CE085 - Estatística Inferencial
Teoria assintótica da verossimilhança

Prof. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação - LEG
Curso de Bacharelado em Estatatística
Universidade Federal do Paraná - UFPR

17 de setembro de 2018

Conteúdo

Conteúdo

- Distribuição assintótica da função escore.
- Estimador de máxima verossimilhança (MLE).
- Familia exponencial.
- Consistência do MLE.
- Eficiência e normalidade assintótica do MLE.
- Distribuição Weibul.
- Modelos de locação.

Normalidade assintótica da função escore

- Principal resultado

$$U(\theta|\mathbf{Y}) \overset{a}{\sim} N(0, I_E(\theta)).$$

- Demonstração.

$$U(\theta|\mathbf{Y}) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(\theta, Y_i).$$

- A função escore é a soma de v.a iid para um dado θ. Pelas igualdades de Bartlet, temos

$$E(U(\theta|\mathbf{Y})) = 0 \quad \text{e} \quad Var(U(\theta|\mathbf{Y})).$$

- Pelo Teorema Central do Limite, temos

$$\frac{U(\theta|\mathbf{Y}) - E(U(\theta|\mathbf{Y}))}{\sqrt{Var(U(\theta|\mathbf{Y}))}} \overset{a}{\sim} N(0, 1).$$

Estimador de máxima verossimilhança (EMV)

- Estimativa de máxima verossimilhança: Seja $L(\underline{\theta}, \underline{y})$ a função de verossimilhança. O valor $\hat{\underline{\theta}} = \hat{\underline{\theta}}(\underline{y})$ é a estimativa de máxima verossimilhança para $\underline{\theta}$ se $L(\hat{\underline{\theta}}) \geq L(\underline{\theta}), \forall \underline{\theta}$.
- Estimador de máxima verossimilhança: Se $\hat{\underline{\theta}}(\underline{y})$ é a estimativa de máxima verossimilhança, então $\hat{\underline{\theta}}(\underline{Y})$ é o estimador de máxima verossimilhança (EMV).
- Em muitos casos $\hat{\theta}$ é um máximo local no interior de Ω e satisfaz

$$U(\hat{\theta}|\mathbf{Y}) = 0.$$

Propriedades do EMV

- $\hat{\theta}$ é consistente, ou seja,

$$\hat{\theta} \xrightarrow{P} \theta, \quad \text{quando} \quad n \to \infty.$$

Isso significa que

$$P(|\hat{\theta} - \theta| \geq \epsilon) \to 0 \quad \text{quando} \quad n \to \infty.$$

- $\hat{\theta}$ é assintóticamente normal e eficiente, ou seja,

$$\hat{\theta} \overset{a}{\sim} N(\theta, I_E(\theta)^{-1}).$$

- $\hat{\theta}$ é o melhor estimador disponível para θ quando o tamanho da amostra é grande.
- Assintóticamente não-viciado e eficiente.
- Exemplos com dados simulados.

Família exponencial de distribuições

- Uma v.a Y é dita ser da familia exponencial se sua fp ou fdp tem a seguinte forma

$$f(y;\theta) = a(y)\exp^{\theta y - \kappa(\theta)},$$

onde θ é chamado parâmetro canônico e $\kappa(\cdot)$ é uma função conhecida.

- Teorema: $E(Y) = \kappa'(\theta)$ e $Var(Y) = \kappa''(\theta)$.
- Demonstração.

Propriedades do EMV na familia exponencial

- Se Y_i são v.a iid com distribuição de probabilidade pertencente a familia exponencial temos:
  1. O EMV $\hat{\theta}$ é consistente para θ.
  2. O EMV é assintóticamente eficiente e normalmente distribuído.
- Demonstração.

Método delta

- Se uma sequência Y_n satisfaz

$$\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2) \quad \text{quando} \quad n \to \infty,$$

e se $g : \Re \to \Re$ diferenciável em θ e $g'(\theta) \neq 0$, então

$$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2 g'(\theta)^2) \quad \text{quando} \quad n \to \infty.$$

- Demonstração.

Consistência: Caso geral

- Lembre-se do Teorema que deu origen a ideia do EMV.

$$P(l(\theta_0) > l(\theta)) \to 1 \quad \text{quando} \quad n \to \infty.$$

  para $\theta \neq \theta_0$ fixado.
- Teorema (Consistência): Com probabilidade tendendo a 1 quando $n \to \infty$, a verossimilhança tem uma solução $\hat{\theta}$ que é consistente.
- Demonstração.

Consistência: Caso geral

- Seja $\delta > 0$ tal que $\theta_0 - \delta$ e $\theta_0 + \delta$ estão em Ω e defina

$$A_n = \{y : l(\theta_0) > l(\theta_0 - \delta) \quad \text{e} \quad l(\theta_0) > l(\theta + \delta)\}.$$

- Assim, pelo Teorema anterior nos podemos concluir que $P(A_n) \to 1$ quando $n \to \infty$.
- De fato, note que para dois eventos quaisquer A e B, nos temos

$$P(A^c \cap B^c) = 1 - P(A \cup B) \geq 1 - P(A) - P(B).$$

- Isto implica que

$$\begin{aligned} P(A_n) &\geq 1 - P(l(\theta_0) \leq l(\theta_0 - \delta)) - P(l(\theta_0) \leq P(l(\theta_0 + \delta))) \\ &\to 1 \quad \text{quando} \quad n \to \infty. \end{aligned}$$

Consistência: Caso geral

- Assim, para qualquer $y \in A_n$ existe $\hat{\theta}(\delta) = (\theta_0 - \delta, \theta_0 + \delta)$, tal que $l'(\hat{\theta}(\delta)) = 0$ e $\hat{\theta}(\delta)$ é um máximo local de $l(\theta)$. Assim,

$$P(|\hat{\theta}(\delta) - \theta_0| < \delta) \geq P(A_n) \to 1 \quad \text{quando} \quad n \to \infty.$$

- Finalmente, precisamos definir uma sequência que não depende de δ. Seja $\hat{\theta}$ a raiz mais próxima de θ_0. Então, para qualquer $\delta > 0$

$$P(|\hat{\theta} - \theta_0| > \delta) \to 1 \quad \text{quando} \quad n \to \infty.$$

- Note que $\hat{\theta}$ não é necessariamente o EMV, mas um ponto estacionário de $l(\theta)$.
- Corolário: Se $\hat{\theta}$ é único, então é consistente.

Eficiência e normalidade assintótica

- Assumimos que existe uma função $M(y)$ tal que

$$\left| \frac{\partial^3}{\partial \theta^3} \log f(y;\theta) \right| < M(y),$$

e $E(M(y)) < \infty$. Então

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N(0, i_E(\theta)^{-1}) \quad \text{quando} \quad n \to \infty.$$

onde $i_E(\theta)$ é a informação de Fisher para uma observação.
- Demonstração.

Exemplos

- Discuta a obtenção do EMV nos seguintes casos:
  1. Considere a distribuição Weibull com densidade

     $$f(y;\theta) = \theta y^{\theta-1} \exp -y^{y\theta}, \quad \text{paray} > 0.$$

  2. Considere o modelo de locação com pdf dada por

     $$f(y;\theta) = f(y-\theta) \quad \text{para} \quad y \in \Re,$$

     onde $\theta \in \Re$ e f é uma dada pdf.