Inferência e Decisão I (LMAC)

2003/04

1º Teste

14 de Outubro de 2003

LMAC

Resolução do teste

(a) Seja X uma v.a. com f.d.p. pertencente a $\mathcal{F}$ e $Y = h(X)$, com $h(x) = |x-\mu|$. Dado que h é uma função com derivada contínua para todo o $x \in I\!R$, com $|h'(x)| = |1_{\{x>\mu\}} - 1_{\{x\leq\mu\}}| = 1$, vem que

$$\begin{aligned}
g_Y(y|\mu,\sigma) &= f_X(y+\mu|\mu,\sigma) + f_X(-y+\mu|\mu,\sigma)\\
&= \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma^2}} + \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma^2}}\\
&= \frac{2}{\sigma\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma^2}} I_{I\!R_0^+}(y)\\
&= g_Y(y|\sigma).
\end{aligned}$$

(b) Seja $\mathcal{G} = \{g_Y(y|\sigma) = \frac{2}{\sigma\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma^2}} I_{I\!R_0^+}(y), \sigma^2 \in I\!R^+\}$.

Note-se que $\mathcal{G}$ pertence à família exponencial, com

$$h(y) = I_{I\!R_0^+}(y), \quad c(\sigma) = \frac{2}{\sigma\sqrt{2\pi}}, \quad k = 1, \quad w(\sigma) = -\frac{1}{2\sigma^2}, \quad t(y) = y^2$$

Logo decorre que a estatística $T(\underline{Y}) = \sum_{i=1}^n Y_i^2$ é uma estatística suficiente completa (e, consequentemente, mínima).

Note-se que se $\underline{X}$ for uma amostra aleatória de uma população com f.d.p. pertencente à família $\mathcal{F}$, então uma estatística suficiente para $\mathcal{G}$ é obtida a partir de $\underline{X}$ através da transformação $\sum_{i=1}^n (X_i - \mu)^2$.

(c) Começamos por provar que a família $\mathcal{G}$ é uma família de escala. Para tal basta verificar se a função de distribuição da v.a. $\frac{Y}{\sigma}$ depende ou não de σ. Note-se que

$$\begin{aligned}
P(\frac{Y}{\sigma} \leq y|\sigma) &= P(Y \leq y\sigma)\\
&= P(|X-\mu| \leq y\sigma) = P(-y\sigma \leq X-\mu \leq y\sigma)\\
&= P(-y \leq \frac{X-\mu}{\sigma} \leq y)\\
&= \Phi(y) - \Phi(-y)
\end{aligned}$$

que efectivamente não depende de σ. Logo $\mathcal{G}$ é uma família de escala pelo que, em particular, qualquer estatística função das estatísticas $\frac{Y_1}{Y_n}, \frac{Y_2}{Y_n}, \ldots, \frac{Y_{n-1}}{Y_n}$ é uma estatística ancilar para $\mathcal{G}$. Como $\sum_{i=1}^{n} Y_i^2$ é uma estatística suficiente completa, então é independente de qualquer estatística ancilar pelo que, em particular, é independente de $\frac{\sum_{i=1}^{n-1} Y_i}{Y_n}$. Logo é também independente de $\left(\frac{\sum_{i=1}^{n-1} Y_i}{Y_n}\right)^2$, pelo que a afirmação é correcta.

Admita agora que $\sigma^2 \in I\!R^+$ está fixo e considere a família $\mathcal{F}_\sigma$ de f.d.p. correspondentes, com

$$\mathcal{F}_\sigma = \{f_\sigma(x|\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \mu \in I\!R\}.$$

(d) A família $\mathcal{F}_\sigma$ pertence, tal como $\mathcal{F}$, à família exponencial, com

$$\begin{aligned} f_\sigma(x|\mu) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}} \\ &= e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{-\mu^2}{2\sigma^2}} e^{\frac{\mu x}{\sigma^2}} \end{aligned}$$

pelo que $c(\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\mu}{\sigma^2}}$, $h(x) = e^{-\frac{x^2}{2\sigma^2}}$, $w(\mu) = \frac{\mu}{2\sigma^2}$ e $t(x) = x$. Logo a estatística $T(\underline{X}) = \sum_{i=1}^{n} X_i$ é uma estatística suficiente completa pelo que $\bar{X}$ (estatística equivalente) é também estatística suficiente completa. Logo é independente de qualquer estatística ancilar. Em particular, dado que $\frac{(n-1)S^2}{\sigma^2} \sim \chi^{(2)}_{n-1}$, então $T_1(\underline{X}) = S^2$ (estatística equivalente a $\frac{(n-1)S^2}{\sigma^2}$ - não esquecer que aqui σ é fixo!) é uma estatística ancilar para $\mathcal{F}_\sigma$, pelo que $\bar{X}$ e S^2 são independentes. Em particular tal significa que

$$E\left[\frac{S^2}{\bar{X}}|\mu\right] = E\left[S^2|\mu\right] E\left[\frac{1}{\bar{X}}|\mu\right].$$

O resultado pode ser estendido a $\mathcal{F}$ uma vez que variando σ para todos os valores possíveis de $I\!R^+$, obtemos o mesmo resultado, pelo que o teorema de Basu prova efectivamente que $\bar{X}$ e S^2 são estatísticas independentes mesmo para o modelo $\mathcal{F}$.

(e) Claramente que a família $\mathcal{H}$ pertence à família exponencial, mas na resolução deste exercício não faremos uso desta informação.

Utilizando o critério da factorização podemos determinar uma estatística suficiente para $\mathcal{H}$, já que, baseado numa só observação:

$$f_X(x|\theta) = (1-\theta)\frac{\theta}{2(1-\theta)}^{|x|} I_{\{-1,0,1\}}(x)$$

pelo que $T(X) = |X|$ constitui uma estatística suficiente para $\mathcal{H}$ baseada numa só observação. Vejamos agora se é completa. Note-se que

$$P(|X| = x) = \begin{cases} (1-\theta) & x = 0 \\ \theta & x = 1 \\ 0 & r.v.x \end{cases}$$

Seja agora $h(|X|)$ uma função mensurável, com $E[h(|X|)|\theta] = 0$, pelo que

$$E[h(|X|)|\theta] = h(0)(1-\theta) + h(1)\theta = h(0) + \theta(h(1) - h(0)) = 0, \quad \forall \theta \in \Theta$$

se e só se $h(1) - h(0) = 0$ e $h(0) = 0$, i.e., $h(0) = h(1) = 0$, pelo que $|X|$ é estatística completa.

Claramente que $T(X) = X$ não é estatística completa, uma vez que $E[X|\theta] = 0$ (pois trata-se de uma v.a. simétrica), mas X não é nula quase certamente.