Cálculo das Probabilidades II - Lista 3 - 2020/02 - Gabarito

Prof. Hugo Carvalho

02/06/2021

Questão 1:

a) Note que maximizar a verossimilhança é equivalente a maximizar o logaritmo da verossimilhança, já que a função log é crescente. Portanto:

$$\begin{aligned}\text{Maximizar } \mathcal{L}_n(\theta) &\iff \text{maximizar } \log \mathcal{L}_n(\theta)\\ &\iff \text{maximizar } M_n(\theta) = \frac{1}{n}\sum_{i=1}^n \log\left(\frac{f(x_i|\theta)}{f(x_i|\theta_\star)}\right),\end{aligned}$$

onde a primeira equivalência já foi discutida acima e a segunda se dá pelo fato de podermos reescrever

$$M_n(\theta) = \frac{1}{n}\sum_{i=1}^n \log f(x_i|\theta) - \frac{1}{n}\sum_{i=1}^n \log f(x_i|\theta_\star).$$

Como o último termo é constante em θ, fica clara a equivalência.

b) Pela Lei Fraca dos Grandes Números (e claro, assumindo as hipóteses adequadas para que ela seja aplicada), temos que:

$$\begin{aligned}M_n(\theta) &\xrightarrow{p} \mathbb{E}_{\theta_\star}\left[\log\left(\frac{f(X|\theta)}{f(X|\theta_\star)}\right)\right]\\ &= \int_{-\infty}^{+\infty} f(x|\theta_\star)\log\left(\frac{f(x|\theta)}{f(x|\theta_\star)}\right)\, dx\\ &= -\int_{-\infty}^{+\infty} f(x|\theta_\star)\log\left(\frac{f(x|\theta_\star)}{f(x|\theta)}\right)\, dx\\ &= -D(\theta_\star, \theta).\end{aligned}$$

c) Pelo item b), se n é suficientemente grande, temos que $M_n(\theta) \approx -D(\theta_\star, \theta)$, que assume valor máximo em $\theta = \theta_\star$, pois $-D(\theta_\star, \theta_\star) = 0$ e $-D(\theta_\star, \theta) < 0$ se $\theta \neq \theta_\star$, pelas hipóteses i) e ii). Dessa forma, no limite, $\theta_\star$ maximiza $M_n(\theta)$, e portanto, $\mathcal{L}_n(\theta)$.

d) (BÔNUS) Note que $\widehat{\theta}_n$ também maximiza $M_n(\theta)$, pois podemos reescrever $M_n(\theta) = n^{-1}(\log \mathcal{L}_n(\theta) - \log \mathcal{L}_n(\theta_\star))$, sendo o segundo termo constante em θ. Dessa forma, temos que $M_n(\widehat{\theta}_n) \geq M_n(\theta_\star)$, pela definição do EMV. Vamos escrever a diferença entre $M(\theta_\star)$ e $M(\widehat{\theta}_n)$, de modo a controlar a distância entre $\widehat{\theta}_n$ e $\theta_\star$:

$$\begin{aligned}M(\theta_\star) - M(\widehat{\theta}_n) &= M(\theta_\star) - M(\widehat{\theta}_n) \pm M_n(\theta_\star)\\ &= (\underbrace{M_n(\theta_\star)}_{\leq M_n(\widehat{\theta}_n)} - M(\widehat{\theta}_n)) + (M(\theta_\star) - M_n(\theta_\star))\\ &\leq (M_n(\widehat{\theta}_n) - M(\widehat{\theta}_n)) + (M(\theta_\star) - M_n(\theta_\star))\\ &\leq \sup_{\theta\in\Theta} |M_n(\theta) - M(\theta)| + (M(\theta_\star) - M_n(\theta_\star)).\end{aligned}$$

Note que o primeiro termo converge em probabilidade para zero, por hipótese, enquanto que o segundo converge para zero em probabilidade devido à Lei Fraca dos Grandes Números. Dessa forma, temos que $M(\theta_\star) - M(\widehat{\theta}_n) \xrightarrow{p} 0$. Isso implica que, para todo $\delta > 0$ vale que

$$\mathbb{P}(M(\hat{\theta}_n) < M(\theta_\star) - \delta) \to 0, \text{ quando } n \to +\infty.$$

Fixe $\varepsilon > 0$, e lembre que por hipótese vale que $\sup_{\{\theta \ | \ |\theta-\theta_\star|\geq\varepsilon\}} M(\theta) < M(\theta_\star)$. Ou seja, existe $\delta > 0$ tal que $|\theta - \theta_\star| \geq \varepsilon$ implica que $M(\theta) < M(\theta_\star) - \delta$. Dessa forma, temos:

$$\mathbb{P}(|\widehat{\theta}_n - \theta_\star| \geq \epsilon) \leq \mathbb{P}(M(\widehat{\theta}_n) < M(\theta_\star) - \delta) \to 0, \text{ quando } n \to +\infty,$$

onde a desigualdade vem da implicação discutida logo acima. O resultado obtido é precisamente a definição de convergência em probabilidade de $\widehat{\theta}_n$ para $\theta_\star$.

Questão 2:

a) Note que $\displaystyle\int_{-\infty}^{+\infty} f(x|\theta)\ dx = 1$. Derivando de ambos os lados em relação a θ e supondo condições de regularidade para poder trocar a derivada com a integral, temos que:

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} \frac{\partial}{\partial\theta} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} \frac{\frac{\partial}{\partial\theta} f(x|\theta)}{f(x|\theta)} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} \underbrace{\frac{\partial}{\partial\theta}\log f(x|\theta)}_{=s(\theta;x).} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} s(\theta;x) f(x|\theta)\ dx \\
&= \mathbb{E}_\theta[s(\theta;X)].
\end{aligned}$$

Dessa forma, temos que $\mathbb{E}_\theta[s(\theta;X)] = 0$.

b) Derivando novamente em θ, temos que:

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty} s(\theta;x) f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} \frac{\partial}{\partial\theta}[s(\theta;x) f(x|\theta)]\ dx \\
&= \int_{-\infty}^{+\infty} s'(\theta;x) f(x|\theta) + s(\theta;x)\frac{\partial}{\partial\theta} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} s'(\theta;x) f(x|\theta) + s(\theta;x) \underbrace{\frac{\frac{\partial}{\partial\theta} f(x|\theta)}{f(x|\theta)}}_{=\frac{\partial}{\partial\theta}\log f(x|\theta) = s(\theta;x).} f(x|\theta)\ dx \\
&= \int_{-\infty}^{+\infty} s'(\theta;x) f(x|\theta) + s(\theta;x)^2 f(x|\theta)\ dx \\
&= \mathbb{E}_\theta[s'(\theta;X)] + \mathbb{E}_\theta[s(\theta;X)^2] \\
&= \mathbb{E}_\theta[s'(\theta;X)] + \mathbb{V}_\theta(s(\theta;X)).
\end{aligned}$$

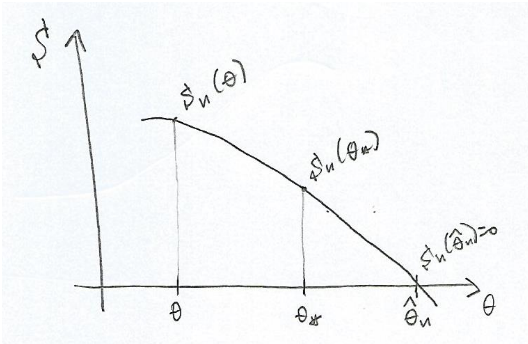
Isso implica que $\mathbb{V}_\theta(s(\theta;X)) = -\mathbb{E}_\theta[s'(\theta;X)]$.

c) Note que o EMV $\widehat{\theta}_n$ satisfaz

$$0 = \frac{\partial}{\partial\theta}\mathcal{L}_n(\widehat{\theta}_n) = \frac{\partial}{\partial\theta}\log\mathcal{L}_n(\widehat{\theta}_n) = S_n(\widehat{\theta}_n).$$

Para n suficientemente grande, temos que $\widehat{\theta}_n \approx \theta_\star$ com alta probabilidade, pois vimos que o EMV é consistente. Dessa forma, a reta secante ao gráfico de S_n que passa pelos pontos $(\theta_\star, S_n(\theta_\star))$ e $(\hat{\theta}_n, S_n(\hat{\theta}_n)) = (\hat{\theta}_n, 0)$ tem inclinação aproximadamente igual à da reta tangente ao gráfico de S_n no ponto $(\theta_\star, S_n(\theta_\star))$:

$$\frac{S_n(\hat{\theta}_n) - S_n(\theta_\star)}{\hat{\theta}_n - \theta_\star} = \frac{-S_n(\theta_\star)}{\hat{\theta}_n - \theta_\star} \approx S_n'(\theta_\star).$$



d) Temos que:

$$\hat{\theta}_n - \theta_\star \approx \frac{-S_n(\theta_\star)}{S_n'(\theta_\star)} = -\frac{\sum_{i=1}^{n} s(\theta_\star; X_i)}{\sum_{i=1}^{n} s'(\theta_\star; X_i)}.$$

Multiplicando ambos os lados por $\sqrt{n}$, temos então que:

$$\sqrt{n}(\hat{\theta}_n - \theta_\star) \approx \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} s(\theta_\star; X_i)}{-\frac{1}{n} \sum_{i=1}^{n} s'(\theta_\star; X_i)}.$$

e) Note que D_n é o negativo da média amostral de $s'(\theta; X_i)$, sendo que cada um desses termos individualmente satisfaz $\mathbb{E}_{\theta_\star}[s'(\theta_\star; X_i)] = -\mathbb{V}_{\theta_\star}(s(\theta_\star; X_i)) =: -I(\theta_\star)$, dita a *informação de Fisher*. Pela Lei Fraca dos Grandes Números, temos que:

$$D_n \xrightarrow{p} \mathbb{E}_{\theta_\star}[D_n] = I(\theta_\star).$$

f) Note que:

$$N_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} s(\theta_\star; X_i) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} s(\theta_\star; X_i) \right),$$

e pelo Teorema Central do Limite, temos que

$$N_n \xrightarrow{d} \sqrt{I(\theta_\star)} Z,$$

onde $Z \sim N(0, 1)$.

g) Pelo Lema de Slutzky, temos que:

$$\sqrt{n}(\hat{\theta}_n - \theta_\star) \xrightarrow{d} \frac{I(\theta_\star) Z}{I(\theta_\star)} = \frac{Z}{\sqrt{I(\theta_\star)}} \sim N\left(0, \frac{1}{I(\theta_\star)}\right).$$

—

Observação: Note que a troca da derivada com a integral e o argumento da reta tangente são pontos imprecisos nessa prova. Tais pontos podem ser tornados rigorosos através de hipóteses adequadas, porém tal prova está fora do escopo do nosso curso. Segue abaixo um enunciado rigoroso:

Teorema (Normalidade Assintótica do EMV)**.** *Assuma que valem as hipóteses abaixo:*

i) $X_1, \ldots, X_n$ *são amostras iid de uma densidade* $f(x|\theta_\star)$*, onde* $\theta_\star \in \Theta \subset \mathbb{R}$ *e* Θ *é um intervalo;*

ii) *(Suavidade) Para todo x no suporte de $f(x|\theta)$ vale que $f(x|\theta) > 0$ e além disso $f(x|\theta)$ é de classe C^2 com respeito a θ;*

iii) *(Momentos) Para todo $\theta \in \Theta$, vale que $\mathbb{E}_\theta[s(\theta; X)] = 0$ e $0 < I(\theta) = \mathbb{V}_\theta(s(\theta; X)) = -\mathbb{E}_\theta[s'(\theta; X)] < +\infty$;*

iv) *(Integrabilidade) Para todo $\theta \in \Theta$, existe $\delta_\theta > 0$ tal que $\mathbb{E}_\theta[\sup_{\{\tau \ | \ |\tau - \theta| \leq \delta_\theta\}} |s'(\tau, X)|] < +\infty$;*

v) *(Consistência) Para todo $\theta \in \Theta$ e todo $\epsilon > 0$, vale que $\mathbb{P}_\theta(|\hat{\theta}_n - \theta| \geq \epsilon) \to 0$, quando $n \to +\infty$.*

Então para todo $\theta_\star$ no interior de Θ o EMV $\hat{\theta}_n$ satisfaz

$$\sqrt{n}(\hat{\theta}_n - \theta_\star) \xrightarrow{d} N\left(0, \frac{1}{I(\theta_\star)}\right).$$