

Inferência Estatística II

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Programa de Pós-Graduação em Estatística Aplicada e Biometria
Universidade Federal de Viçosa
Campus UFV - Viçosa



- 1 Razão de Verossimilhança Monótona
 - Exemplo 1

- 2 Teorema de Karlin-Rubin
 - Exemplo 1
 - Exemplo 2
 - Exemplo 3

Razão de Verossimilhança Monótona

A Razão de Verossimilhança Monótona (RVM) está relacionada à comparação de dois modelos estatísticos por meio da razão de suas funções de verossimilhança.

Razão de Verossimilhança Monótona

A Razão de Verossimilhança Monótona (RVM) está relacionada à comparação de dois modelos estatísticos por meio da razão de suas funções de verossimilhança.

A razão de verossimilhança é definida como o quociente

$$\frac{L(\theta_0 | T(X))}{L(\theta_1 | T(X))}$$

para dois valores de parâmetro θ_0 e θ_1 . Dizemos que a razão de verossimilhança é **monótona** se, para quaisquer dois valores $\theta_1 > \theta_0$, a razão de verossimilhança for uma função **monótona não-decrescente** em relação a estatística $T(X)$.

Razão de Verossimilhança Monótona

A monotonicidade aqui é uma propriedade importante, pois implica que, se um modelo é mais provável do que outro para uma dada estatística $T(X)$, ele permanecerá mais provável à medida que a estatística $T(X)$ aumenta. Isso é fundamental em testes de hipóteses e inferência estatística, pois garante uma ordem consistente de preferência entre os modelos em estudo.

Razão de Verossimilhança Monótona

Definição 1

Uma família de funções de densidade de probabilidade $\{f_{\mathbf{X}}(\cdot; \theta), \theta \in \Theta\}$, $\Theta \in \mathbb{R}$, é dita ter Razão de Verossimilhança Monótona (RVM) se existir uma estatística $T = t(\mathbf{X})$ tal que, para todo $\theta_1 > \theta_0$,

$$\frac{L(\theta_0 \mid T(\mathbf{X}))}{L(\theta_1 \mid T(\mathbf{X}))}$$

é uma função monótona (não decrescente ou não crescente) em $t(\mathbf{x})$, para $\mathbf{x} \in \{f(\mathbf{x}; \theta_0) > 0 \mid f(\mathbf{x}; \theta_1) > 0\}$, em que $f_{\mathbf{X}}(\mathbf{x}; \theta_0) \neq f_{\mathbf{X}}(\mathbf{x}; \theta_1)$.

Exemplo 1

Seja X_1, \dots, X_n uma amostra de $X \sim \text{Exp}(\theta)$. Segue que,

$$\frac{L(\theta_0 | T(\mathbf{X}))}{L(\theta_1 | T(\mathbf{X}))} = \frac{\theta_1^n}{\theta_0^n} \exp \left(- \left(\frac{1}{\theta_1} - \frac{1}{\theta_0} \right) \sum_{i=1}^n x_i \right)$$

Exemplo 1

Seja X_1, \dots, X_n uma amostra de $X \sim \text{Exp}(\theta)$. Segue que,

$$\frac{L(\theta_0 | T(\mathbf{x}))}{L(\theta_1 | T(\mathbf{x}))} = \frac{\theta_1^n}{\theta_0^n} \exp \left(- \left(\frac{1}{\theta_1} - \frac{1}{\theta_0} \right) \sum_{i=1}^n x_i \right)$$

Para $\theta_1 > \theta_0$ fixados, a função acima é monótona não crescente em $t(\mathbf{x}) = \sum_{i=1}^n x_i$, logo, $\text{Exp}(\theta)$ tem Razão de Verossimilhança Monótona

(RVM) não crescente em $t(\mathbf{x}) = \sum_{i=1}^n x_i$. De outro lado, $\text{Exp}(\theta)$ tem

RVM não decrescente em $t^*(\mathbf{x}) = -\sum_{i=1}^n x_i$.

Exemplo 2

Seja X_1, \dots, X_n uma amostra de $X \sim U(0, \theta)$. Considere $y_n = \max\{X_1, \dots, X_n\}$,

$$L(x; \theta) = L(y; \theta) = \frac{1}{\theta^n} 1_{(0, \theta)}(y_n)$$

Assim, se $\theta_0 < \theta_1$, temos que:

$$\frac{L(y; \theta_0)}{L(y; \theta_1)} = \left(\frac{\theta_1}{\theta_0} \right)^n \frac{1_{(0, \theta_0)}(y_n)}{1_{(0, \theta_1)}(y_n)}$$

Exemplo 2

Portanto, note que, $U(0, \theta)$ tem Razão de Verossimilhança Monótona (RVM) não decrescente em y_n e,

$$g(y_n) = \begin{cases} \left(\frac{\theta_0}{\theta_1}\right)^n, & \text{se } 0 < y_n < \theta_0 \\ \text{Indeterminado}, & \text{se } \theta_0 \leq y_n < \theta_1 \end{cases}$$

Página 350 do Casella e Berger (Tradução)

Muitas famílias de distribuição comuns têm uma RVM. Por exemplo, a distribuição normal (Variância conhecida, média desconhecida), de Poisson, e a Binomial, todas têm. Na verdade, qualquer família exponencial regular com $f(t|\theta) = h(t)c(\theta) \exp w(\theta)t$ tem uma RVM se $w(\theta)$ for uma função não decrescente.

Teorema de Karlin-Rubin

Considere testar $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$. Seja $L(\theta; \mathbf{x})$ a função de verosimilhança de uma distribuição com espaço de parâmetros Θ e estatística suficiente $T = t(\mathbf{x})$ para θ , tal que, para quaisquer θ_0 e θ_1 em Θ tais que $\theta_0 < \theta_1$, a razão de verosimilhança

$$\Lambda(\mathbf{x}; \theta_0, \theta_1) = \frac{L(\theta_0; \mathbf{x})}{L(\theta_1; \mathbf{x})}$$

é uma RVM (Ou seja, é uma função monótona (não decrescente ou não crescente) em $t(\mathbf{x})$). Sob tais condições, para um apropriado valor t_0 ,

$$C = \{\mathbf{x}; T(\mathbf{x}) > t_0\}$$

é uma região crítica para um teste uniformemente mais poderoso para $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ de tamanho $\alpha = P_{\theta_0}(T(\mathbf{x}) > t_0)$.

Observação:

Sob as condições do teorema anterior podemos testar $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$. Nesse caso,

$$C = \{x; T(x) < t_0\}$$

é uma região crítica para um teste uniformemente mais poderoso para $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$ de tamanho $\alpha = P_{\theta_0}(T(x) < t_0)$.

Exemplo 1

Considere $X_1, \dots, X_n \sim \text{Poisson}(\theta)$. Encontre um TUMP de nível α para testar

$$H_0 : \theta \leq \theta_0 \text{ contra } H_1 : \theta > \theta_0$$

Exemplo 1

Considere $X_1, \dots, X_n \sim \text{Poisson}(\theta)$. Encontre um TUMP de nível α para testar

$$H_0 : \theta \leq \theta_0 \text{ contra } H_1 : \theta > \theta_0$$

Considere $0 < \theta_0 < \theta_1 < \infty$ e

$$\frac{L(t; \theta_0)}{L(t; \theta_1)} = e^{n(\theta_1 - \theta_0)} \left(\frac{\theta_0}{\theta_1} \right)^t, \text{ com } t = T(\mathbf{X}) = \sum_{i=1}^n X_i.$$

Sabemos que $T(\mathbf{X}) = \sum_{i=1}^n X_i$ é uma estatística suficiente para θ , como a razão anterior é uma RVM de t , para $\theta_0 < \theta_1$, temos que, pelo teorema de Karlin-Rubin, o teste que rejeita H_0 para toda amostra \mathbf{X} tal que $\sum_{i=1}^n X_i > t_0$ é um teste uniformemente mais poderoso de tamanho $\alpha = P_{\theta_0} \left(\sum_{i=1}^n X_i > t_0 \right)$.

Exemplo2

Suponha

$$H_0 : \theta \leq 1 \text{ contra } H_1 : \theta > 1$$

$$n = 100$$

$$\alpha = 0.05$$

Segue que $\sum_{i=1}^n X_i \sim \text{Poisson}(100\theta)$, sob H_0 , $\sum_{i=1}^n X_i \sim \text{Poisson}(100)$.

No software R, obtemos $t_0 = 117$ e $P_{\theta_0}(\sum_{i=1}^n X_i > t_0) \approx 0,043 \leq \alpha$. Logo, o teste que rejeita H_0 para toda amostra de tamanho $n = 100$ quando $\sum_{i=1}^n X_i > 117$ é um TUMP de nível 0.05. O tamanho do teste é $P_{\theta_0}(\sum_{i=1}^n X_i > 117) \approx 0,043$.

Exemplo 3

Seja X_1, \dots, X_n uma amostra aleatória de uma distribuição exponencial(θ). Encontre um TUMP de

$$H_0 : \theta \geq \theta_0 \text{ contra } H_1 : \theta < \theta_0$$

Exemplo 3

Seja X_1, \dots, X_n uma amostra aleatória de uma distribuição exponencial(θ). Encontre um TUMP de

$$H_0 : \theta \geq \theta_0 \text{ contra } H_1 : \theta < \theta_0$$

Sabemos que $T(\mathbf{X}) = \sum_{i=1}^n X_i$ é uma estatística suficiente para θ e que a distribuição de $T(\mathbf{X}) = \sum_{i=1}^n X_i$ é Gamma(n, θ). Para $0 < \theta_0 < \theta_1 < \infty$

$$\frac{L(t; \theta_0)}{L(t; \theta_1)} = \left(\frac{\theta_1}{\theta_0}\right)^n e^{t\left(\frac{1}{\theta_1} - \frac{1}{\theta_0}\right)}, \quad \frac{1}{\theta_1} - \frac{1}{\theta_0} < 0,$$

é uma RVM de t . Logo, a família gamma tem a propriedade RVM.

Pelo teorema de Karlin-Rubin, o teste que rejeita H_0 para toda amostra \mathbf{X} tal que $\sum_{i=1}^n X_i < 3$ é um teste uniformemente mais poderoso de tamanho $\alpha = P_{\theta_0}(\sum_{i=1}^n X_i < t_0)$.

Pelo teorema de Karlin-Rubin, o teste que rejeita H_0 para toda amostra \mathbf{X} tal que $\sum_{i=1}^n X_i < 3$ é um teste uniformemente mais poderoso de tamanho $\alpha = P_{\theta_0}(\sum_{i=1}^n X_i < t_0)$.

Considere, por exemplo,

$$H_0 : \theta \geq 3 \text{ contra } H_1 : \theta < 3$$

$$n = 5$$

$$\alpha = 0.05$$

Segue que $\sum_{i=1}^n X_i \sim \text{Gamma}(5, \theta)$, sob H_0 , $\sum_{i=1}^n X_i \sim \text{Gamma}(5, 3)$.

No software R, obtemos $t_0 = 3,05$ e $P_{\theta_0}(\sum_{i=1}^n X_i > t_0) \approx 0,05$. Logo, o teste que rejeita H_0 para toda amostra de tamanho $n = 5$ quando $\sum_{i=1}^n X_i < 3,05$ é um TUMP de tamanho 0.05.

Referências I



Casella, George e Roger L Berger (2021). *Statistical inference*. Cengage Learning.



Hogg, RV, J McKean e AT Craig (2019). *Introduction to Mathematical Statistics*.