Intervalos de confiança

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 17 de abril de 2024

Roteiro

Roteiro

Modelo de regressão linear

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

$$Y_\ell = \boldsymbol{x}_\ell^\top \beta + \varepsilon_\ell,\ \ell = 1, 2, \ldots, n, \tag{1}$$

Modelo de regressão linear

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

$$Y_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell,\ \ell = 1, 2, \ldots, n, \tag{1}$$

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, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 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

$$Y_\ell = \boldsymbol{x}_\ell^\top \boldsymbol{\beta} + \varepsilon_\ell,\ \ell = 1, 2, \ldots, n, \tag{1}$$

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, $\varepsilon_\ell \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ são variáveis aleatórias independentes e com a mesma variância σ^2, também desconhecida, a ser estimada.

Forma matricial

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

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

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:

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

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

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

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

Método de máxima verossimilhança

O **estimador de máxima verossimilhança** (EMV) de $\boldsymbol{\beta}$ e σ^2 são dados, respectivamente, por:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}, \tag{3}$$

$$\hat{\sigma}^2_{\mathrm{MV}} = \frac{1}{n} \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta} \right)^\top \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta} \right).$$

Método de máxima verossimilhança

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

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y}, \tag{3}$$

$$\hat{\sigma}^2_{\mathrm{MV}} = \frac{1}{n}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}\right)^\top \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta}\right).$$

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:

$$\widehat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}, \tag{3}$$

$$\hat{\sigma}^2_{\mathrm{MV}} = \frac{1}{n} \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta} \right)^\top \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\beta} \right).$$

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

Sob as condições de regularidades, nós temos que:

$$\widehat{\beta} \sim \mathcal{N}_p \left\{ \beta, \sigma^2 \left(\boldsymbol{X}^\top \boldsymbol{X} \right)^{-1} \right\}, \tag{4}$$

$$\hat{\sigma}^2_{\mathrm{MV}} \sim \mathcal{N} \left\{ \sigma^2, \frac{2(\sigma^2)^2}{n} \right\},$$

Método de máxima verossimilhança

Sob as condições de regularidades, nós temos que:

$$\widehat{\boldsymbol{\beta}} \sim \mathcal{N}_{\boldsymbol{p}}\left\{\boldsymbol{\beta}, \sigma^2\left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1}\right\}, \tag{4}$$

$$\hat{\sigma}^2_{\mathrm{MV}} \sim \mathcal{N}\left\{\sigma^2, \frac{2(\sigma^2)^2}{n}\right\},$$

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

Método de máxima verossimilhança

Sob as condições de regularidades, nós temos que:

$$\widehat{\boldsymbol{\beta}} \sim \mathcal{N}_p \left\{ \boldsymbol{\beta}, \sigma^2 \left(\boldsymbol{X}^\top \boldsymbol{X} \right)^{-1} \right\}, \tag{4}$$

$$\hat{\sigma}^2_{\mathrm{MV}} \sim \mathcal{N} \left\{ \sigma^2, \frac{2(\sigma^2)^2}{n} \right\},$$

quando o tamanho de amostra é grande, adicionalmente, $\widehat{\boldsymbol{\beta}}$ e $\hat{\sigma}^2_{\mathrm{MV}}$ 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 Intervalo de confiança para β_m

3 Intervalos de confiança para μ_h e σ^2

4 Intervalos de confiança a partir de estatísticas do teste

5 Aplicação

6 Referências bibliográficas

Intervalo de confiança

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

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

para $m = 1, 2, \ldots, p$.

Intervalo de confiança

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

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

para $m = 1, 2, \ldots, p$.

Intervalo de confiança

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

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

Intervalo de confiança

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

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

em que $t(n-p)$ é a distribuição t de Student (Gosset "Student", 1908), com $n-p$ graus de liberdade, $m = 1, 2, \ldots, p$.

Intervalo de confiança

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

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

em que $t(n-p)$ é a distribuição t de Student (Gosset "Student", 1908), com $n-p$ graus de liberdade, $m = 1, 2, \ldots, p$.

Intervalo de confiança para β_m

Por (6), o intervalo de confiança para β_m, com coeficiente de confiança $(1-\alpha)$ é dado por

$$\text{IC}(\beta_m; 1-\alpha) = \hat{\beta}_m \mp t(1-\alpha/2; n-p)\sqrt{\text{QMRes}\, c_{mm}},$$

Intervalo de confiança para β_m

Por (6), o intervalo de confiança para β_m, com coeficiente de confiança $(1-\alpha)$ é dado por

$$\mathrm{IC}(\beta_m; 1-\alpha) = \hat{\beta}_m \mp t(1-\alpha/2; n-p)\sqrt{\mathrm{QMRes}\, c_{mm}},$$

em que $t(1-\alpha/2; n-p)$ é o quantil $100(1-\alpha/2)\%$ da distribuição t de Student, com $n-p$ graus de liberdade, $m = 1, 2, \ldots, p$.

Intervalo de confiança para β_m

Por (6), o intervalo de confiança para β_m, com coeficiente de confiança $(1-\alpha)$ é dado por

$$\text{IC}(\beta_m; 1-\alpha) = \hat{\beta}_m \mp t(1-\alpha/2; n-p)\sqrt{\text{QMRes}\, c_{mm}},$$

em que $t(1-\alpha/2; n-p)$ é o quantil $100(1-\alpha/2)\%$ da distribuição t de Student, com $n-p$ graus de liberdade, $m = 1, 2, \ldots, p$.

Roteiro

Inferência sobre μ_h

Nós podemos também nos interessar pela a estimação de μ_h, a média de Y_h, para um vetor de observações $\boldsymbol{x}_h = (x_{h1}, x_{h2}, \ldots, x_{hp})^\top$, que, pelo princípio da invariância, tem seu estimador dado por

$$\hat{\mu}_h = \boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}. \quad (7)$$

Inferência sobre μ_h

Nós podemos também nos interessar pela a estimação de μ_h, a média de Y_h, para um vetor de observações $\boldsymbol{x}_h = (x_{h1}, x_{h2}, \ldots, x_{hp})^\top$, que, pelo princípio da invariância, tem seu estimador dado por

$$\hat{\mu}_h = \boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}. \tag{7}$$

Inferência sobre μ_h

Por (7), nós temos que

$$\mathbb{E}\left(\hat{\mu}_h\right) = \mathbb{E}\left(\boldsymbol{x}_h^\top \widehat{\beta}\right) = \boldsymbol{x}_h^\top \mathbb{E}\left(\widehat{\beta}\right) = \boldsymbol{x}_h^\top \beta = \mu_h$$

Inferência sobre μ_h

Por (7), nós temos que

$$\mathbb{E}\left(\hat{\mu}_h\right) = \mathbb{E}\left(\boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \mathbb{E}\left(\widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \boldsymbol{\beta} = \mu_h$$

e

$$\text{Var}\left(\hat{\mu}_h\right) = \text{Var}\left(\boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \text{Var}\left(\widehat{\boldsymbol{\beta}}\right) \boldsymbol{x}_h = \sigma^2 \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h.$$

Inferência sobre μ_h

Por (7), nós temos que

$$\mathbb{E}\left(\hat{\mu}_h\right) = \mathbb{E}\left(\boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \mathbb{E}\left(\widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \boldsymbol{\beta} = \mu_h$$

e

$$\mathrm{Var}\left(\hat{\mu}_h\right) = \mathrm{Var}\left(\boldsymbol{x}_h^\top \widehat{\boldsymbol{\beta}}\right) = \boldsymbol{x}_h^\top \mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\right) \boldsymbol{x}_h = \sigma^2 \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h.$$

Inferência sobre μ_h

Por (7) e pelo método delta (Dorfman, 1938), a distribuição de $\hat{\mu}_h$ é dada por:

$$\hat{\mu}_h \sim \mathcal{N}\left\{\mu_h,\, \sigma^2 \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h\right\}.$$

Inferência sobre μ_h

Por (7) e pelo método delta (Dorfman, 1938), a distribuição de $\hat{\mu}_h$ é dada por:

$$\hat{\mu}_h \sim \mathcal{N}\left\{\mu_h,\, \sigma^2 \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h\right\}.$$

O estimador da variância de $\hat{\mu}_h$ é dada por:

$$\hat{\text{V}}\text{ar}(\hat{\mu}_h) = \text{QMRes}\, \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h.$$

Inferência sobre μ_h

Por (7) e pelo método delta (Dorfman, 1938), a distribuição de $\hat{\mu}_h$ é dada por:

$$\hat{\mu}_h \sim \mathcal{N}\left\{\mu_h,\, \sigma^2 \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h\right\}.$$

O estimador da variância de $\hat{\mu}_h$ é dada por:

$$\hat{\text{V}}\text{ar}(\hat{\mu}_h) = \text{QMRes}\, \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h.$$

Intervalo de confiança para μ_h

O intervalo de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp t(1-\alpha/2; n-p)\sqrt{\text{QMRes}\, \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

Intervalo de confiança para μ_h

O intervalo de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp t(1-\alpha/2; n-p)\sqrt{\text{QMRes}\, \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

em que $t(1-\alpha/2; n-p)$ é o quantil $100(1-\alpha/2)\%$ da distribuição t de Student, com $n-p$ graus de liberdade.

Intervalo de confiança para μ_h

O intervalo de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp t(1-\alpha/2; n-p)\sqrt{\text{QMRes}\, \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

em que $t(1-\alpha/2; n-p)$ é o quantil $100(1-\alpha/2)\%$ da distribuição t de Student, com $n-p$ graus de liberdade.

Intervalo de confiança para μ_h

Uma alternativa ao IC anterior, o método de Working-Hotelling (Working e Hotelling, 1929) é utilizado para a construção de uma banda de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp W \sqrt{\text{QMRes}\ \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

Intervalo de confiança para μ_h

Uma alternativa ao IC anterior, o método de Working-Hotelling (Working e Hotelling, 1929) é utilizado para a construção de uma banda de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp W\sqrt{\text{QMRes}\ \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

em que $W^2 = pF(1-\alpha; p, n-p)$ e $F(1-\alpha; p, n-p)$ é o quantil $100(1-\alpha)\%$ da distribuição F, com p e $n-p$ graus de liberdade.

Intervalo de confiança para μ_h

Uma alternativa ao IC anterior, o método de Working-Hotelling (Working e Hotelling, 1929) é utilizado para a construção de uma banda de confiança para μ_h:

$$\text{IC}(\mu_h; 1-\alpha) = \hat{\mu}_h \mp W\sqrt{\text{QMRes}\ \boldsymbol{x}_h^\top \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{x}_h},$$

em que $W^2 = pF(1-\alpha; p, n-p)$ e $F(1-\alpha; p, n-p)$ é o quantil $100(1-\alpha)\%$ da distribuição F, com p e $n-p$ graus de liberdade.

Intervalo de confiança para μ_h

Observação

O método também é conhecido como Bandas de Confiança Working-Hotelling-Scheffé, pela semelhança do Método de Working-Hotelling com o método de Scheffé (1959) na análise de constraste em uma Análise de Variância.

Intervalo de confiança para σ^2

O intervalo de confiança para σ^2:

$$\text{IC}(\sigma^2; 1-\alpha) = \left[\frac{(n-p)\text{QMRes}}{\chi^2(\alpha/2; n-p)}; \frac{(n-p)\text{QMRes}}{\chi^2(1-\alpha/2; n-p)}\right],$$

Intervalo de confiança para σ^2

O intervalo de confiança para σ^2:

$$\text{IC}(\sigma^2; 1-\alpha) = \left[\frac{(n-p)\text{QMRes}}{\chi^2(\alpha/2; n-p)}; \frac{(n-p)\text{QMRes}}{\chi^2(1-\alpha/2; n-p)}\right],$$

em que $\chi^2(\alpha/2; n-p)$ e $\chi^2(1-\alpha/2; n-p)$ são, respectivamente, os quantis $100\alpha/2\%$ e $100(1-\alpha/2)\%$ da distribuição qui-quadrado, com $n-p$ graus de liberdade.

Intervalo de confiança para σ^2

O intervalo de confiança para σ^2:

$$\text{IC}(\sigma^2; 1-\alpha) = \left[\frac{(n-p)\text{QMRes}}{\chi^2(\alpha/2; n-p)}; \ \frac{(n-p)\text{QMRes}}{\chi^2(1-\alpha/2; n-p)}\right],$$

em que $\chi^2(\alpha/2; n-p)$ e $\chi^2(1-\alpha/2; n-p)$ são, respectivamente, os quantis $100\alpha/2\%$ e $100(1-\alpha/2)\%$ da distribuição qui-quadrado, com $n-p$ graus de liberdade.

Roteiro

Intervalos de confiança a partir de estatísticas do teste

Sejam S_1 a S_4, respectivamente, as estatísticas dos testes da razão de verossimilhanças (Wilks, 1938), Wald (Wald, 1943), escore (Rao, 1948) e gradiente (Terrell, 2002). Nós vimos que, quando $n \to \infty$,

$$S_i \sim \chi^2_q,\ i = 1, 2, 3, 4,$$

Intervalos de confiança a partir de estatísticas do teste

Sejam S_1 a S_4, respectivamente, as estatísticas dos testes da razão de verossimilhanças (Wilks, 1938), Wald (Wald, 1943), escore (Rao, 1948) e gradiente (Terrell, 2002). Nós vimos que, quando $n \to \infty$,

$$S_i \sim \chi^2_q,\ i = 1, 2, 3, 4,$$

em que q é o número de parâmetros na hipótese nula, a dimensão do vetor $\beta_1^{(0)}$.

Intervalos de confiança a partir de estatísticas do teste

Sejam S_1 a S_4, respectivamente, as estatísticas dos testes da razão de verossimilhanças (Wilks, 1938), Wald (Wald, 1943), escore (Rao, 1948) e gradiente (Terrell, 2002). Nós vimos que, quando $n \to \infty$,

$$S_i \sim \chi^2_q,\ i = 1, 2, 3, 4,$$

em que q é o número de parâmetros na hipótese nula, a dimensão do vetor $\boldsymbol{\beta}_1^{(0)}$.

Intervalos de confiança a partir de estatísticas do teste

Intervalos de confiança a partir das estatística S_i é construído através de um conjunto de valores tais que,

$$S_i \leq \chi^2(1-\alpha; q),$$

Intervalos de confiança a partir de estatísticas do teste

Intervalos de confiança a partir das estatística S_i é construído através de um conjunto de valores tais que,

$$S_i \leq \chi^2(1 - \alpha; q),$$

em que $\chi^2(1 - \alpha; q)$ é o quantil $100(1 - \alpha)\%$ de uma distribuição qui-quadrado com q graus de liberdade, $i = 1, 2, 3, 4$.

Intervalos de confiança a partir de estatísticas do teste

Intervalos de confiança a partir das estatística S_i é construído através de um conjunto de valores tais que,

$$S_i \leq \chi^2(1-\alpha; q),$$

em que $\chi^2(1-\alpha; q)$ é o quantil $100(1-\alpha)\%$ de uma distribuição qui-quadrado com q graus de liberdade, $i = 1, 2, 3, 4$.

Intervalos de confiança a partir de estatísticas do teste

Suponham que o nosso interesse é testar,

$$\begin{cases} \mathcal{H}: & \beta_1 = \beta^* \\ \mathcal{A}: & \beta_1 \neq \beta^* \end{cases}, \qquad (8)$$

Intervalos de confiança a partir de estatísticas do teste

Suponham que o nosso interesse é testar,

$$\begin{cases} \mathcal{H}: & \beta_1 = \beta^* \\ \mathcal{A}: & \beta_1 \neq \beta^* \end{cases}, \tag{8}$$

a estatística da razão de verossimilhança (RV) pode ser escrita como

$$S_1 = 2\left[\ell(\hat{\beta}_1, \widehat{\boldsymbol{\beta}}_2, \hat{\sigma}^2) - \ell(\beta^*, \widetilde{\boldsymbol{\beta}}_2, \tilde{\sigma}^2)\right].$$

Intervalos de confiança a partir de estatísticas do teste

Suponham que o nosso interesse é testar,

$$\begin{cases} \mathcal{H}: & \beta_1 = \beta^* \\ \mathcal{A}: & \beta_1 \neq \beta^* \end{cases}, \tag{8}$$

a estatística da razão de verossimilhança (RV) pode ser escrita como

$$S_1 = 2\left[\ell(\hat{\beta}_1, \widehat{\boldsymbol{\beta}}_2, \hat{\sigma}^2) - \ell(\beta^*, \widetilde{\boldsymbol{\beta}}_2, \tilde{\sigma}^2)\right].$$

IC baseado na estatística RV

O IC construído a partir da estatística RV é construído da seguinte forma, (algoritmo adaptado de Hosmer et al., 2013, p. 19),

1. definir um conjunto de β^* tal que $\mathcal{H}$, em (8), não seja rejeitado;

IC baseado na estatística RV

O IC construído a partir da estatística RV é construído da seguinte forma, (algoritmo adaptado de Hosmer et al., 2013, p. 19),

1. definir um conjunto de β^* tal que $\mathcal{H}$, em (8), não seja rejeitado;
2. os limites inferior e superior do intervalo serão os percentis 2,5 e 97,5 do conjunto de valores de β^*.

IC baseado na estatística RV

O IC construído a partir da estatística RV é construído da seguinte forma, (algoritmo adaptado de Hosmer et al., 2013, p. 19),

1. definir um conjunto de β^* tal que $\mathcal{H}$, em (8), não seja rejeitado;
2. os limites inferior e superior do intervalo serão os percentis 2,5 e 97,5 do conjunto de valores de β^*.

IC baseado na estatística RV

Observações:

1. O procedimento é válido para todos os β_m, $m = 1, 2, \ldots, p$;

IC baseado na estatística RV

Observações:

1. O procedimento é válido para todos os β_m, $m = 1, 2, \ldots, p$;
2. Para a construção de regiões de confiança, somente será necessário mudar as restrições em $\mathcal{H}$. Por exemplo, $\mathcal{H} : \beta_m = \beta^*, \beta_{m'} = \beta^{**}$;

IC baseado na estatística RV

Observações:

1. O procedimento é válido para todos os β_m, $m = 1, 2, \ldots, p$;
2. Para a construção de regiões de confiança, somente será necessário mudar as restrições em $\mathcal{H}$. Por exemplo, $\mathcal{H} : \beta_m = \beta^*, \beta_{m'} = \beta^{**}$;
3. Como mencionado, IC baseados nas estatísticas de Wald, escore e gradiente também podem ser construído. Em alguns casos, poderão ter forma fechada.

IC baseado na estatística RV

Observações:

1. O procedimento é válido para todos os β_m, $m = 1, 2, \ldots, p$;
2. Para a construção de regiões de confiança, somente será necessário mudar as restrições em $\mathcal{H}$. Por exemplo, $\mathcal{H} : \beta_m = \beta^*, \beta_{m'} = \beta^{**}$;
3. Como mencionado, IC baseados nas estatísticas de Wald, escore e gradiente também podem ser construído. Em alguns casos, poderão ter forma fechada.

Roteiro

Aplicação

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

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

Aplicação

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, $\ell = 1, 2, \ldots, 10$.

Aplicação

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, $\ell = 1, 2, \ldots, 10$.

Exemplo

Nós temos também 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

QMRes $= 7{,}5$.

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 8) = 2{,}31$. Então, o intervalo de confiança baseado na estatística t, é dado por:

$$\text{IC}(\beta_1; 0{,}95) = [4{,}23;\ 15{,}77]\,,$$

$$\text{IC}(\beta_2; 0{,}95) = [1{,}89;\ 2{,}11]\,.$$

Aplicação

Aplicação 2. (Neter et al., 1983, p. 247) Conjunto de dados da *Zarthan Company*. Após o ajuste, nós temos o seguinte modelo estimado,

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

Aplicação 2. (Neter et al., 1983, p. 247) Conjunto de dados da *Zarthan Company*. Após o ajuste, nós temos o seguinte modelo estimado,

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

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

Aplicação 2. (Neter et al., 1983, p. 247) Conjunto de dados da *Zarthan Company*. Após o ajuste, nós temos o seguinte modelo estimado,

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

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

Exemplo

Nós temos também 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

QMRes $= 4{,}74$.

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 12) = 2{,}179$. Então, o intervalo de confiança baseado na estatística t, é dado por:

$$\text{IC}(\beta_1; 0{,}95) = [-1{,}843;\ 8{,}749]\,,$$

$$\text{IC}(\beta_2; 0{,}95) = [0{,}483;\ 0{,}509]\,,$$

$$\text{IC}(\beta_3; 0{,}95) = [0{,}007;\ 0{,}011]\,.$$

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 12) = 2{,}179$. Então, o intervalo de confiança baseado na estatística t, é dado por:

$$\text{IC}(\beta_1; 0{,}95) = [-1{,}843;\ 8{,}749]\,,$$

$$\text{IC}(\beta_2; 0{,}95) = [0{,}483;\ 0{,}509]\,,$$

$$\text{IC}(\beta_3; 0{,}95) = [0{,}007;\ 0{,}011]\,.$$

Como nós podemos ver, o IC para β_1 contém o zero, isto é, o intercepto não foi significativo.

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 12) = 2{,}179$. Então, o intervalo de confiança baseado na estatística t, é dado por:

$$\text{IC}(\beta_1; 0{,}95) = [-1{,}843;\ 8{,}749]\,,$$

$$\text{IC}(\beta_2; 0{,}95) = [0{,}483;\ 0{,}509]\,,$$

$$\text{IC}(\beta_3; 0{,}95) = [0{,}007;\ 0{,}011]\,.$$

Como nós podemos ver, o IC para β_1 contém o zero, isto é, o intercepto não foi significativo.

Exemplo

A Tabela 3 apresenta o ajuste considerando um modelo sem intercepto:

Tabela 3: Estimativas do parâmetros.

Parâmetro	Estimativa	EP	t_c
β_2	0,495	0,006	78,99
β_3	0,010	0,001	18,83

QMRes $= 5{,}111$.

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 12) = 2{,}18$. Então, o intervalo de confiança base-ado na estatística t, é dado por:

$$\text{IC}(\beta_2; 0{,}95) = [0{,}482;\ 0{,}509]\,,$$

$$\text{IC}(\beta_3; 0{,}95) = [0{,}009;\ 0{,}012]\,.$$

Exemplo

Para $\alpha = 0{,}05$, $t(0{,}975; 12) = 2{,}18$. Então, o intervalo de confiança baseado na estatística t, é dado por:

$$\text{IC}(\beta_2; 0{,}95) = [0{,}482;\ 0{,}509]\,,$$

$$\text{IC}(\beta_3; 0{,}95) = [0{,}009;\ 0{,}012]\,.$$

Roteiro

Referências bibliográficas I

Dorfman, R. (1938), 'A note on the δ-method for finding variance formulae', *The Biometric Bulletin* **1**, 129–137.

Gosset "Student", W. S. (1908), 'The probable error of a mean', *Biometrika* **6**(1), 1–25.

Hosmer, D. W., Lemeshow, S. e Sturdivant, R. X. (2013), *Applied logistic regression*, 3rd edn, Wiley, New Jersey.

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.

Scheffé, H. (1959), *The Analysis of Variance*, Wiley, New York.

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.

Working, H. e Hotelling, H. (1929), 'Applications of the theory of error to the interpretation of trends', *Journal of the American Statistical Association* **24**(165A), 73–85.

Obrigado!

✉ tiago.magalhaes@ufjf.br

ufjf.br/tiago_magalhaes

Departamento de Estatística, Sala 319