3ª Lista de Exercícios - SME0812 Modelos Lineares - 14/09/2012
Entregar Exercícios 10, 13 e 14 em 28/09/2012

Exercício 1. Considere o modelo $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ com $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 I_n)$ e X de posto completo $k+1$. Se $\widehat{\boldsymbol{\epsilon}} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - X\widehat{\boldsymbol{\beta}}$, prove que

(i) $\mathrm{Cov}(\widehat{\boldsymbol{\epsilon}}, \widehat{\boldsymbol{\beta}}) = \mathbf{0}$

(ii) $\mathrm{Cov}(\boldsymbol{\epsilon}, \widehat{\boldsymbol{\beta}}) = X(X'X)^{-1}\sigma^2$

(iii) $\mathrm{Cov}(\boldsymbol{\epsilon}, \mathbf{Y}) = [I - X(X'X)^{-1}X']\sigma^2$

(iv) $\mathrm{Cov}(\boldsymbol{\epsilon}, \widehat{\mathbf{Y}}) = \mathbf{0}$.

Com base nos itens (iii) e (iv), discuta por que é preferível a análise gráfica de $\boldsymbol{\epsilon} \times \widehat{\mathbf{Y}}$ em relação à análise gráfica de $\boldsymbol{\epsilon} \times \mathbf{Y}$.

Exercício 2. Seja $\mathbf{Y} \sim N_n(X\boldsymbol{\beta}, \sigma^2 I_n)$ com X de posto completo p. Prove que

$$\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \frac{X'X}{\sigma^2} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \sim \chi^2_p.$$

Exercício 3. Considere um modelo linear geral $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, com intercepto, e as partições

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \boldsymbol{\beta}_1 \end{bmatrix} \text{ com } \boldsymbol{\beta}_1 = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_k \end{bmatrix}' \text{ e } X = [\mathbf{1} \;\; X_1].$$

Considere $S = X_1'X_1 - n\overline{\mathbf{x}}\overline{\mathbf{x}}'$ e $\overline{\mathbf{x}}' = [\overline{x}_1 \; \overline{x}_2 \; \dots \; \overline{x}_k]$, em que $\overline{x}_i$ é a média dos elementos da $(i+1)$-ésima coluna de X, $i = 1, \dots, k$. Prove que

(i) $\widehat{\boldsymbol{\beta}}_1 = S^{-1}\left[X_1'\mathbf{Y} - \dfrac{X_1'\mathbf{1}\mathbf{1}'\mathbf{Y}}{n}\right] = S^{-1}(X_1'\mathbf{Y} - n\overline{Y}\overline{\mathbf{x}}')$

(ii) $\widehat{\beta}_0 = \overline{Y} - \overline{\mathbf{x}}'S^{-1}(X_1'\mathbf{Y} - n\overline{Y}\overline{\mathbf{x}}) = \overline{Y} - \overline{\mathbf{x}}'\widehat{\boldsymbol{\beta}}_1$

Exercício 4. No modelo $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, com $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 I_n)$, determine a distribuição do vetor de resíduos $\widehat{\boldsymbol{\epsilon}} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - X\widehat{\boldsymbol{\beta}}$.

Exercício 5. Para o modelo $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, \dots, 20$, onde $\displaystyle\sum_{i=1}^{20} y_i^2 = 90$ e $\epsilon_i \sim N(0, \sigma^2)$ independentes, o sistema de equações normais obtido foi

$$\begin{bmatrix} 20 & 10 \\ 10 & 6 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 40 \\ 22 \end{bmatrix}$$

(a) Calcule $\widehat{\boldsymbol{\beta}}'X'\mathbf{Y}$ e SQRes$=\mathbf{Y}'\mathbf{Y} - \widehat{\boldsymbol{\beta}}'X'\mathbf{Y}$.

(b) Construa um intervalo de confiança com coeficiente de confiança 0,95 para $E(Y|x=3)$.

(c) Construa o correspondente intervalo de predição.

Exercício 6. Sejam

$$\begin{aligned} Y_1 &= \theta + \epsilon_1 \\ Y_2 &= 2\theta - \gamma + \epsilon_2 \\ Y_3 &= \theta + 2\gamma + \epsilon_3 \end{aligned}$$

onde $E(\epsilon_i) = 0$, $i = 1, \dots, 3$. Determine os estimadores de mínimos quadrados de θ e γ.

Exercício 7. Prove que o coeficiente de explicação do modelo, R^2, em um modelo de regressão linear simples é o quadrado do coeficiente de correlação entre $\mathbf{Y}$ e $\widehat{\mathbf{Y}}$.

Exercício 8. No modelo de regressão linear simples com as suposições usuais, determine o estimador de máxima verossimilhança de β_0 e β_1 sob a hipótese $H_0 : \beta_0 = 2$.

Exercício 9. Num problema de estocar sorvetes a baixa temperatura, verifica-se que a perda de sorvete Y está relacionada com o tempo de estocagem através do modelo $Y = \beta t + \epsilon$. Foi conduzido um experimento em que registrou-se a perda de sorvete (Y, em cm^3) em função do tempo (t, em semanas), com os seguintes resultados

t_i	1	2	3	4	5	6	7	8
y_i	2,1	2,81	3,04	3,1	6,24	8,01	5,79	8,38

Admitindo que $\epsilon_i \sim N(0, \sigma^2)$, $i = 1, \ldots, 8$ e que são v. a.'s independentes,

(a) Estime β e σ^2.

(b) Determine um intervalo com 95% de confiança para β e σ^2.

(c) Calcule o coeficiente de explicação do modelo.

(d) Teste, ao nível de significância 0,05, as hipóteses $H_0 : \beta = 0$ contra $H_1 : \beta \neq 0$.

(e) Determine o poder do teste para $\beta/\sigma = 0,2$.

Exercício 10. Considere um modelo linear em que

$$\begin{aligned}
\mathrm{E}(Y_{1i}) &= \theta,\ i = 1, \ldots, m,\\
\mathrm{E}(Y_{2i}) &= \theta + \gamma,\ i = 1, \ldots, m,\\
\mathrm{E}(Y_{3i}) &= \theta - 2\gamma,\ i = 1, \ldots, n
\end{aligned}$$

em que todas as observações estão sujeitas a erros independentes com média 0 e variância σ^2.

(a) Determine os estimadores de mínimos quadrados de θ e γ.

(b) Prove que esses estimadores são não correlacionados se $m = 2n$.

Exercício 11. Mostre que $(\mathbf{Y} - X\boldsymbol{\beta})'(\mathbf{Y} - X\boldsymbol{\beta}) = (\mathbf{Y} - X\widehat{\boldsymbol{\beta}})'(\mathbf{Y} - X\widehat{\boldsymbol{\beta}}) + (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'X'X(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ com $\widehat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{Y}$ e conclua que $(\mathbf{Y} - X\boldsymbol{\beta})'(\mathbf{Y} - X\boldsymbol{\beta})$ é minimizada para $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$.

Exercício 12. Suponha que, em um modelo linear geral com intercepto, multiplicamos todos os valores das variáveis independentes de modo que $x_{ij} = k_j w_{ij}$ para todo i e j. Expressando a matrix X em termos de uma nova matriz W, prove que $\widehat{\mathbf{Y}}$ permanece inalterado.

Exercício 13. Considere o modelo linear geral $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ com $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 I_n)$ e $\widehat{\boldsymbol{\beta}}_0$ o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ sob a hipótese de que $C\boldsymbol{\beta} = \mathbf{m}$. Prove que $C\widehat{\boldsymbol{\beta}}_0 = \mathbf{m}$.

Exercício 14. Decidiu-se ajustar um modelo do tipo $Y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$ sendo que os erros são variáveis aleatórias independentes com distribuição $N(0, \sigma^2)$ com base nos dados a seguir

x_1	1	1	1	2	2	2	0	0
x_2	1	-1	0	1	0	-1	0	1
Y	8,5	3	5,5	12,5	10	7	2	5

(a) Obtenha o modelo ajustado $\widehat{Y} = \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2$.

(b) Prove que, nesse caso, os estimadores $\widehat{\beta}_1$ e $\widehat{\beta}_2$ são variáveis aleatórias independentes e determine suas distribuições marginais.

(c) Construa a tabela de análise de variância e teste ao nível de significância de 0,05 a hipótese $H_0 : \beta_1 = \beta_2 = 0$ contra $H_1 :$ pelo menos um β_i diferente de zero, $i = 1, 2$.

(d) Prove que o poder deste teste para hipótese alternativa $H_1 : \beta_1 = \beta_2 = 1$, admitindo que $\sigma^2 = 2$ é dado por $P(W \geq 5,14)$ em que $W \sim F_{2,6,5}$.