

Testes de hipóteses

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 08 de abril de 2026



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado
- 5 Aplicação
- 6 Referências bibliográficas



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado
- 5 Aplicação
- 6 Referências bibliográficas



Modelo de regressão linear

Suponham que Y_1, Y_2, \dots, Y_n tais que

$$Y_\ell = \mathbf{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \ell = 1, 2, \dots, n, \quad (1)$$



Modelo de regressão linear

Suponham que Y_1, Y_2, \dots, Y_n tais que

$$Y_\ell = \mathbf{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \ell = 1, 2, \dots, n, \quad (1)$$

em que $\mathbf{x}_\ell = (x_{\ell 1}, x_{\ell 2}, \dots, x_{\ell p})^\top$ é conhecido, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos a serem estimados, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ são variáveis aleatórias independentes e com a mesma variância σ^2 , também desconhecida, a ser estimada.



Modelo de regressão linear

Suponham que Y_1, Y_2, \dots, Y_n tais que

$$Y_\ell = \mathbf{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \ell = 1, 2, \dots, n, \quad (1)$$

em que $\mathbf{x}_\ell = (x_{\ell 1}, x_{\ell 2}, \dots, x_{\ell p})^\top$ é conhecido, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos a serem estimados, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ são variáveis aleatórias independentes e com a mesma variância σ^2 , também desconhecida, a ser estimada.



Forma matricial

A Equação (1) é o que nós definimos como **modelo de regressão normal linear** (MNL), podendo ser escrita de forma matricial, da seguinte forma:

$$Y = X\beta + \varepsilon, \quad (2)$$



Forma matricial

A Equação (1) é o que nós definimos como **modelo de regressão normal linear** (MNL), podendo ser escrita de forma matricial, da seguinte forma:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2)$$

em que $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^\top$ e $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^\top$, $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, sendo $\mathbf{0}$ o vetor nulo de dimensão n , \mathbf{I}_n a matriz identidade de ordem n e $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$, a matriz de planejamento.



Forma matricial

A Equação (1) é o que nós definimos como **modelo de regressão normal linear** (MNL), podendo ser escrita de forma matricial, da seguinte forma:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2)$$

em que $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^\top$ e $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^\top$, $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, sendo $\mathbf{0}$ o vetor nulo de dimensão n , \mathbf{I}_n a matriz identidade de ordem n e $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$, a matriz de planejamento.



Método de máxima verossimilhança

O **estimador de máxima verossimilhança (EMV)** de β e σ^2 são dados, respectivamente, por:

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}, \\ \hat{\sigma}_{\text{MV}}^2 &= \frac{1}{n} (\mathbf{Y} - \mathbf{X}\hat{\beta})^\top (\mathbf{Y} - \mathbf{X}\hat{\beta}).\end{aligned}\tag{3}$$

Método de máxima verossimilhança

O **estimador de máxima verossimilhança (EMV)** de β e σ^2 são dados, respectivamente, por:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}, \quad (3)$$
$$\hat{\sigma}_{\text{MV}}^2 = \frac{1}{n} (\mathbf{Y} - \mathbf{X}\hat{\beta})^T (\mathbf{Y} - \mathbf{X}\hat{\beta}).$$

Observação: o EMV de σ^2 também pode ser escrito como função do quadrado médio do resíduo (QMRes).



Método de máxima verossimilhança

O **estimador de máxima verossimilhança (EMV)** de β e σ^2 são dados, respectivamente, por:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}, \quad (3)$$
$$\hat{\sigma}_{\text{MV}}^2 = \frac{1}{n} (\mathbf{Y} - \mathbf{X}\hat{\beta})^T (\mathbf{Y} - \mathbf{X}\hat{\beta}).$$

Observação: o EMV de σ^2 também pode ser escrito como função do quadrado médio do resíduo (QMRes).



Método de máxima verossimilhança

Sob as condições de regularidades, nós temos que:

$$\begin{aligned}\hat{\beta} &\sim \mathcal{N}_p \left\{ \beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1} \right\}, \\ \hat{\sigma}_{\text{MV}}^2 &\sim \mathcal{N} \left\{ \sigma^2, \frac{2(\sigma^2)^2}{n} \right\},\end{aligned}\tag{4}$$

quando o tamanho de amostra é grande.

Método de máxima verossimilhança

Sob as condições de regularidades, nós temos que:

$$\begin{aligned}\hat{\beta} &\sim \mathcal{N}_p \left\{ \beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1} \right\}, \\ \hat{\sigma}_{\text{MV}}^2 &\sim \mathcal{N} \left\{ \sigma^2, \frac{2(\sigma^2)^2}{n} \right\},\end{aligned}\tag{4}$$

quando o tamanho de amostra é grande. Adicionalmente, $\hat{\beta}$ e $\hat{\sigma}_{\text{MV}}^2$ são ortogonais.

Método de máxima verossimilhança

Sob as condições de regularidades, nós temos que:

$$\begin{aligned}\hat{\beta} &\sim \mathcal{N}_p \left\{ \beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1} \right\}, \\ \hat{\sigma}_{\text{MV}}^2 &\sim \mathcal{N} \left\{ \sigma^2, \frac{2(\sigma^2)^2}{n} \right\},\end{aligned}\tag{4}$$

quando o tamanho de amostra é grande. Adicionalmente, $\hat{\beta}$ e $\hat{\sigma}_{\text{MV}}^2$ são ortogonais.



Introdução

Motivação

O resultado (4) é muito importante, pois ele permite a construção de procedimentos inferenciais como intervalos de confiança e testes de hipóteses.



Roteiro

- 1 Introdução
- 2 Teste de hipóteses**
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado
- 5 Aplicação
- 6 Referências bibliográficas



Teste de hipóteses

Seja $c_{mm'}$, o (m, m') -ésimo elemento da matriz $\mathbf{C} = (\mathbf{X}^\top \mathbf{X})^{-1}$, com $m, m' = 1, 2, \dots, p$, i.e., $\text{Cov}(\hat{\beta}_m, \hat{\beta}_{m'}) = \sigma^2 c_{mm'}$. Por (4), nós temos que,

$$\frac{\hat{\beta}_m - \beta_m}{\sqrt{\sigma^2 c_{mm}}} \sim \mathcal{N}(0, 1), \quad (5)$$

com $m = 1, 2, \dots, p$.



Teste de hipóteses

Seja $c_{mm'}$, o (m, m') -ésimo elemento da matriz $\mathbf{C} = (\mathbf{X}^\top \mathbf{X})^{-1}$, com $m, m' = 1, 2, \dots, p$, i.e., $\text{Cov}(\hat{\beta}_m, \hat{\beta}_{m'}) = \sigma^2 c_{mm'}$. Por (4), nós temos que,

$$\frac{\hat{\beta}_m - \beta_m}{\sqrt{\sigma^2 c_{mm}}} \sim \mathcal{N}(0, 1), \quad (5)$$

com $m = 1, 2, \dots, p$.



Testes de hipóteses

Como a estatística em (5) depende do parâmetro desconhecido σ^2 , nós precisamos substituí-lo por um estimativa não viesada, no caso, pelo QMRes, e, assim, nós temos que

$$\frac{\hat{\beta}_m - \beta_m}{\sqrt{\text{QMRes } c_{mm}}} \sim t(n - p), \quad (6)$$

com $m = 1, 2, \dots, p$.



Testes de hipóteses

Como a estatística em (5) depende do parâmetro desconhecido σ^2 , nós precisamos substituí-lo por um estimativa não viesada, no caso, pelo QMRes, e, assim, nós temos que

$$\frac{\hat{\beta}_m - \beta_m}{\sqrt{\text{QMRes } c_{mm}}} \sim t(n - p), \quad (6)$$

com $m = 1, 2, \dots, p$.



Teste de hipóteses para β_m

E assim, a partir de (6), nós podemos construir um procedimento para testar hipóteses sobre o parâmetro β_m . O procedimento é o seguinte:

$$\begin{cases} \mathcal{H} : \beta_m = \beta_m^{(0)} \\ \mathcal{A} : \beta_m \neq \beta_m^{(0)} \end{cases},$$

em que $\beta_m^{(0)}$ é um valor especificado, $m = 1, 2, \dots, p$.

Teste de hipóteses para β_m

E assim, a partir de (6), nós podemos construir um procedimento para testar hipóteses sobre o parâmetro β_m . O procedimento é o seguinte:

$$\begin{cases} \mathcal{H} : \beta_m = \beta_m^{(0)} \\ \mathcal{A} : \beta_m \neq \beta_m^{(0)} \end{cases},$$

em que $\beta_m^{(0)}$ é um valor especificado, $m = 1, 2, \dots, p$.

Teste de hipóteses para β_m

E por (6), nós podemos definir a seguinte estatística do teste como

$$t_c = \frac{\hat{\beta}_m - \beta_m}{\sqrt{\text{QMRes } c_{mm}}}.$$

Teste de hipóteses para β_m

E por (6), nós podemos definir a seguinte estatística do teste como

$$t_c = \frac{\hat{\beta}_m - \beta_m}{\sqrt{\text{QMRes } c_{mm}}}.$$

Sob \mathcal{H} , $t_c \sim t(n - p)$. E, assim, a hipótese nula será rejeitada, para um dado nível nominal α , se $|t_c|$ for maior que o quantil $100(1 - \alpha/2)\%$ de uma $t(n - p)$.

Teste de hipóteses para β_m

E por (6), nós podemos definir a seguinte estatística do teste como

$$t_c = \frac{\hat{\beta}_m - \beta_m}{\sqrt{\text{QMRes } c_{mm}}}.$$

Sob \mathcal{H} , $t_c \sim t(n - p)$. E, assim, a hipótese nula será rejeitada, para um dado nível nominal α , se $|t_c|$ for maior que o quantil $100(1 - \alpha/2)\%$ de uma $t(n - p)$.

Teste de hipóteses para β_m

Observações:

- Esse é um teste parcial (ou marginal), pois β_m depende de todas as outras covariáveis do modelo;



Teste de hipóteses para β_m

Observações:

- Esse é um teste parcial (ou marginal), pois β_m depende de todas as outras covariáveis do modelo;
- Na prática, o teste é feito com $\beta_m^{(0)} = 0$, isto é, para verificar se o coeficiente β_m é significativo, $m = 1, 2, \dots, p$. Esse é um teste da contribuição de x_m dado que os demais preditoras estão no modelo;

Teste de hipóteses para β_m

Observações:

- Esse é um teste parcial (ou marginal), pois β_m depende de todas as outras covariáveis do modelo;
- Na prática, o teste é feito com $\beta_m^{(0)} = 0$, isto é, para verificar se o coeficiente β_m é significativo, $m = 1, 2, \dots, p$. Esse é um teste da contribuição de x_m dado que os demais preditoras estão no modelo;
- Testes unilaterais também podem ser feitos. Bastando ajustar a região de rejeição de \mathcal{H} .



Teste de hipóteses para β_m

Observações:

- Esse é um teste parcial (ou marginal), pois β_m depende de todas as outras covariáveis do modelo;
- Na prática, o teste é feito com $\beta_m^{(0)} = 0$, isto é, para verificar se o coeficiente β_m é significativo, $m = 1, 2, \dots, p$. Esse é um teste da contribuição de x_m dado que os demais preditoras estão no modelo;
- Testes unilaterais também podem ser feitos. Bastando ajustar a região de rejeição de \mathcal{H} .



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F**
- 4 Estatísticas qui-quadrado
- 5 Aplicação
- 6 Referências bibliográficas



Teste de hipóteses para um subconjunto de regressoras

Suponha que nós podemos fazer as seguintes partições no MNL,

$$\beta = (\beta_1^\top, \beta_2^\top)^\top \text{ e } \mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2],$$



Teste de hipóteses para um subconjunto de regressoras

Suponha que nós podemos fazer as seguintes partições no MNL,

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top \text{ e } \mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2],$$

em que $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_q)^\top$ é um vetor q -dimensional, $\boldsymbol{\beta}_2 = (\beta_{q+1}, \dots, \beta_p)^\top$, um vetor de parâmetros $(p-q)$ -dimensional, \mathbf{X}_1 , \mathbf{X}_2 sendo $n \times q$ e $n \times (p-q)$, respectivamente.



Teste de hipóteses para um subconjunto de regressoras

Suponha que nós podemos fazer as seguintes partições no MNL,

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top \text{ e } \mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2],$$

em que $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_q)^\top$ é um vetor q -dimensional, $\boldsymbol{\beta}_2 = (\beta_{q+1}, \dots, \beta_p)^\top$, um vetor de parâmetros $(p-q)$ -dimensional, \mathbf{X}_1 , \mathbf{X}_2 sendo $n \times q$ e $n \times (p-q)$, respectivamente.



Teste de hipóteses para um subconjunto de regressoras

O MNL em (2) pode ser escrito da seguinte forma:

$$Y = X_1\beta_1 + X_2\beta_2 + \varepsilon. \quad (7)$$



Teste de hipóteses para um subconjunto de regressoras

O MNL em (2) pode ser escrito da seguinte forma:

$$\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon. \quad (7)$$

Nós vamos denominar (7) como modelo completo. Agora, se nós supomos que $\beta_1 = \mathbf{0}_q$, nós passamos a ter um modelo reduzido,



Teste de hipóteses para um subconjunto de regressoras

O MNL em (2) pode ser escrito da seguinte forma:

$$\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon. \quad (7)$$

Nós vamos denominar (7) como modelo completo. Agora, se nós supomos que $\beta_1 = \mathbf{0}_q$, nós passamos a ter um modelo reduzido, i.e.,

$$\mathbf{Y} = \mathbf{X}_2\beta_2 + \varepsilon.$$



Teste de hipóteses para um subconjunto de regressoras

O MNL em (2) pode ser escrito da seguinte forma:

$$\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon. \quad (7)$$

Nós vamos denominar (7) como modelo completo. Agora, se nós supomos que $\beta_1 = \mathbf{0}_q$, nós passamos a ter um modelo reduzido, i.e.,

$$\mathbf{Y} = \mathbf{X}_2\beta_2 + \varepsilon.$$



Teste F parcial

Nós podemos utilizar a soma de quadrados extra para construir um procedimento para testar se $\beta_1 = \mathbf{0}_q$, da seguinte forma:

$$\begin{cases} \mathcal{H}: \beta_1 = \mathbf{0}_q \\ \mathcal{A}: \beta_1 \neq \mathbf{0}_q \end{cases} .$$



Teste F parcial

Nós podemos utilizar a soma de quadrados extra para construir um procedimento para testar se $\beta_1 = \mathbf{0}_q$, da seguinte forma:

$$\begin{cases} \mathcal{H} : \beta_1 = \mathbf{0}_q \\ \mathcal{A} : \beta_1 \neq \mathbf{0}_q \end{cases} .$$

Com a estatística do teste é dada por:

$$F_c = \frac{\text{QMReg}(\mathbf{X}_1 | \mathbf{X}_2)}{\text{QMRes}} .$$



Teste F parcial

Nós podemos utilizar a soma de quadrados extra para construir um procedimento para testar se $\beta_1 = \mathbf{0}_q$, da seguinte forma:

$$\begin{cases} \mathcal{H} : \beta_1 = \mathbf{0}_q \\ \mathcal{A} : \beta_1 \neq \mathbf{0}_q \end{cases} .$$

Com a estatística do teste é dada por:

$$F_c = \frac{\text{QMReg}(\mathbf{X}_1 | \mathbf{X}_2)}{\text{QMRes}} .$$



Teste F parcial

Sob \mathcal{H} , $F_c \sim F(q, n - p)$. E assim, nós rejeitaremos \mathcal{H} se $F_c > F(1 - \alpha; q, n - p)$, isto é, a hipótese nula será rejeitada, para um dado nível nominal α , se a estatística do teste for maior que o quantil $100(1 - \alpha)\%$ de uma $F(q, n - p)$.



Teste F geral

Suponham que nós estamos interessados em um procedimento para testar

$$\begin{cases} \mathcal{H}: \mathbf{B}\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)} \\ \mathcal{A}: \mathbf{B}\boldsymbol{\beta} \neq \boldsymbol{\beta}^{(0)} \end{cases},$$

em que \mathbf{B} é uma matriz $s \times p$, com posto igual a $s \leq p$.



Teste F geral

Suponham que nós estamos interessados em um procedimento para testar

$$\left\{ \begin{array}{l} \mathcal{H}: \mathbf{B}\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)} \\ \mathcal{A}: \mathbf{B}\boldsymbol{\beta} \neq \boldsymbol{\beta}^{(0)} \end{array} \right. ,$$

em que \mathbf{B} é uma matriz $s \times p$, com posto igual a $s \leq p$.



Teste F geral

Exemplo 1. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H} : \beta_1 = \beta_2 \\ \mathcal{A} : \beta_1 \neq \beta_2 \end{array} \right. ,$$

Teste F geral

Exemplo 1. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H} : \beta_1 = \beta_2 \\ \mathcal{A} : \beta_1 \neq \beta_2 \end{array} \right. ,$$

nós teríamos:

Teste F geral

Exemplo 1. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H} : \beta_1 = \beta_2 \\ \mathcal{A} : \beta_1 \neq \beta_2 \end{array} \right. ,$$

nós teríamos:

$$B = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}.$$

Teste F geral

Exemplo 1. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H} : \beta_1 = \beta_2 \\ \mathcal{A} : \beta_1 \neq \beta_2 \end{array} \right. ,$$

nós teríamos:

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}.$$

Teste F geral

Exemplo 2. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H}: \beta_1 = \beta_2 \text{ e } \beta_3 = 0 \\ \mathcal{A}: \text{Pelo menos uma igualdade não é verdadeira} \end{array} \right. ,$$

Teste F geral

Exemplo 2. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H}: \beta_1 = \beta_2 \text{ e } \beta_3 = 0 \\ \mathcal{A}: \text{ Pelo menos uma igualdade não é verdadeira} \end{array} \right. ,$$

nós teríamos:

$$B = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Teste F geral

Exemplo 2. Se $Y_\ell = \beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \beta_3 x_{\ell 3}$ e o interesse for em testar

$$\left\{ \begin{array}{l} \mathcal{H}: \beta_1 = \beta_2 \text{ e } \beta_3 = 0 \\ \mathcal{A}: \text{ Pelo menos uma igualdade não é verdadeira} \end{array} \right. ,$$

nós teríamos:

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Teste F geral

O procedimento para testar:

$$\left\{ \begin{array}{l} \mathcal{H} : \mathbf{B}\beta = \beta^{(0)} \\ \mathcal{A} : \mathbf{B}\beta \neq \beta^{(0)} \end{array} \right. ,$$

Teste F geral

O procedimento para testar:

$$\begin{cases} \mathcal{H}: \mathbf{B}\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)} \\ \mathcal{A}: \mathbf{B}\boldsymbol{\beta} \neq \boldsymbol{\beta}^{(0)} \end{cases},$$

tem a estatística do teste dada por:

$$F_c = \frac{1}{s\hat{\sigma}^2} \left(\mathbf{B}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{(0)} \right)^\top \left[\mathbf{B} \left(\mathbf{X}^\top \mathbf{X} \right)^{-1} \mathbf{B}^\top \right]^{-1} \left(\mathbf{B}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{(0)} \right).$$

Teste F geral

O procedimento para testar:

$$\begin{cases} \mathcal{H} : \mathbf{B}\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)} \\ \mathcal{A} : \mathbf{B}\boldsymbol{\beta} \neq \boldsymbol{\beta}^{(0)} \end{cases},$$

tem a estatística do teste dada por:

$$F_c = \frac{1}{s\hat{\sigma}^2} \left(\mathbf{B}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{(0)} \right)^\top \left[\mathbf{B} \left(\mathbf{X}^\top \mathbf{X} \right)^{-1} \mathbf{B}^\top \right]^{-1} \left(\mathbf{B}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{(0)} \right).$$

Teste F geral

Sob \mathcal{H} , $F_c \sim F(s, n-p)$. E assim, nós rejeitaremos \mathcal{H} se $F_c > F(1-\alpha; s, n-p)$, isto é, a hipótese nula será rejeitada, para um dado nível nominal α , se a estatística do teste for maior que o quantil $100(1-\alpha)\%$ de uma $F(s, n-p)$.



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado**
- 5 Aplicação
- 6 Referências bibliográficas



Estatísticas qui-quadrado

Suponham as mesmas partições discutidas na seção anterior e o nosso interesse é testar:

$$\begin{cases} \mathcal{H} : \beta_1 = \beta_1^{(0)} \\ \mathcal{A} : \beta_1 \neq \beta_1^{(0)} \end{cases},$$

Estatísticas qui-quadrado

Suponham as mesmas partições discutidas na seção anterior e o nosso interesse é testar:

$$\begin{cases} \mathcal{H} : \beta_1 = \beta_1^{(0)} \\ \mathcal{A} : \beta_1 \neq \beta_1^{(0)} \end{cases},$$

sendo $\beta_1^{(0)}$ um vetor especificado. β_1 e β_2 são denominados como vetores de parâmetros de interesse e perturbação, respectivamente.



Estatísticas qui-quadrado

Suponham as mesmas partições discutidas na seção anterior e o nosso interesse é testar:

$$\left\{ \begin{array}{l} \mathcal{H} : \beta_1 = \beta_1^{(0)} \\ \mathcal{A} : \beta_1 \neq \beta_1^{(0)} \end{array} \right. ,$$

sendo $\beta_1^{(0)}$ um vetor especificado. β_1 e β_2 são denominados como vetores de parâmetros de interesse e perturbação, respectivamente.



Estatísticas qui-quadrado

Sejam $\hat{\beta} = (\hat{\beta}_1^\top, \hat{\beta}_2^\top)^\top$ e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$, as estimativas de máxima verossimilhança irrestrita e restrita de β a \mathcal{H} , respectivamente.

A partição $\beta = (\beta_1^\top, \beta_2^\top)^\top$ conduz às correspondentes partições no vetor escore $\mathbf{U}_\beta = (\mathbf{U}_{\beta_1}^\top, \mathbf{U}_{\beta_2}^\top)^\top$, com

Estatísticas qui-quadrado

Sejam $\hat{\beta} = (\hat{\beta}_1^\top, \hat{\beta}_2^\top)^\top$ e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$, as estimativas de máxima verossimilhança irrestrita e restrita de β a \mathcal{H} , respectivamente.

A partição $\beta = (\beta_1^\top, \beta_2^\top)^\top$ conduz às correspondentes partições no vetor escore $\mathbf{U}_\beta = (\mathbf{U}_{\beta_1}^\top, \mathbf{U}_{\beta_2}^\top)^\top$, com

$$\mathbf{U}_{\beta_1} = \frac{1}{\sigma^2} \mathbf{X}_1^\top (\mathbf{Y} - \mathbf{X}_1 \beta_1) \text{ e } \mathbf{U}_{\beta_2} = \frac{1}{\sigma^2} \mathbf{X}_2^\top (\mathbf{Y} - \mathbf{X}_2 \beta_2),$$

Estatísticas qui-quadrado

Sejam $\hat{\beta} = (\hat{\beta}_1^\top, \hat{\beta}_2^\top)^\top$ e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$, as estimativas de máxima verossimilhança irrestrita e restrita de β a \mathcal{H} , respectivamente.

A partição $\beta = (\beta_1^\top, \beta_2^\top)^\top$ conduz às correspondentes partições no vetor escore $\mathbf{U}_\beta = (\mathbf{U}_{\beta_1}^\top, \mathbf{U}_{\beta_2}^\top)^\top$, com

$$\mathbf{U}_{\beta_1} = \frac{1}{\sigma^2} \mathbf{X}_1^\top (\mathbf{Y} - \mathbf{X}_1 \beta_1) \text{ e } \mathbf{U}_{\beta_2} = \frac{1}{\sigma^2} \mathbf{X}_2^\top (\mathbf{Y} - \mathbf{X}_2 \beta_2),$$



Estatísticas qui-quadrado

na matriz de informação de Fisher e na sua inversa, dadas por

$$\mathbf{K}_\beta = \begin{pmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{pmatrix} = \frac{1}{\sigma^2} \begin{pmatrix} \mathbf{x}_1^\top \mathbf{x}_1 & \mathbf{x}_1^\top \mathbf{x}_2 \\ \mathbf{x}_2^\top \mathbf{x}_1 & \mathbf{x}_2^\top \mathbf{x}_2 \end{pmatrix} \quad (8)$$

$$\text{e } \mathbf{K}_\beta^{-1} = \sigma^2 \begin{pmatrix} \mathbf{K}^{11} & \mathbf{K}^{12} \\ \mathbf{K}^{21} & \mathbf{K}^{22} \end{pmatrix}.$$

A matriz inversa \mathbf{K}_β^{-1} pode ser obtida usando propriedades de inversão para matrizes bloco-diagonais (Lu e Shiu, 2002).



Estatísticas qui-quadrado

na matriz de informação de Fisher e na sua inversa, dadas por

$$\mathbf{K}_\beta = \begin{pmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{pmatrix} = \frac{1}{\sigma^2} \begin{pmatrix} \mathbf{x}_1^\top \mathbf{x}_1 & \mathbf{x}_1^\top \mathbf{x}_2 \\ \mathbf{x}_2^\top \mathbf{x}_1 & \mathbf{x}_2^\top \mathbf{x}_2 \end{pmatrix} \quad (8)$$

$$\text{e } \mathbf{K}_\beta^{-1} = \sigma^2 \begin{pmatrix} \mathbf{K}^{11} & \mathbf{K}^{12} \\ \mathbf{K}^{21} & \mathbf{K}^{22} \end{pmatrix}.$$

A matriz inversa \mathbf{K}_β^{-1} pode ser obtida usando propriedades de inversão para matrizes bloco-diagonais (Lu e Shiou, 2002).



Estatísticas qui-quadrado

As estatísticas da razão de verossimilhanças (Wilks, 1938), Wald (Wald, 1943), escore (Rao, 1948) e gradiente (Terrell, 2002) para testar \mathcal{H} são definidas, respectivamente, como

$$\begin{aligned} S_1 &= 2 \left[\ell(\hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}^2) - \ell(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\sigma}^2) \right], \\ S_2 &= \left(\hat{\beta}_1 - \beta_1^{(0)} \right)^\top \left(\hat{\mathbf{K}}^{11} \right)^{-1} \left(\hat{\beta}_1 - \beta_1^{(0)} \right), \\ S_3 &= \tilde{\mathbf{U}}_{\beta_1}^\top \tilde{\mathbf{K}}^{11} \tilde{\mathbf{U}}_{\beta_1}, \\ S_4 &= \tilde{\mathbf{U}}_{\beta_1}^\top \left(\hat{\beta}_1 - \beta_1^{(0)} \right). \end{aligned} \quad (9)$$

Estatísticas qui-quadrado

As estatísticas da razão de verossimilhanças (Wilks, 1938), Wald (Wald, 1943), escore (Rao, 1948) e gradiente (Terrell, 2002) para testar \mathcal{H} são definidas, respectivamente, como

$$\begin{aligned}S_1 &= 2 \left[\ell(\hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}^2) - \ell(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\sigma}^2) \right], \\S_2 &= \left(\hat{\beta}_1 - \beta_1^{(0)} \right)^\top \left(\hat{\mathbf{K}}^{11} \right)^{-1} \left(\hat{\beta}_1 - \beta_1^{(0)} \right), \\S_3 &= \tilde{\mathbf{U}}_{\beta_1}^\top \tilde{\mathbf{K}}^{11} \tilde{\mathbf{U}}_{\beta_1}, \\S_4 &= \tilde{\mathbf{U}}_{\beta_1}^\top \left(\hat{\beta}_1 - \beta_1^{(0)} \right).\end{aligned}\tag{9}$$



Estatísticas qui-quadrado

Sob \mathcal{H} , as quatro estatísticas têm distribuição qui-quadrado com q graus de liberdade. A hipótese nula será rejeitada, para um dado nível nominal α e uma estatística especificada, se a estatística do teste for maior que o quantil $100(1 - \alpha)\%$ de uma χ_q^2 .



Estatísticas qui-quadrado

Buse (1982) apresenta de forma didática a interpretação geométrica dos testes da razão de verossimilhanças, escore e Wald para o caso de hipóteses simples, mais recentemente, Montoril e Souza (2013) fazem a mesma análise incluindo a estatística gradiente.



Estatísticas qui-quadrado

Especificamente, para o MNL as quatro estatísticas, apresentadas em (9), podem ser escritas da seguinte forma:

$$\begin{aligned}S_1 &= n \log \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) + \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\beta})^\top (\mathbf{Y} - \mathbf{X}\tilde{\beta}) - \frac{1}{\hat{\sigma}^2} (\mathbf{Y} - \mathbf{X}\hat{\beta})^\top (\mathbf{Y} - \mathbf{X}\hat{\beta}), \\S_2 &= \frac{1}{\hat{\sigma}^2} (\hat{\beta}_1 - \beta_1^{(0)})^\top (\hat{\mathbf{R}}^\top \hat{\mathbf{R}}) (\hat{\beta}_1 - \beta_1^{(0)}), \\S_3 &= \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\beta})^\top \mathbf{X}_1 (\tilde{\mathbf{R}}^\top \tilde{\mathbf{R}})^{-1} \mathbf{X}_1^\top (\mathbf{Y} - \mathbf{X}\tilde{\beta}), \\S_4 &= \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\beta})^\top \mathbf{X}_1 (\hat{\beta}_1 - \beta_1^{(0)}),\end{aligned}\tag{10}$$



Estatísticas qui-quadrado

Especificamente, para o MNL as quatro estatísticas, apresentadas em (9), podem ser escritas da seguinte forma:

$$\begin{aligned}S_1 &= n \log \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) + \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}})^\top (\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) - \frac{1}{\hat{\sigma}^2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^\top (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}), \\S_2 &= \frac{1}{\hat{\sigma}^2} (\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)})^\top (\hat{\mathbf{R}}^\top \hat{\mathbf{R}}) (\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)}), \\S_3 &= \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}})^\top \mathbf{X}_1 (\tilde{\mathbf{R}}^\top \tilde{\mathbf{R}})^{-1} \mathbf{X}_1^\top (\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}), \\S_4 &= \frac{1}{\tilde{\sigma}^2} (\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}})^\top \mathbf{X}_1 (\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)}),\end{aligned}\tag{10}$$

Estatísticas qui-quadrado

em que $\mathbf{R} = \mathbf{X}_1 - \mathbf{X}_2\mathbf{C}$, com $\mathbf{C} = (\mathbf{X}_2^\top \mathbf{X}_2)^{-1} \mathbf{X}_2^\top \mathbf{X}_1$. Como não há parâmetros envolvidos em \mathbf{R} , $\mathbf{R} = \hat{\mathbf{R}} = \tilde{\mathbf{R}}$.

A matriz de dimensão $q \times (p - q)$ \mathbf{C} pode ser interpretada como coeficientes de regressão estimados, obtido de um MNL entre as colunas de \mathbf{X}_1 e a matriz planejamento \mathbf{X}_2 .



Estatísticas qui-quadrado

em que $\mathbf{R} = \mathbf{X}_1 - \mathbf{X}_2\mathbf{C}$, com $\mathbf{C} = (\mathbf{X}_2^\top \mathbf{X}_2)^{-1} \mathbf{X}_2^\top \mathbf{X}_1$. Como não há parâmetros envolvidos em \mathbf{R} , $\mathbf{R} = \hat{\mathbf{R}} = \tilde{\mathbf{R}}$.

A matriz de dimensão $q \times (p - q)$ \mathbf{C} pode ser interpretada como coeficientes de regressão estimados, obtido de um MNL entre as colunas de \mathbf{X}_1 e a matriz planejamento \mathbf{X}_2 .



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado
- 5 Aplicação**
- 6 Referências bibliográficas



Aplicação 1. Conjunto de dados *Westwood Company* (Neter et al., 1983, p. 36). Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = 10 + 2x_{\ell 2},$$

em que Y_ℓ : horas trabalhadas, $x_{\ell 2}$: tamanho do lote, $\ell = 1, 2, \dots, 10$.

Aplicação 1. Conjunto de dados *Westwood Company* (Neter et al., 1983, p. 36). Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = 10 + 2x_{\ell 2},$$

em que Y_ℓ : horas trabalhadas, $x_{\ell 2}$: tamanho do lote, $\ell = 1, 2, \dots, 10$.

Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 1: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	10	2,503	3,995
β_2	2	0,047	42,583

Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 1: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	10	2,503	3,995
β_2	2	0,047	42,583

Para $\alpha = 0,05$, o quantil 95% de uma $t(8) = 2,31$. Logo, os parâmetros β_1 e β_2 são significativos.



Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 1: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	10	2,503	3,995
β_2	2	0,047	42,583

Para $\alpha = 0,05$, o quantil 95% de uma $t(8) = 2,31$. Logo, os parâmetros β_1 e β_2 são significativos.



Aplicação

Aplicação 2. (Neter et al., 1983, p. 247) A *Zarthan Company* vende cremes especiais. E estes são vendidos em 15 distritos.

Com o interesse em prever o número de vendas (Y , dúzia de potes), foi verificado o tamanho da população (x_2 , por mil habitantes) e a renda per capita (x_3 , em dólares) de cada um dos 15 distritos.



Aplicação

Aplicação 2. (Neter et al., 1983, p. 247) A *Zarthan Company* vende cremes especiais. E estes são vendidos em 15 distritos.

Com o interesse em prever o número de vendas (Y , dúzia de potes), foi verificado o tamanho da população (x_2 , por mil habitantes) e a renda per capita (x_3 , em dólares) de cada um dos 15 distritos.



Aplicação

Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = 3,453 + 0,496x_{\ell 2} + 0,009x_{\ell 3},$$

em que Y_ℓ : dúzias vendidas, $x_{\ell 2}$: tamanho da população (por mil habitantes), $x_{\ell 3}$: renda per capita (em dólares), $\ell = 1, 2, \dots, 15$.

Aplicação

Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = 3,453 + 0,496x_{\ell 2} + 0,009x_{\ell 3},$$

em que Y_ℓ : dúzias vendidas, $x_{\ell 2}$: tamanho da população (por mil habitantes), $x_{\ell 3}$: renda per capita (em dólares), $\ell = 1, 2, \dots, 15$.



Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 2: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	3,453	2,431	1,420
β_2	0,496	0,006	81,924
β_3	0,009	0,001	9,502

Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 2: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	3,453	2,431	1,420
β_2	0,496	0,006	81,924
β_3	0,009	0,001	9,502

Para $\alpha = 0,05$, o quantil 95% de uma $t(12) = 2,179$. Logo, os parâmetros β_2 e β_3 são significativos, enquanto, β_1 não é.



Exemplo

Para os dados, utilizando (4) e (6), nós temos que:

Tabela 2: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_1	3,453	2,431	1,420
β_2	0,496	0,006	81,924
β_3	0,009	0,001	9,502

Para $\alpha = 0,05$, o quantil 95% de uma $t(12) = 2,179$. Logo, os parâmetros β_2 e β_3 são significativos, enquanto, β_1 não é.



Roteiro

- 1 Introdução
- 2 Teste de hipóteses
- 3 Estatísticas F
- 4 Estatísticas qui-quadrado
- 5 Aplicação
- 6 Referências bibliográficas



Referências bibliográficas I

- Buse, A. (1982), 'The likelihood ratio, wald and lagrange multiplier tests: an expository note', *The American Statistician* **36**(3), 153–157.
- Lu, T.-T. e Shiou, S.-H. (2002), 'Inverses of 2×2 block matrices', *Computers & Mathematics with Applications* **43**(1), 119–129.
- Montoril, M. H. e Souza, E. A. (2013), 'Estatística gradiente: propriedades e aplicações', *Revista Brasileira de Biometria* **31**(1), 43–60.
- Neter, J., Wasserman, W. e Kutner, M. H. (1983), *Applied linear regression models*, Richard D. Irwin Inc, Homewood, Illinois.



Referências bibliográficas II

- Rao, C. R. (1948), 'Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation', *Proceedings of the Cambridge Philosophical Society* **44**(1), 50–57.
- Terrell, G. R. (2002), 'The gradient statistic', *Computing Science and Statistics* **34**, 206–215.
- Wald, A. (1943), 'Test of statistical hypotheses concerning several parameter when the number of observations is large', *Transactions of the American Mathematical Society* **54**(3), 426–482.



Referências bibliográficas III

Wilks, S. S. (1938), 'The large-sample distribution of likelihood ratio for testing composite hypotheses', *The Annals of Mathematical Statistics* **9**(1), 60–62.



Obrigado!

✉ tiago.magalhaes@ufjf.br

📄 ufjf.br/tiago_magalhaes

🌐 Departamento de Estatística, Sala 319

