesmo poder.

③ A v.a. X tem pdf $f(x)=e^{-x}$, $x>0$. Uma observação é obtida de $Y=X^{\theta}$, e precisamos construir um teste para $H_0: \theta=1$ vs $H_1: \theta=2$. Encontre o teste UMP de nível $\alpha=0.1$ e compute a probabilidade de erro do tipo II.

A pdf de Y é

$$f(y|\theta)=\frac{1}{\theta}\, y^{(1/\theta)}\, e^{-y^{1/\theta}}, \quad y>0.$$

Pelo Lema de Neyman-Pearson, o ~~t~~ teste UMP é rejeitado se

$$\frac{1}{2} y^{-1/2} e^{y-y^{1/2}} = \frac{f(y|2)}{f(y|1)} > k.$$

Para encontrar a região de rejeição, computamos

$$\frac{\partial}{\partial y}\left(\frac{1}{2} y^{-1/2} e^{y-y^{1/2}}\right) = \frac{1}{2} y^{-3/2} e^{y-y^{1/2}}\left(y-\frac{y^{1/2}}{2}-\frac{1}{2}\right)$$

é negativo para $y<1$ e positivo para $y>1$. Daí $f(y|2)/f(y|1)$ é decrescente para $y\le 1$ e crescente para $y>1$. Portanto, rejeitar para $f(y|2)/f(y|1)>k$ é equivalente a rejeitar para $y\le c_0$ ou $y\ge c_1$. Para obter o tamanho α do teste, as constantes c_0 e c_1 devem satisfazer

$$\cancel{\alpha = P(c_0<Y<c_1|\theta=2)} = \int_{c_0}^{c_1} \frac{1}{2} y^{-1/2} e^{-y^{1/2}}\, dy$$

$$\cancel{= -e^{-y^{1/2}}\Big|_{c_0}^{c_1}}$$

$$\alpha = P(Y\le c_0|H_0) + P(Y\ge c_1|H_0) = 1-e^{-c_0}+e^{-c_1}$$

e

/ /

$$\frac{f(c_0|2)}{f(c_0|1)} = \frac{f(c_1|2)}{f(c_1|1)}.$$

Resolvendo numericamente, para $\alpha = 0.1$, temos $c_0 \simeq 0.0765$ e $c_1 \simeq 3.6378$. E a probabilidade do erro do Tipo II é

$$P(c_0 < Y < c_1 | \theta = 2) = \int_{c_0}^{c_1} \frac{1}{2} y^{-\frac{1}{2}} e^{-y^{1/2}} dy$$

$$= -e^{-y^{1/2}} \Big|_{c_0}^{c_1} \simeq 0.6098.$$

/ /

(4) Seja X uma v.a. onde a pmf sobre H_0 e H_1 são dados por

x	1	2	3	4	5	6	7
$f(x\|H_0)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f(x\|H_1)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

Use o lema de Neyman-Pearson para encontrar o teste mais poderoso para H_0 vs H_1 com tamanho $\alpha = 0.04$. Calcule a probabilidade de erro do tipo II.

Pelo lema de Neyman-Pearson, o teste UMP rejeita para valores altos de $f(x|H_1)/f(x|H_0)$. Calculando temos

x	1	2	3	4	5	6	7
$\frac{f(x\|H_1)}{f(x\|H_0)}$	6	5	4	3	2	1	0.84

A razão é decrescente em X. Então rejeitamos para valores grandes de $f(x|H_1)/f(x|H_0)$ corresponde a rejeitar para valores pequenos de X. Para conseguir tamanho α temos que escolher c de forma que $P(X < c|H_0) = \alpha$. Escolhendo $c = 4$ temos que $P(X \leq 4|H_0) = 0.04$. E a probabilidade de erro do tipo II é $P(X = 5, 6, 7 | H_1) = 0.82$.