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

Sumário

1. Introdução
2. Teste da Razão de Verossimilhança para a Distribuição Exponencial
3. Exemplo 2 - Distribuição Normal
4. Teorema do Teste da Razão de Verossimilhança
5. Teste do Tipo Wald
6. Teste do Tipo Escore
7. Exemplos TRV, Wald e Escore

Considere $X_1, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas com função de densidade de probabilidade $f(x;\theta)$ para $\theta \in \Omega$. Considere, ainda, as hipóteses bilaterais:

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$

em que θ_0 é um valor especificado. Lembre-se de que a função de verossimilhança e seu logaritmo são dados por:

$$L(\theta) = \prod_{i=1} f(X_i;\theta) \text{ e } \ell(\theta) = \sum_{i=1} \log f(X_i;\theta)$$

Com $\hat{\theta}$ sendo a estimativa de máxima verossimilhança de θ.

Sabemos, que se θ_0 é o valor verdadeiro de θ, então, assintoticamente, $L(\theta_0)$ é o valor máximo de $L(\theta)$. Considere a razão de duas funções de verossimilhança, a saber,

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})}, \ \hat{\theta} \in \Theta. \tag{1}$$

Sabemos, que se θ_0 é o valor verdadeiro de θ, então, assintoticamente, $L(\theta_0)$ é o valor máximo de $L(\theta)$. Considere a razão de duas funções de verossimilhança, a saber,

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})}, \ \hat{\theta} \in \Theta. \tag{1}$$

Observe que $\Lambda \leq 1$, pois $L(\theta_0)$ é uma verossimilhança restrita a Θ_0 e $L(\hat{\theta})$ é uma verossimilhança irrestrita a Θ. Assim, se θ_0 é o verdadeiro valor do parâmetro, $L(\hat{\theta})$ atingirá $L(\theta_0)$ e $\Lambda = 1$, caso contrário, $L(\theta_0) < L(\hat{\theta})$ e $\Lambda < 1$.

Para um nível de significância especificado α, isso leva à regra de decisão intuitiva:

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \Lambda \leq c,$$

em que c é tal que $\alpha = P_{\theta_0}(\Lambda \leq c)$. Chamamos isso de teste da razão de verossimilhança (TRV).

Teste da Razão de Verossimilhança para a Distribuição Exponencial

Suponha que $X_1, \ldots, X_n$ são variáveis i.i.d com densidade de probabilidade

$$f(x;\theta) = \theta^{-1} e^{-x/\theta}, \quad \text{para} \quad x, \theta > 0$$

As hipóteses são dadas por

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0.$$

A função de verossimilhança pode ser escrita como:

$$L(\theta) = \theta^{-n} e^{-n\bar{X}/\theta}$$

É fácil ver que a estimativa de máxima verossimilhança de θ é $\bar{X}$.

Após alguma simplificação, a estatística do teste da razão de verossimilhança pode ser escrita como:

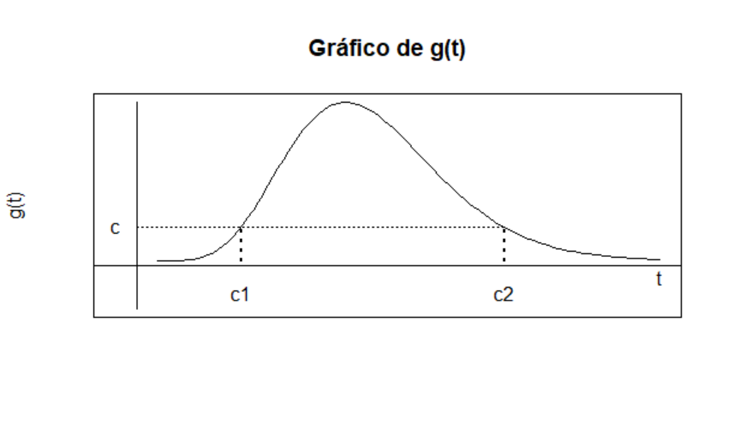
$$\Lambda = e^n \left(\frac{\bar{X}}{\theta_0} \right)^n e^{-n\bar{X}/\theta_0}.$$

Além da constante e^n, a estatística do teste tem a forma:

$$g(t) = t^n e^{-nt}, \quad t > 0$$

em que $t = \bar{X}/\theta_0$.

Como mostra a Figura abaixo, $g(t) \leq c$ se, e somente se, $t \leq c_1$ ou $t \geq c_2$. Isso leva a:
$\Lambda \leq c$ se e somente se $\bar{X}/\theta_0 \leq c_1$ ou $\bar{X}/\theta_0 \geq c_2$.



Notem que $M_{X_i}(t) = (1 - t\theta)^{-1}$, é a função geradora de momentos de X_i, $i = 1, 2, \ldots, n$. Assim, a função geradora de momentos de $Y = 2\sum_{i=1}^{n} X_i/\theta_0$ é dada por $M_Y(t) = (1 - 2t)^{-2n/2}$, que é a função geradora de momentos de uma variável $\chi^2(2n)$.

Notem que $M_{X_i}(t) = (1 - t\theta)^{-1}$, é a função geradora de momentos de X_i, $i = 1, 2, \ldots, n$. Assim, a função geradora de momentos de $Y = 2\sum_{i=1}^{n} X_i/\theta_0$ é dada por $M_Y(t) = (1 - 2t)^{-2n/2}$, que é a função geradora de momentos de uma variável $\chi^2(2n)$.

Ou seja, $\bar{X}/\theta_0 \leq c_1 \iff Y = 2\sum_{i=1}^{n} X_i/\theta_0 \leq k_1 = 2nc_1$ ou

Notem que $M_{X_i}(t) = (1 - t\theta)^{-1}$, é a função geradora de momentos de X_i, $i = 1, 2, \ldots, n$. Assim, a função geradora de momentos de $Y = 2\sum_{i=1}^{n} X_i/\theta_0$ é dada por $M_Y(t) = (1 - 2t)^{-2n/2}$, que é a função geradora de momentos de uma variável $\chi^2(2n)$.

Ou seja, $\bar{X}/\theta_0 \leq c_1 \iff Y = 2\sum_{i=1}^{n} X_i/\theta_0 \leq k_1 = 2nc_1$ ou

$$\bar{X}/\theta_0 \geq c_2 \iff Y = 2\sum_{i=1}^{n} X_i/\theta_0 \geq k_2 = 2nc_2.$$

Baseado nas observações anteriores, podemos utilizar a seguinte regra de decisão para o teste de nível α:

Baseado nas observações anteriores, podemos utilizar a seguinte regra de decisão para o teste de nível α:

Rejeitar H_0 se $(2/\theta_0)\sum_{i=1} X_i \leq \chi^2_{1-\alpha/2}(2n)$ ou $(2/\theta_0)\sum_{i=1} X_i \geq \chi^2_{\alpha/2}(2n)$,

em que $\chi^2_{1-\alpha/2}(2n)$ é o quantil inferior $\alpha/2$ de uma distribuição χ^2 com $2n$ graus de liberdade e $\chi^2_{\alpha/2}(2n)$ é o quantil superior $\alpha/2$ de uma distribuição χ^2 com $2n$ graus de liberdade.

Exemplo 2 - Distribuição Normal

Considere uma amostra aleatória $X_1, X_2, \ldots, X_n$ de uma distribuição $N(\theta, \sigma^2)$, em que $-\infty < \theta < \infty$ e $\sigma^2 > 0$ são conhecidos. Vamos considerar as hipóteses:

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0,$$

em que θ_0 é especificado. A função de verossimilhança é dada por:

$$\begin{aligned} L(\theta) &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right\} \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right\} \exp\left\{-\frac{1}{2\sigma^2} n(\bar{x} - \theta)^2\right\}, \end{aligned}$$

Em $\Omega = \{\theta : -\infty < \theta < \infty\}$, o estimador de máxima verossimilhança é $\hat{\theta} = \bar{x}$, e, portanto, a razão de verossimilhança é:

$$\begin{aligned}\Lambda &= \frac{L(\theta_0)}{L(\hat{\theta})}\\ &= \exp\left\{-\frac{1}{2\sigma^2}n(\bar{X}-\theta_0)^2\right\}.\end{aligned}$$

Em $\Omega = \{\theta : -\infty < \theta < \infty\}$, o estimador de máxima verossimilhança é $\hat{\theta} = \bar{x}$, e, portanto, a razão de verossimilhança é:

$$\begin{aligned}\Lambda &= \frac{L(\theta_0)}{L(\hat{\theta})}\\ &= \exp\left\{-\frac{1}{2\sigma^2} n(\bar{X} - \theta_0)^2\right\}.\end{aligned}$$

Logo, $\Lambda \leq c$ é equivalente a $-2\log\Lambda \geq -2\log c$. No entanto,

$$-2\log\Lambda = \left(\frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}}\right)^2,$$

que segue uma distribuição $\chi^2(1)$ sob H_0.

Portanto, o teste da razão de verossimilhança com nível de significância α afirma que rejeitamos H_0 e aceitamos H_1 quando

$$\left(\frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}}\right)^2 \geq \chi^2_\alpha(1).$$

Teorema do Teste da Razão de Verossimilhança

Teorema 1

Suponha que as condições de regularidade R0 a R5 são satisfeitas. Sob a hipótese nula, $H0 : \theta = \theta_0$, *temos que* $\{-2 \log \Lambda\} \xrightarrow{D} \chi^2(1)$.

Teorema do Teste da Razão de Verossimilhança

Teorema 1

Suponha que as condições de regularidade R0 a R5 são satisfeitas. Sob a hipótese nula, $H0 : \theta = \theta_0$*, temos que* $\{-2\log\Lambda\} \xrightarrow{D} \chi^2(1)$.

Demonstração:

Prova: Expanda a função $\ell(\theta)$ em uma série de Taylor em torno de θ_0 de ordem 1 e avalie-a no estimador de máxima verossimilhança, $\hat{\theta}$. Isso resulta em

$$\ell(\hat{\theta}) = \ell(\theta_0) + (\hat{\theta} - \theta_0)\ell'(\hat{\theta}_0) + \frac{1}{2}(\hat{\theta} - \theta_0)^2\ell''(\theta_n^*), \qquad (2)$$

em que θ_n^* está entre $\hat{\theta}$ e θ_0.

Demonstração:

Como $\hat{\theta} \xrightarrow{P} \theta_0$, segue que $\theta_n^* \xrightarrow{P} \theta_0$. Isso, além do fato de que a função $I'(\theta)$ é contínua e a equação (6.2.22) do Teorema 6.2.2 do livro texto (Hogg 8° edição) implicam que

$$-\frac{1}{n}\ell''(\theta_n^*) \xrightarrow{P} I(\theta_0). \tag{3}$$

Pelo Corolário 6.2.3, do livro texto,

$$\frac{1}{\sqrt{n}}\ell'(\hat{\theta}_0) = \sqrt{n}(\hat{\theta} - \theta_0)I(\theta_0) + R_n, \tag{4}$$

em que $R_n \xrightarrow{P} 0$.

Se substituirmos (3) e (4) na expressão (2) e fizermos algumas simplificações, obtemos

$$-2\log\Lambda = 2(\ell(\hat{\theta}) - \ell(\theta_0)) = \left[\sqrt{nI(\theta_0)}(\hat{\theta} - \theta_0)\right]^2 + R_n^*,$$

em que $R_n^* \xrightarrow{P} 0$. Pelo Teorema 5.2.4 e pelo Teorema 6.2.2, o primeiro termo no lado direito da equação acima converge em distribuição para uma distribuição χ^2 com um grau de liberdade. cqd ■

Logo, defina a estatística de teste $\chi_L^2 = -2\log\Lambda$. Para as hipóteses,

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0,$$

este teorema sugere a regra de decisão:

$$\text{Rejeitar } H0 \text{ em favor de } H1 \text{ se } \chi_L^2 \geq \chi_\alpha^2(1).$$

Pelo último teorema, este teste possui nível assintótico α. Se não conseguirmos obter a estatística de teste ou sua distribuição em uma forma fechada, podemos usar este teste assintótico.

Teste do Tipo Wald

Além do teste da razão de verossimilhança, na prática, são empregados outros dois testes relacionados à verossimilhança. Uma estatística de teste natural é baseada na distribuição assintótica de $\hat{\theta}$. Considere a estatística

$$\chi_W^2 = \left\{ \sqrt{nI(\hat{\theta})}(\hat{\theta} - \theta_0) \right\}^2.$$

como $I(\theta)$ é uma função contínua, $I(\hat{\theta}) \xrightarrow{P} I(\theta_0)$ sob a hipótese nula. Portanto, sob $H0, \chi_W^2$ possui uma distribuição assintótica χ^2 com um grau de liberdade.

Teste do Tipo Wald

Além do teste da razão de verossimilhança, na prática, são empregados outros dois testes relacionados à verossimilhança. Uma estatística de teste natural é baseada na distribuição assintótica de $\hat{\theta}$. Considere a estatística

$$\chi^2_W = \left\{ \sqrt{nI(\hat{\theta})}(\hat{\theta} - \theta_0) \right\}^2.$$

como $I(\theta)$ é uma função contínua, $I(\hat{\theta}) \xrightarrow{P} I(\theta_0)$ sob a hipótese nula. Portanto, sob $H0, \chi^2_W$ possui uma distribuição assintótica χ^2 com um grau de liberdade.

Isso sugere a regra de decisão

$$\text{Rejeitar } H0 \text{ em favor de } H1 \text{ se } \chi^2_W \geq \chi^2_\alpha(1). \tag{5}$$

O teste (5) é frequentemente referido como um teste do tipo Wald, em homenagem a **Abraham Wald** (nasceu em 31 de outubro de 1902, morreu em 1950 de acidente de avião), que foi um estatístico do século XX. Após sua morte, Wald foi criticado por Sir Ronald A. Fisher. O trabalho de Wald foi defendido, posteriormente, por Jerzy Neyman e outros acadêmicos proeminentes.

Teste do Tipo Escore

O terceiro teste é chamado de teste de escores de Rao, em homenagem a **Calyampudi Radhakrishna Rao**(10 de setembro de 1920 - 22 de agosto de 2023). A American Statistical Association o descreveu como "uma lenda viva cujo trabalho influenciou não apenas as estatísticas, mas teve implicações de longo alcance para diversos outros campos." O Times of India listou Rao como um dos 10 maiores cientistas indianos de todos os tempos.

Os escores são os componentes do vetor

$$S(\theta) = \left\{ \frac{\partial \log f(X_1;\theta)}{\partial \theta}, \ldots, \frac{\partial \log f(X_n;\theta)}{\partial \theta} \right\}.$$

Em nossa notação, temos

$$\frac{1}{\sqrt{n}} \ell'(\hat{\theta}_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(X_i;\theta_0)}{\partial \theta}.$$

Defina a estatística

$$\chi_R^2 = \left\{ \frac{\ell'(\theta_0)}{\sqrt{nI(\theta_0)}} \right\}^2.$$

Sob H0, segue da expressão (4) que

$$\chi_R^2 = \chi_W^2 + R_{0n},$$

em que R_{0n} converge para 0 em probabilidade. Portanto, a próxima regra de decisão define um teste.

Defina a estatística

$$\chi_R^2 = \left\{ \frac{\ell'(\theta_0)}{\sqrt{nI(\theta_0)}} \right\}^2 .$$

Sob H0, segue da expressão (4) que

$$\chi_R^2 = \chi_W^2 + R_{0n},$$

em que R_{0n} converge para 0 em probabilidade. Portanto, a próxima regra de decisão define um teste.

$$\text{Rejeite } H0 \text{ em favor de } H1 \text{ se } \chi_R^2 \geq \chi_\alpha^2(1). \tag{6}$$

Exemplos TRV, Wald e Escore

Seja $X_1, \ldots, X_n$ uma amostra aleatória com uma distribuição beta$(\theta, 1)$. Considere as hipóteses:

$$H_0 : \theta = 1 \text{ versus } H_1 : \theta \neq 1$$

Exemplos TRV, Wald e Escore

Seja $X_1, \ldots, X_n$ uma amostra aleatória com uma distribuição beta$(\theta, 1)$. Considere as hipóteses:

$$H_0 : \theta = 1 \text{ versus } H_1 : \theta \neq 1$$

Sob H_0, $f(x;\theta)$ Uniforme$(0,1)$. Lembre-se de que $\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log X_i}$ é o estimador de máxima verossimilhança de θ. Após alguma simplificação, o valor da função de verossimilhança no estimador de máxima verossimilhança é dado por:

$$L(\hat{\theta}) = \left(-\sum_{i=1}^{n} \log X_i\right)^{-n} \exp\left\{n(\log n - 1) - \sum_{i=1}^{n} \log X_i\right\}.$$

Exemplos TRV, Wald e Escore

Além disso, $L(1) = 1$, portanto, a estatística do teste da razão de verossimilhança é $\Lambda = \frac{1}{L(\hat{\theta})}$, de modo que:

$$\chi^2_L = -2\log\Lambda = 2\left(-\sum_{i=1}^{n}\log X_i - n\log\left(-\sum_{i=1}^{n}\log X_i\right) - n + n\log n\right)$$

Lembre-se de que a informação para esta distribuição é $I(\theta) = \theta^{-2}$.

Exemplos TRV, Wald e Escore

Para o teste do tipo Wald, podemos estimar isso consistentemente como $\hat{\theta}^{-2}$. O teste do tipo Wald simplifica-se para:

$$\chi^2_W = \sqrt{\frac{n}{\hat{\theta}^2}}(\hat{\theta} - 1) = n\left(1 - \frac{1}{\hat{\theta}}\right)^2$$

Exemplos TRV, Wald e Escore

Por fim, para o teste do tipo escores, $\ell'(1)$ é dado por:

$$\ell'(1) = \sum_{i=1}^{n} \log X_i + n$$

Portanto, a estatística do teste do tipo escores é:

$$\chi^2_R = \left\{ \frac{\left(\sum_{i=1}^{n} \log X_i + n \right)}{\sqrt{n}} \right\}^2 .$$

Para 🏠

- **Exercícios da seção 6.3:** 6, 9, 10, 11, 12, 13, 16, 18, 19.

Referências I

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