Estimação

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 20 de março de 2024

Roteiro

Roteiro

Modelo de regressão linear

Suponham que $Y_1, Y_2, \ldots, Y_n$ tais que

$$Y_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \tag{1}$$

Modelo de regressão linear

Suponham que $Y_1, Y_2, \ldots, Y_n$ tais que

$$Y_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \tag{1}$$

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

Modelo de regressão linear

Suponham que $Y_1, Y_2, \ldots, Y_n$ tais que

$$Y_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell, \tag{1}$$

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

Forma matricial

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

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{\varepsilon}, \tag{2}$$

Forma matricial

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

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

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

Forma matricial

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

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{\varepsilon}, \tag{2}$$

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

Forma matricial

ou

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}.$$

A matriz $\boldsymbol{X}$ tem n linhas (tamanho da amostra) e p colunas (número de variáveis preditoras), de posto p ($n \geq p$), $\boldsymbol{X}$ é chamada de matriz de planejamento.

Forma matricial

ou

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}.$$

A matriz $\boldsymbol{X}$ tem n linhas (tamanho da amostra) e p colunas (número de variáveis preditoras), de posto p ($n \geq p$), $\boldsymbol{X}$ é chamada de matriz de planejamento.

Modelo de regressão normal linear

Para que procedimentos inferenciais possam ser aplicados, nós precisamos assumir que, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\ell = 1, 2, \ldots, n$ e $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n)$, em (1) e (2), respectivamente. O MRL passa a ser denominado de **modelo de regressão normal linear** (MNL).

Modelo de regressão normal linear

Para que procedimentos inferenciais possam ser aplicados, nós precisamos assumir que, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\ell = 1, 2, \ldots, n$ e $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n)$, em (1) e (2), respectivamente. O MRL passa a ser denominado de **modelo de regressão normal linear** (MNL).

Roteiro

Método dos mínimos quadrados

Definição

Os estimadores de mínimos quadrados (EMQ) obtém estimativas para os parâmetros desconhecidos **minimizando** a soma de quadrado dos erros.

Método dos mínimos quadrados

Definição

Os estimadores de mínimos quadrados (EMQ) obtém estimativas para os parâmetros desconhecidos **minimizando** a soma de quadrado dos erros.

Método dos mínimos quadrados

Se $\widehat{\boldsymbol{\beta}} = \left(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p\right)^\top$ é o EMQ de $\boldsymbol{\beta}$, ele deve ser tal que minimize

$$Q = \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p})\right]^2 . \tag{3}$$

Método dos mínimos quadrados

Se $\widehat{\boldsymbol{\beta}} = \left(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p\right)^\top$ é o EMQ de $\boldsymbol{\beta}$, ele deve ser tal que minimize

$$Q = \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p})\right]^2 . \qquad (3)$$

Método dos mínimos quadrados

Derivando (3), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial Q}{\partial \beta_m} = -2 \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \qquad (4)$$

Método dos mínimos quadrados

Derivando (3), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial Q}{\partial \beta_m} = -2 \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \qquad (4)$$

Igualando (4) a zero,

Método dos mínimos quadrados

Derivando (3), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial Q}{\partial \beta_m} = -2 \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \tag{4}$$

Igualando (4) a zero, nós temos que

$$\hat{\beta}_1 \sum_{\ell=1}^{n} x_{\ell 1} x_{\ell m} + \cdots + \hat{\beta}_m \sum_{\ell=1}^{n} x_{\ell m}^2 + \cdots + \hat{\beta}_p \sum_{\ell=1}^{n} x_{\ell p} x_{\ell m} = \sum_{\ell=1}^{n} x_{\ell m} Y_\ell. \tag{5}$$

Método dos mínimos quadrados

Derivando (3), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial Q}{\partial \beta_m} = -2 \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \tag{4}$$

Igualando (4) a zero, nós temos que

$$\hat{\beta}_1 \sum_{\ell=1}^{n} x_{\ell 1} x_{\ell m} + \cdots + \hat{\beta}_m \sum_{\ell=1}^{n} x_{\ell m}^2 + \cdots + \hat{\beta}_p \sum_{\ell=1}^{n} x_{\ell p} x_{\ell m} = \sum_{\ell=1}^{n} x_{\ell m} Y_\ell. \tag{5}$$

Método dos mínimos quadrados

As p equações em (5) são chamadas de **equações normais**. E o **estimador de mínimos quadrados** é dado por:

Método dos mínimos quadrados

As p equações em (5) são chamadas de **equações normais**. E o **estimador de mínimos quadrados** é dado por:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{6}$$

Método dos mínimos quadrados

As p equações em (5) são chamadas de **equações normais**. E o **estimador de mínimos quadrados** é dado por:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{6}$$

Método dos mínimos quadrados

Uma alternativa para a obtenção de (6), é observar que (3) pode ser escrito como:

$$Q = (\boldsymbol{Y} - \boldsymbol{X}\beta)^\top(\boldsymbol{Y} - \boldsymbol{X}\beta).$$

Método dos mínimos quadrados

Uma alternativa para a obtenção de (6), é observar que (3) pode ser escrito como:

$$Q = (\boldsymbol{Y} - \boldsymbol{X}\beta)^\top (\boldsymbol{Y} - \boldsymbol{X}\beta).$$

e derivar Q em relação à β.

Método dos mínimos quadrados

Uma alternativa para a obtenção de (6), é observar que (3) pode ser escrito como:

$$Q = (\boldsymbol{Y} - \boldsymbol{X}\beta)^\top(\boldsymbol{Y} - \boldsymbol{X}\beta).$$

e derivar Q em relação à $\boldsymbol{\beta}$.

Método dos mínimos quadrados

Observação

O EMQ existirá se existir a inversa $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$.

Método dos mínimos quadrados

Observação

O EMQ existirá se existir a inversa $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$. E a matriz $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$ existe se as variáveis preditoras forem linearmente independentes, isto é, se nenhuma coluna da matriz $\boldsymbol{X}$ é uma combinação linear das outras.

Método dos mínimos quadrados

Observação

O EMQ existirá se existir a inversa $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$. E a matriz $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$ existe se as variáveis preditoras forem linearmente independentes, isto é, se nenhuma coluna da matriz $\boldsymbol{X}$ é uma combinação linear das outras.

Propriedades do EMQ

A esperança de $\widehat{\beta}$ é dada por

$$
\begin{aligned}
\mathbb{E}(\widehat{\beta}) &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}(\boldsymbol{Y}) = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}(\boldsymbol{X}\beta + \varepsilon) \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \left[\boldsymbol{X}\beta + \mathbb{E}(\varepsilon)\right] \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\beta = \beta.
\end{aligned}
$$

Propriedades do EMQ

A esperança de $\widehat{\beta}$ é dada por

$$
\begin{aligned}
\mathbb{E}(\widehat{\beta}) &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}(\boldsymbol{Y}) = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}(\boldsymbol{X}\beta + \varepsilon) \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \left[\boldsymbol{X}\beta + \mathbb{E}(\varepsilon)\right] \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\beta = \beta.
\end{aligned}
$$

Propriedades do EMQ

E a variância de $\widehat{\boldsymbol{\beta}}$ é dada por

$$\begin{aligned}
\text{Var}(\widehat{\boldsymbol{\beta}}) &= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\text{Var}(\boldsymbol{Y})\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\text{Var}(\boldsymbol{X}\beta+\varepsilon)\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\text{Var}(\varepsilon)\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\sigma^2\boldsymbol{I}_n\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} \\
&= \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}.
\end{aligned}$$

Propriedades do EMQ

E a variância de $\widehat{\beta}$ é dada por

$$
\begin{aligned}
\mathrm{Var}(\widehat{\beta}) &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathrm{Var}(\boldsymbol{Y})\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathrm{Var}(\boldsymbol{X}\beta + \boldsymbol{\varepsilon})\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathrm{Var}(\boldsymbol{\varepsilon})\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \sigma^2 \boldsymbol{I}_n \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \\
&= \sigma^2 (\boldsymbol{X}^\top \boldsymbol{X})^{-1}.
\end{aligned}
$$

Propriedades do EMQ

Pelo Teorema de Gauss-Markov (Plackett, 1949), os EMQ

- são não viesados;

Propriedades do EMQ

Pelo Teorema de Gauss-Markov (Plackett, 1949), os EMQ

- são não viesados;
- e têm menor variância entre todos os estimadores lineares não viesados.

Propriedades do EMQ

Pelo Teorema de Gauss-Markov (Plackett, 1949), os EMQ

- são não viesados;
- e têm menor variância entre todos os estimadores lineares não viesados.

Valores ajustados

E assim, o valor ajustado (estimado) da ℓ-ésima observação é dado por

Valores ajustados

E assim, o valor ajustado (estimado) da ℓ-ésima observação é dado por

$$\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}},$$

$\ell = 1, 2, \ldots, n.$

Valores ajustados

E assim, o valor ajustado (estimado) da ℓ-ésima observação é dado por

$$\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}},$$

$\ell = 1, 2, \ldots, n$. De forma matricial,

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\beta} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y} = \boldsymbol{H}\boldsymbol{Y}.$$

Valores ajustados

E assim, o valor ajustado (estimado) da ℓ-ésima observação é dado por

$$\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}},$$

$\ell = 1, 2, \ldots, n$. De forma matricial,

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y} = \boldsymbol{H}\boldsymbol{Y}.$$

A matriz $\boldsymbol{H}$, $n \times n$, é chamada de matriz chapéu (*hat matrix*).

Valores ajustados

E assim, o valor ajustado (estimado) da ℓ-ésima observação é dado por

$$\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}},$$

$\ell = 1, 2, \ldots, n$. De forma matricial,

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y} = \boldsymbol{H}\boldsymbol{Y}.$$

A matriz $\boldsymbol{H}$, $n \times n$, é chamada de matriz chapéu (*hat matrix*).

Matriz chapéu

Propriedades da matriz $\boldsymbol{H}$.

- $\boldsymbol{H}$ é simétrica e idempotente;

Matriz chapéu

Propriedades da matriz $\boldsymbol{H}$.

- $\boldsymbol{H}$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})$ é simétrica e idempotente;

Matriz chapéu

Propriedades da matriz $\boldsymbol{H}$.

- $\boldsymbol{H}$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{H} = \boldsymbol{0}$.

Matriz chapéu

Propriedades da matriz $\boldsymbol{H}$.

- $\boldsymbol{H}$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{H} = \mathbf{0}$.

Observação: uma matriz $\boldsymbol{M}$ é dita idempotente se e somente se $\boldsymbol{M}\boldsymbol{M} = \boldsymbol{M}$.

Matriz chapéu

Propriedades da matriz $\boldsymbol{H}$.

- $\boldsymbol{H}$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})$ é simétrica e idempotente;
- $(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{H} = \mathbf{0}$.

Observação: uma matriz $\boldsymbol{M}$ é dita idempotente se e somente se $\boldsymbol{M}\boldsymbol{M} = \boldsymbol{M}$.

Resíduos

É a diferença entre o valor observado pelo correspondente valor ajustado.

$$\begin{aligned} \boldsymbol{e} &= \boldsymbol{Y} - \widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\beta} \\ &= \boldsymbol{Y} - \boldsymbol{H}\boldsymbol{Y} = (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y}, \end{aligned}$$

Resíduos

É a diferença entre o valor observado pelo correspondente valor ajustado.

$$\boldsymbol{e} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}$$

$$= \boldsymbol{Y} - \boldsymbol{H}\boldsymbol{Y} = (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y},$$

em que $\boldsymbol{e} = (e_1, e_2, \ldots, e_n)^\top$ e $e_\ell = Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\beta}$ é o ℓ-ésimo resíduo, $\ell = 1, 2, \ldots, n$.

Resíduos

É a diferença entre o valor observado pelo correspondente valor ajustado.

$$\boldsymbol{e} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}$$

$$= \boldsymbol{Y} - \boldsymbol{H}\boldsymbol{Y} = (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y},$$

em que $\boldsymbol{e} = (e_1, e_2, \ldots, e_n)^\top$ e $e_\ell = Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\beta}$ é o ℓ-ésimo resíduo, $\ell = 1, 2, \ldots, n$.

Resíduos

A esperança de $\boldsymbol{e}$ é dada por

$$
\begin{aligned}
\mathbb{E}(\boldsymbol{e}) &= (\boldsymbol{I}_n - \boldsymbol{H})\mathbb{E}(\boldsymbol{Y}) = (\boldsymbol{I}_n - \boldsymbol{H})\mathbb{E}(\boldsymbol{X}\beta + \varepsilon) \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\left[\boldsymbol{X}\beta + \mathbb{E}(\varepsilon)\right] = (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{H}\boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{X}\beta = \mathbf{0}.
\end{aligned}
$$

Resíduos

A esperança de $\boldsymbol{e}$ é dada por

$$
\begin{aligned}
\mathbb{E}(\boldsymbol{e}) &= (\boldsymbol{I}_n - \boldsymbol{H})\mathbb{E}(\boldsymbol{Y}) = (\boldsymbol{I}_n - \boldsymbol{H})\mathbb{E}(\boldsymbol{X}\beta + \varepsilon) \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\left[\boldsymbol{X}\beta + \mathbb{E}(\varepsilon)\right] = (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{H}\boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\beta \\
&= \boldsymbol{X}\beta - \boldsymbol{X}\beta = \mathbf{0}.
\end{aligned}
$$

Resíduos

E a variância de $\boldsymbol{e}$ é dada por

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{e}) &= (\boldsymbol{I}_n - \boldsymbol{H})\mathrm{Var}(\boldsymbol{Y})(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\mathrm{Var}(\boldsymbol{X}\beta + \varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\mathrm{Var}(\varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\sigma^2\boldsymbol{I}_n(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= \sigma^2(\boldsymbol{I}_n - \boldsymbol{H}).
\end{aligned}
$$

Resíduos

E a variância de $\boldsymbol{e}$ é dada por

$$\begin{aligned}
\text{Var}(\boldsymbol{e}) &= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\boldsymbol{Y})(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\boldsymbol{X}\beta + \varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\sigma^2\boldsymbol{I}_n(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= \sigma^2(\boldsymbol{I}_n - \boldsymbol{H}).
\end{aligned}$$

Lembrando que $(\boldsymbol{I}_n - \boldsymbol{H})$ é idempotente e simétrica.

Resíduos

E a variância de $\boldsymbol{e}$ é dada por

$$\begin{aligned}
\text{Var}(\boldsymbol{e}) &= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\boldsymbol{Y})(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\boldsymbol{X}\beta + \varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\text{Var}(\varepsilon)(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= (\boldsymbol{I}_n - \boldsymbol{H})\sigma^2\boldsymbol{I}_n(\boldsymbol{I}_n - \boldsymbol{H})^\top \\
&= \sigma^2(\boldsymbol{I}_n - \boldsymbol{H}).
\end{aligned}$$

Lembrando que $(\boldsymbol{I}_n - \boldsymbol{H})$ é idempotente e simétrica.

Resíduos

Seja h_{ij}, o (i, j)-ésimo termo da matriz $\boldsymbol{H}$, $i, j = 1, 2, \ldots, n$, então nós temos que:

- $\mathbb{E}(e_i) = 0;$

Resíduos

Seja h_{ij}, o (i, j)-ésimo termo da matriz $\boldsymbol{H}$, $i, j = 1, 2, \ldots, n$, então nós temos que:

- $\mathbb{E}(e_i) = 0$;
- $\text{Var}(e_i) = \sigma^2(1 - h_{ii})$;

Resíduos

Seja h_{ij}, o (i,j)-ésimo termo da matriz $\boldsymbol{H}$, $i,j = 1, 2, \ldots, n$, então nós temos que:

- $\mathbb{E}(e_i) = 0$;
- $\text{Var}(e_i) = \sigma^2(1 - h_{ii})$;
- $\text{Cov}(e_i, e_j) = -\sigma^2 h_{ij}$.

Resíduos

Seja h_{ij}, o (i, j)-ésimo termo da matriz $\boldsymbol{H}$, $i, j = 1, 2, \ldots, n$, então nós temos que:

- $\mathbb{E}(e_i) = 0$;
- $\mathrm{Var}(e_i) = \sigma^2(1 - h_{ii})$;
- $\mathrm{Cov}(e_i, e_j) = -\sigma^2 h_{ij}$.

Estimação da variância do erro

O estimador da variância, σ^2, é dado por

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\beta}\right)^2}{n-p} = \frac{1}{n-p}\boldsymbol{e}^\top \boldsymbol{e} \tag{7}$$

$$= \frac{1}{n-p}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}\right)^\top \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}\right).$$

Estimação da variância do erro

O estimador da variância, σ^2, é dado por

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2}{n-p} = \frac{1}{n-p}\boldsymbol{e}^\top\boldsymbol{e} \tag{7}$$
$$= \frac{1}{n-p}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)^\top\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right).$$

O estimador $\hat{\sigma}^2$, dado em (7), também é denominado de **quadrado médio do resíduo** (QMRes).

Estimação da variância do erro

O estimador da variância, σ^2, é dado por

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2}{n-p} = \frac{1}{n-p}\boldsymbol{e}^\top\boldsymbol{e} \tag{7}$$

$$= \frac{1}{n-p}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)^\top\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right).$$

O estimador $\hat{\sigma}^2$, dado em (7), também é denominado de **quadrado médio do resíduo** (QMRes). E $\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2$ é denominado de **soma de quadrados do resíduo** (SQRes).

Estimação da variância do erro

O estimador da variância, σ^2, é dado por

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2}{n-p} = \frac{1}{n-p}\boldsymbol{e}^\top\boldsymbol{e} \tag{7}$$
$$= \frac{1}{n-p}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)^\top\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right).$$

O estimador $\hat{\sigma}^2$, dado em (7), também é denominado de **quadrado médio do resíduo** (QMRes). E $\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2$ é denominado de **soma de quadrados do resíduo** (SQRes).

Estimação da variância do erro

A esperança de $\hat{\sigma}^2$ é dada por

$$
\begin{aligned}
\mathbb{E}(\hat{\sigma}^2) &= \frac{1}{n-p}\mathbb{E}(\boldsymbol{e}^\top \boldsymbol{e}) = \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y}]^\top (\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y}\right\} \\
&= \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{X}\beta + \varepsilon)]^\top (\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{X}\beta + \varepsilon)\right\} \\
&\overset{*}{=} \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon]^\top (\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right\} \\
&= \frac{1}{n-p}\mathbb{E}\left[\varepsilon^\top (\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right] = \frac{1}{n-p}\mathbb{E}\left[\varepsilon^\top (\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right] \\
&\overset{**}{=} \frac{1}{n-p}\mathrm{tr}\left[(\boldsymbol{I}_n - \boldsymbol{H})\sigma^2 \boldsymbol{I}_n\right] = \frac{\sigma^2}{n-p}\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{H}) \\
&\overset{***}{=} \frac{\sigma^2}{n-p}(n-p) = \sigma^2.
\end{aligned}
$$

Estimação da variância do erro

A esperança de $\hat{\sigma}^2$ é dada por

$$
\begin{aligned}
\mathbb{E}(\hat{\sigma}^2) &= \frac{1}{n-p}\mathbb{E}(\boldsymbol{e}^\top\boldsymbol{e}) = \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y}]^\top(\boldsymbol{I}_n - \boldsymbol{H})\boldsymbol{Y}\right\} \\
&= \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{X}\beta + \varepsilon)]^\top(\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{X}\beta + \varepsilon)\right\} \\
&\overset{*}{=} \frac{1}{n-p}\mathbb{E}\left\{[(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon]^\top(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right\} \\
&= \frac{1}{n-p}\mathbb{E}\left[\varepsilon^\top(\boldsymbol{I}_n - \boldsymbol{H})(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right] = \frac{1}{n-p}\mathbb{E}\left[\varepsilon^\top(\boldsymbol{I}_n - \boldsymbol{H})\varepsilon\right] \\
&\overset{**}{=} \frac{1}{n-p}\mathrm{tr}\left[(\boldsymbol{I}_n - \boldsymbol{H})\sigma^2\boldsymbol{I}_n\right] = \frac{\sigma^2}{n-p}\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{H}) \\
&\overset{***}{=} \frac{\sigma^2}{n-p}(n-p) = \sigma^2.
\end{aligned}
$$

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

$$\mathbb{E}\left[\boldsymbol{v}^\top \boldsymbol{W} \boldsymbol{v}\right] = \mathrm{tr}\left[\boldsymbol{W}\boldsymbol{C}\right] + \boldsymbol{m}^\top \boldsymbol{W} \boldsymbol{m}.$$

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

$$\mathbb{E}\left[\boldsymbol{v}^\top \boldsymbol{W}\boldsymbol{v}\right] = \operatorname{tr}\left[\boldsymbol{W}\boldsymbol{C}\right] + \boldsymbol{m}^\top \boldsymbol{W}\boldsymbol{m}.$$

$***$: $\operatorname{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = \operatorname{tr}(\boldsymbol{I}_n) - \operatorname{tr}(\boldsymbol{H})$.

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

$$\mathbb{E}\left[\boldsymbol{v}^\top \boldsymbol{W}\boldsymbol{v}\right] = \text{tr}\left[\boldsymbol{W}\boldsymbol{C}\right] + \boldsymbol{m}^\top \boldsymbol{W}\boldsymbol{m}.$$

$***$: $\text{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = \text{tr}(\boldsymbol{I}_n) - \text{tr}(\boldsymbol{H})$. Como $\text{tr}(\boldsymbol{I}_n) = n$ e

$$\text{tr}(\boldsymbol{H}) = \text{tr}\left[\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right] = \text{tr}\left[(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\right] = \text{tr}(\boldsymbol{I}_p) = p.$$

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

$$\mathbb{E}\left[\boldsymbol{v}^\top \boldsymbol{W}\boldsymbol{v}\right] = \mathrm{tr}\left[\boldsymbol{W}\boldsymbol{C}\right] + \boldsymbol{m}^\top \boldsymbol{W}\boldsymbol{m}.$$

$***$: $\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = \mathrm{tr}(\boldsymbol{I}_n) - \mathrm{tr}(\boldsymbol{H})$. Como $\mathrm{tr}(\boldsymbol{I}_n) = n$ e

$$\mathrm{tr}(\boldsymbol{H}) = \mathrm{tr}\left[\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top\right] = \mathrm{tr}\left[(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\right] = \mathrm{tr}(\boldsymbol{I}_p) = p.$$

Logo, $\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = n - p$.

Estimação da variância do erro

$*$: Ver $\mathbb{E}(\boldsymbol{e})$.

$**$: Sejam $\boldsymbol{v}$ um vetor aleatório de comprimento v, com vetor de médias $\boldsymbol{m}$ e matriz de covariâncias $\boldsymbol{C}$ e $\boldsymbol{W}$ uma matriz de dimensão $v \times v$, então

$$\mathbb{E}\left[\boldsymbol{v}^\top \boldsymbol{W}\boldsymbol{v}\right] = \operatorname{tr}\left[\boldsymbol{W}\boldsymbol{C}\right] + \boldsymbol{m}^\top \boldsymbol{W}\boldsymbol{m}.$$

$***$: $\operatorname{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = \operatorname{tr}(\boldsymbol{I}_n) - \operatorname{tr}(\boldsymbol{H}).$ Como $\operatorname{tr}(\boldsymbol{I}_n) = n$ e

$$\operatorname{tr}(\boldsymbol{H}) = \operatorname{tr}\left[\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right] = \operatorname{tr}\left[(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\right] = \operatorname{tr}(\boldsymbol{I}_p) = p.$$

Logo, $\operatorname{tr}(\boldsymbol{I}_n - \boldsymbol{H}) = n - p$.

Estimação da variância do erro

E assim, nós temos que o QMRes, dado em (7), é um estimador não viciado para σ^2.

Roteiro

Método da máxima verossimilhança

Definição

Os estimadores de máxima verossimilhança (EMV) obtém estimativas para parâmetros desconhecidos que **maximizam** a função de verossimilhança.

Método da máxima verossimilhança

Definição

Os estimadores de máxima verossimilhança (EMV) obtém estimativas para parâmetros desconhecidos que **maximizam** a função de verossimilhança.

Método da máxima verossimilhança

Sejam $Y_1, Y_2 \ldots, Y_n$ variáveis aleatórias independentes, com $Y_\ell \sim \mathcal{N}(\mu_\ell, \sigma^2)$,

Método da máxima verossimilhança

Sejam $Y_1, Y_2 \ldots, Y_n$ variáveis aleatórias independentes, com $Y_\ell \sim \mathcal{N}(\mu_\ell, \sigma^2)$,

em que $\mu_\ell = \boldsymbol{x}_\ell^\top \beta$, $\ell = 1, 2, \ldots, n$,

Método da máxima verossimilhança

Sejam $Y_1, Y_2 \ldots, Y_n$ variáveis aleatórias independentes, com $Y_\ell \sim \mathcal{N}(\mu_\ell, \sigma^2)$, em que $\mu_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta}$, $\ell = 1, 2, \ldots, n$, a função de verossimilhança é dada por

$$L(\beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \beta \right)^2 \right\}. \qquad (8)$$

Método da máxima verossimilhança

Sejam $Y_1, Y_2 \ldots, Y_n$ variáveis aleatórias independentes, com $Y_\ell \sim \mathcal{N}(\mu_\ell, \sigma^2)$, em que $\mu_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta}$, $\ell = 1, 2, \ldots, n$, a função de verossimilhança é dada por

$$L(\boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \boldsymbol{\beta}\right)^2\right\}. \tag{8}$$

Método da máxima verossimilhança

Maximizar (8) é equivalente a maximizar o seu logaritmo, dado por

$$l(\beta, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \beta\right)^2. \tag{9}$$

Método da máxima verossimilhança

Maximizar (8) é equivalente a maximizar o seu logaritmo, dado por

$$l(\boldsymbol{\beta}, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \boldsymbol{\beta}\right)^2. \qquad (9)$$

Método da máxima verossimilhança

Derivando (9), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial l(\beta, \sigma^2)}{\partial \beta_m} = \frac{1}{\sigma^2} \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \qquad (10)$$

Método da máxima verossimilhança

Derivando (9), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \beta_m} = \frac{1}{\sigma^2} \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \tag{10}$$

Igualando (10) a zero,

Método da máxima verossimilhança

Derivando (9), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \beta_m} = \frac{1}{\sigma^2} \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \qquad (10)$$

Igualando (10) a zero, nós temos que

$$\hat{\beta}_1 \sum_{\ell=1}^{n} x_{\ell 1} x_{\ell m} + \cdots + \hat{\beta}_m \sum_{\ell=1}^{n} x_{\ell m}^2 + \cdots + \hat{\beta}_p \sum_{\ell=1}^{n} x_{\ell p} x_{\ell m} = \sum_{\ell=1}^{n} x_{\ell m} Y_\ell. \quad (11)$$

Método da máxima verossimilhança

Derivando (9), em relação ao parâmetro β_m, $m = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \beta_m} = \frac{1}{\sigma^2} \sum_{\ell=1}^{n} \left[Y_\ell - (\beta_1 x_{\ell 1} + \beta_2 x_{\ell 2} + \cdots + \beta_p x_{\ell p}) \right] x_{\ell m}. \qquad (10)$$

Igualando (10) a zero, nós temos que

$$\hat{\beta}_1 \sum_{\ell=1}^{n} x_{\ell 1} x_{\ell m} + \cdots + \hat{\beta}_m \sum_{\ell=1}^{n} x_{\ell m}^2 + \cdots + \hat{\beta}_p \sum_{\ell=1}^{n} x_{\ell p} x_{\ell m} = \sum_{\ell=1}^{n} x_{\ell m} Y_\ell. \qquad (11)$$

Método da máxima verossimilhança

O **estimador de máxima verossimilhança** de β_m, obtido da solução das p equações em (11), é dado por:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{12}$$

Método da máxima verossimilhança

O **estimador de máxima verossimilhança** de β_m, obtido da solução das p equações em (11), é dado por:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{12}$$

O EMV em (12) coincide com o EMQ, apresentado em (6).

Método da máxima verossimilhança

O **estimador de máxima verossimilhança** de β_m, obtido da solução das p equações em (11), é dado por:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{12}$$

O EMV em (12) coincide com o EMQ, apresentado em (6).

Método da máxima verossimilhança

Agora, derivando (9), em relação ao parâmetro σ^2, nós temos:

$$\frac{\partial l(\beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \beta \right)^2 . \tag{13}$$

Método da máxima verossimilhança

Agora, derivando (9), em relação ao parâmetro σ^2, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \boldsymbol{\beta} \right)^2 . \tag{13}$$

Igualando (13) a zero,

Método da máxima verossimilhança

Agora, derivando (9), em relação ao parâmetro σ^2, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \boldsymbol{\beta} \right)^2. \tag{13}$$

Igualando (13) a zero, nós temos que

$$\frac{1}{2(\hat{\sigma}^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}} \right)^2 = \frac{n}{2\hat{\sigma}^2}. \tag{14}$$

Agora, derivando (9), em relação ao parâmetro σ^2, nós temos:

$$\frac{\partial l(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \boldsymbol{\beta} \right)^2 . \tag{13}$$

Igualando (13) a zero, nós temos que

$$\frac{1}{2(\hat{\sigma}^2)^2} \sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}} \right)^2 = \frac{n}{2\hat{\sigma}^2}. \tag{14}$$

Método da máxima verossimilhança

O **estimador de máxima verossimilhança** de σ^2, obtido da solução da equação (14), é dado por:

$$\hat{\sigma}^2_{\mathrm{MV}} = \frac{\sum_{\ell=1}^{n} \left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}} \right)^2}{n} = \frac{n-p}{n} \mathrm{QMRes}. \tag{15}$$

Método da máxima verossimilhança

O **estimador de máxima verossimilhança** de σ^2, obtido da solução da equação (14), é dado por:

$$\hat{\sigma}^2_{\mathrm{MV}} = \frac{\sum_{\ell=1}^{n}\left(Y_\ell - \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}\right)^2}{n} = \frac{n-p}{n}\mathrm{QMRes}. \tag{15}$$

Método da máxima verossimilhança

O estimador para β obtido pelo método da máxima verossimilhança coincide com o obtido pelo método dos mínimos quadrados, por conta disso, o EMV de β é **não viesado** e tem menor variância entre todos os estimadores lineares não viesados.

Método da máxima verossimilhança

O estimador para β obtido pelo método da máxima verossimilhança coincide com o obtido pelo método dos mínimos quadrados, por conta disso, o EMV de β é **não viesado** e tem menor variância entre todos os estimadores lineares não viesados.

Método da máxima verossimilhança

O mesmo não acontece com o EMV de σ^2. De fato, $\hat{\sigma}^2_{\mathrm{MV}}$ é um **estimador viciado**,

Método da máxima verossimilhança

O mesmo não acontece com o EMV de σ^2. De fato, $\hat{\sigma}^2_{\mathrm{MV}}$ é um **estimador viciado**, para tamanhos de amostras pequenos, esse viés é considerável.

Método da máxima verossimilhança

O mesmo não acontece com o EMV de σ^2. De fato, $\hat{\sigma}^2_{\mathrm{MV}}$ é um **estimador viciado**, para tamanhos de amostras pequenos, esse viés é considerável.

Adicionalmente, os EMV de β e σ^2 são estimadores consistentes e suficientes.

Método da máxima verossimilhança

O mesmo não acontece com o EMV de σ^2. De fato, $\hat{\sigma}^2_{\mathrm{MV}}$ é um **estimador viciado**, para tamanhos de amostras pequenos, esse viés é considerável.

Adicionalmente, os EMV de $\boldsymbol{\beta}$ e σ^2 são estimadores consistentes e suficientes.

A suposição de normalidade e a consequente estimação por máxima verossimilhança permitem a construção de procedimentos inferenciais, como intervalos de confiança.

Método da máxima verossimilhança

O mesmo não acontece com o EMV de σ^2. De fato, $\hat{\sigma}^2_{\mathrm{MV}}$ é um **estimador viciado**, para tamanhos de amostras pequenos, esse viés é considerável.

Adicionalmente, os EMV de $\boldsymbol{\beta}$ e σ^2 são estimadores consistentes e suficientes.

A suposição de normalidade e a consequente estimação por máxima verossimilhança permitem a construção de procedimentos inferenciais, como intervalos de confiança.

Método da máxima verossimilhança

Derivando (9), uma segunda vez, em relação ao parâmetro $\beta_{m'}$, $m' = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial^2 l(\beta, \sigma^2)}{\partial \beta_m \partial \beta_{m'}} = -\frac{1}{\sigma^2} \sum_{\ell=1}^{n} x_{\ell m} x_{\ell m'}. \tag{16}$$

Derivando (9), uma segunda vez, em relação ao parâmetro $\beta_{m'}$, $m' = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial^2 l(\boldsymbol{\beta}, \sigma^2)}{\partial \beta_m \partial \beta_{m'}} = -\frac{1}{\sigma^2} \sum_{\ell=1}^{n} x_{\ell m} x_{\ell m'}. \tag{16}$$

Se nós aplicamos a esperança e multiplicamos por -1 em (16), obtemos (m, m')-ésimo elemento da matriz de informação de Fisher do EMV de β.

Método da máxima verossimilhança

Derivando (9), uma segunda vez, em relação ao parâmetro $\beta_{m'}$, $m' = 1, 2, \ldots, p$, nós temos:

$$\frac{\partial^2 l(\boldsymbol{\beta}, \sigma^2)}{\partial \beta_m \partial \beta_{m'}} = -\frac{1}{\sigma^2} \sum_{\ell=1}^{n} x_{\ell m} x_{\ell m'}. \tag{16}$$

Se nós aplicamos a esperança e multiplicamos por -1 em (16), obtemos (m, m')-ésimo elemento da matriz de informação de Fisher do EMV de $\boldsymbol{\beta}$.

Lembrando que, a equação (10) é a função escore. Em notação matricial, o vetor escore e a matriz de informação de Fisher são dadas, respectivamente, por

Lembrando que, a equação (10) é a função escore. Em notação matricial, o vetor escore e a matriz de informação de Fisher são dadas, respectivamente, por

$$\boldsymbol{U}_\beta = \frac{1}{\sigma^2}\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{X}\beta) \text{ e } \boldsymbol{K}_\beta = \frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{X}. \tag{17}$$

Método da máxima verossimilhança

Lembrando que, a equação (10) é a função escore. Em notação matricial, o vetor escore e a matriz de informação de Fisher são dadas, respectivamente, por

$$\boldsymbol{U}_\beta = \frac{1}{\sigma^2}\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{X}\beta) \text{ e } \boldsymbol{K}_\beta = \frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{X}. \tag{17}$$

Considerações

Embora seja comum utilizar $\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}$,

Considerações

Embora seja comum utilizar $\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}$, o que, de fato, nós estamos estimando é μ_ℓ.

Considerações

Embora seja comum utilizar $\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}$, o que, de fato, nós estamos estimando é μ_ℓ. Basta lembrar que, $\mu_\ell = \mathbb{E}(Y_\ell) = \boldsymbol{x}_\ell^\top \boldsymbol{\beta}$.

Considerações

Embora seja comum utilizar $\hat{Y}_\ell = \boldsymbol{x}_\ell^\top \widehat{\boldsymbol{\beta}}$, o que, de fato, nós estamos estimando é μ_ℓ. Basta lembrar que, $\mu_\ell = \mathbb{E}(Y_\ell) = \boldsymbol{x}_\ell^\top \boldsymbol{\beta}$.

Roteiro

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

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

Aplicações

Aplicação 1. Conjunto de dados *Westwood Company* (Neter et al., 1983, p. 36). Após o ajuste (estimação dos parâmetros), 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, $\ell = 1, 2, \ldots, 10$ e QMRes = 7,5.

Aplicações

Aplicação 1. Conjunto de dados *Westwood Company* (Neter et al., 1983, p. 36). Após o ajuste (estimação dos parâmetros), 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, $\ell = 1, 2, \ldots, 10$ e QMRes $= 7{,}5$.

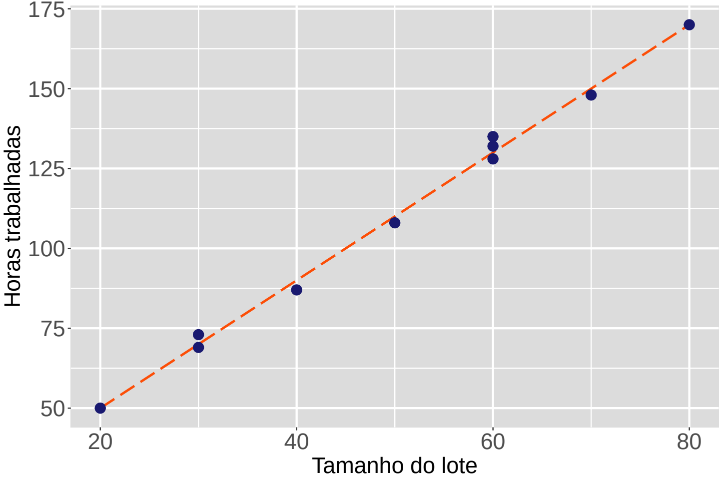

Figura 1: GD com a reta ajustada para os dados *Westwood Company*.

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = 10$:

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = 10$: se o tamanho do lote fosse zero, haveria 10 horas trabalhadas;

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = 10$: se o tamanho do lote fosse zero, haveria 10 horas trabalhadas;
- $\hat{\beta}_2 = 2$:

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = 10$: se o tamanho do lote fosse zero, haveria 10 horas trabalhadas;
- $\hat{\beta}_2 = 2$: a cada uma unidade aumentada no lote, ocorreria um aumento médio de 2 horas trabalhadas.

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = 10$: se o tamanho do lote fosse zero, haveria 10 horas trabalhadas;
- $\hat{\beta}_2 = 2$: a cada uma unidade aumentada no lote, ocorreria um aumento médio de 2 horas trabalhadas.

Aplicações

Aplicação 2. Conjunto de dados Walker (1962). Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = -7{,}211 + 3{,}603x_{\ell 2} - 10{,}065x_{\ell 3},$$

Aplicação 2. Conjunto de dados Walker (1962). Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = -7{,}211 + 3{,}603 x_{\ell 2} - 10{,}065 x_{\ell 3},$$

em que Y_ℓ: número de pulsações, $x_{\ell 2}$: temperatura, $x_{\ell 3}$ a espécie do grilo, $x_{\ell 3} \in \{\textit{Oecanthus exclamationis (ex)}, \textit{Oecanthus niveus (niv)}\}$, $\ell = 1, 2, \ldots, 31$ e QMRes $= 12{,}764$.

Aplicação 2. Conjunto de dados Walker (1962). Após o ajuste, nós temos o seguinte modelo estimado,

$$\hat{Y}_\ell = -7{,}211 + 3{,}603x_{\ell 2} - 10{,}065x_{\ell 3},$$

em que Y_ℓ: número de pulsações, $x_{\ell 2}$: temperatura, $x_{\ell 3}$ a espécie do grilo, $x_{\ell 3} \in \{\textit{Oecanthus exclamationis (ex)}, \textit{Oecanthus niveus (niv)}\}$, $\ell = 1, 2, \ldots, 31$ e QMRes $= 12{,}764$.

Aplicações

Uma forma alternativa de apresentar o modelo ajustado é a seguinte:

$$\begin{cases} \textit{Oecanthus niveus}: & \hat{Y}_\ell = -17{,}276 + 3{,}603 x_{\ell 2} \\ \textit{Oecanthus exclamationis}: & \hat{Y}_\ell = -7{,}211 + 3{,}603 x_{\ell 2} \end{cases}.$$

Uma forma alternativa de apresentar o modelo ajustado é a seguinte:

$$\begin{cases} \textit{Oecanthus niveus}: & \hat{Y}_\ell = -17{,}276 + 3{,}603 x_{\ell 2} \\ \textit{Oecanthus exclamationis}: & \hat{Y}_\ell = -7{,}211 + 3{,}603 x_{\ell 2} \end{cases}.$$

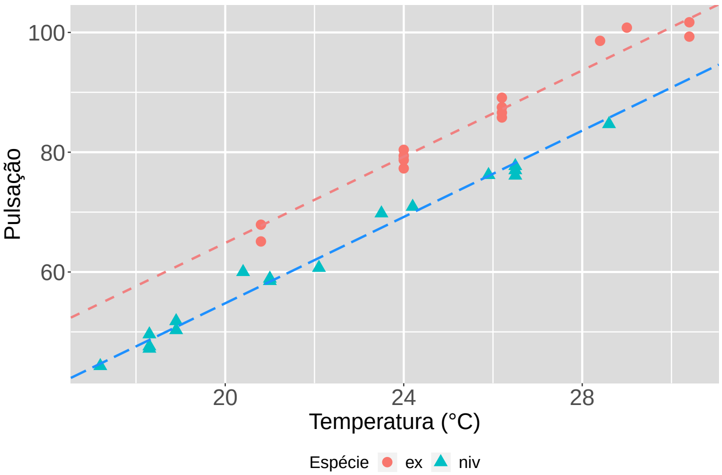

Figura 2: GD com as retas ajustadas para os dados de Walker (1962).

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$:

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);
- $\hat{\beta}_2 = 3{,}603$:

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);
- $\hat{\beta}_2 = 3{,}603$: a cada uma unidade aumentada na temperatura, haveria um aumento médio de 3,603 nas pulsações;

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);
- $\hat{\beta}_2 = 3{,}603$: a cada uma unidade aumentada na temperatura, haveria um aumento médio de 3,603 nas pulsações;
- $\hat{\beta}_3 = -10{,}065$:

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);
- $\hat{\beta}_2 = 3{,}603$: a cada uma unidade aumentada na temperatura, haveria um aumento médio de 3,603 nas pulsações;
- $\hat{\beta}_3 = -10{,}065$: a espécie *Oecanthus niveus* tem, em média, - 10,065 pulsações a menos que a *Oecanthus exclamationis*.

Aplicações

Os parâmetros podem ser interpretados da seguinte forma

- $\hat{\beta}_1 = -7{,}211$: se a temperatura fosse zero e a espécie do grilo fosse a *Oecanthus exclamationis*, o número médio de pulsações seria de -7,211 (faz sentido?);
- $\hat{\beta}_2 = 3{,}603$: a cada uma unidade aumentada na temperatura, haveria um aumento médio de 3,603 nas pulsações;
- $\hat{\beta}_3 = -10{,}065$: a espécie *Oecanthus niveus* tem, em média, - 10,065 pulsações a menos que a *Oecanthus exclamationis*.

Roteiro

Estudo de simulação

Suponham o seguinte MNL,

$$Y_\ell = 1 + x_{\ell 2} + \varepsilon_\ell,$$

Estudo de simulação

Suponham o seguinte MNL,

$$Y_\ell = 1 + x_{\ell 2} + \varepsilon_\ell,$$

em que $x_{\ell 2}$ é quantidade conhecida e $\varepsilon_\ell \sim \mathcal{N}(0, 4)$, $\ell = 1, 2, \ldots, n$.

Estudo de simulação

Suponham o seguinte MNL,

$$Y_\ell = 1 + x_{\ell 2} + \varepsilon_\ell,$$

em que $x_{\ell 2}$ é quantidade conhecida e $\varepsilon_\ell \sim \mathcal{N}(0, 4)$, $\ell = 1, 2, \ldots, n$. E nós vamos assumir que $x_{\ell 2} \sim$ uniforme no intervalo (0,1).

Estudo de simulação

Suponham o seguinte MNL,

$$Y_\ell = 1 + x_{\ell 2} + \varepsilon_\ell,$$

em que $x_{\ell 2}$ é quantidade conhecida e $\varepsilon_\ell \sim \mathcal{N}(0, 4)$, $\ell = 1, 2, \ldots, n$. E nós vamos assumir que $x_{\ell 2} \sim$ uniforme no intervalo (0,1).

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2,

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2, neste último, utilizando o QMRes e o EMV,

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2, neste último, utilizando o QMRes e o EMV, para $n \in \{10, 20, 40, 80, 160\}$.

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2, neste último, utilizando o QMRes e o EMV, para $n \in \{10, 20, 40, 80, 160\}$. Sem perda generalidade, o viés absoluto para o estimador de θ é dada por

$$\text{Viés absoluto} = \left|\hat{\theta}_{\mathrm{MC}} - \theta\right|,$$

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2, neste último, utilizando o QMRes e o EMV, para $n \in \{10, 20, 40, 80, 160\}$. Sem perda generalidade, o viés absoluto para o estimador de θ é dada por

$$\text{Viés absoluto} = \left|\hat{\theta}_{\mathrm{MC}} - \theta\right|,$$

em que $\hat{\theta}_{\mathrm{MC}} = 5.000^{-1} \sum_{i=1}^{5.000} \hat{\theta}_i$.

Estudo de simulação

Nós faremos um estudo de Monte Carlo, com 5.000 réplicas, onde avaliaremos o viés absoluto das estimativas dos parâmetros β_1, β_2 e σ^2, neste último, utilizando o QMRes e o EMV, para $n \in \{10, 20, 40, 80, 160\}$. Sem perda generalidade, o viés absoluto para o estimador de θ é dada por

$$\text{Viés absoluto} = \left|\hat{\theta}_{\mathrm{MC}} - \theta\right|,$$

em que $\hat{\theta}_{\mathrm{MC}} = 5.000^{-1} \sum_{i=1}^{5.000} \hat{\theta}_i$.

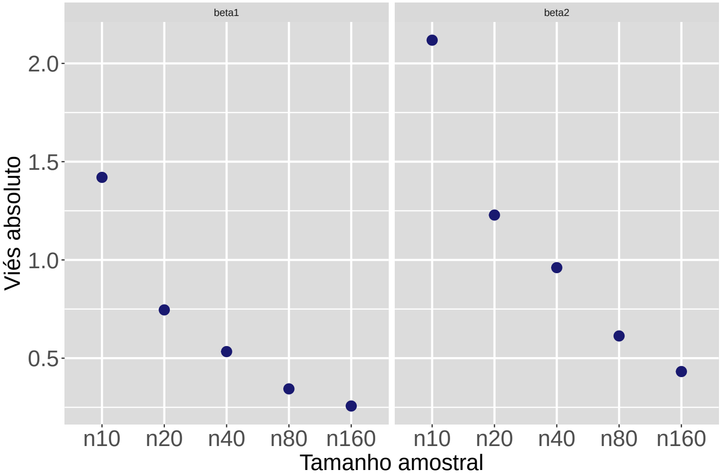

Figura 3: Viés absoluto para β.

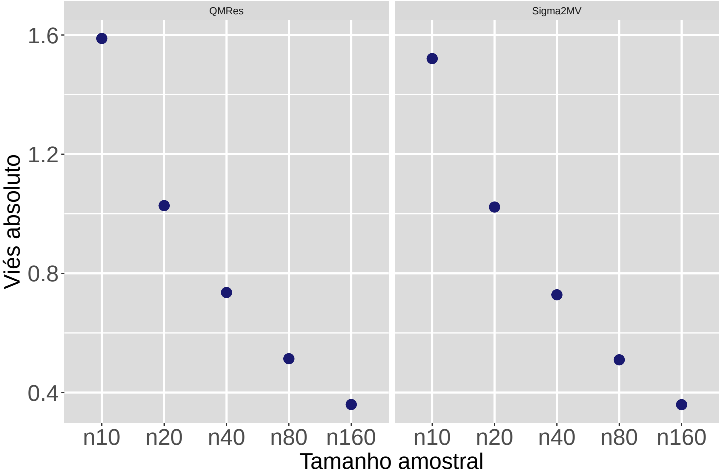

Figura 4: Viés absoluto para σ^2, utilizando o QMRes e o EMV.

Estudo de simulação

Comentários

Pelas Figuras 3 e 4, nós podemos perceber que, a medida que o tamanho amostral cresce, o viés absoluto converge para zero.

Estudo de simulação

Comentários

Pelas Figuras 3 e 4, nós podemos perceber que, a medida que o tamanho amostral cresce, o viés absoluto converge para zero.

Roteiro

Referências bibliográficas I

Neter, J., Wasserman, W. e Kutner, M. H. (1983), *Applied linear regression models*, Richard D. Irwin Inc, Homewood, Illinois.

Plackett, R. L. (1949), 'A historical note on the method of least squares', *Biometrika* **36**(3/4), 458–460.

Walker, T. J. (1962), 'The Taxonomy and Calling Songs of United States Tree Crickets (Orthoptera: Gryllidae: Oecanthinae). I. The Genus Neoxabea and the niveus and varicornis Groups of the Genus Oecanthus', *Annals of the Entomological Society of America* **55**(3), 303–322.

Obrigado!

✉ tiago.magalhaes@ufjf.br

🖱 ufjf.br/tiago_magalhaes

🌍 Departamento de Estatística, Sala 319