MI 634 - Análise de dados longitudinais
Segundo semestre de 2018
Lista de Exercícios I.

1. Resolva TODOS os exercícios deixados em sala.

2. Considere o modelo $Y_i = \beta_0 + \beta_1 x_i + \xi_i, i = 1, 2, ..., n$, em que $\xi_i \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. Defina $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$, em que $\widehat{\beta}_j, j = 0, 1$ é o estimador de MQO de $\beta_j, j = 0, 1$ e $\widehat{\xi}_i = Y_i - \widehat{Y}_i$. Prove que:

   a) Nesse caso, os estimadores de MV coincidem com os estimadores de MQO.
   b) $\sum_{i=1}^{n} \widehat{\xi}_i = 0$.
   c) $\sum_{i=1}^{n} x_i \widehat{\xi}_i = 0$.
   d) $\sum_{i=1}^{n} \widehat{Y}_i \widehat{\xi}_i = 0$.

3. Considere o modelo $Y_i = \beta_0 + \beta_1(x_i - \overline{x}) + \xi_i, i = 1, 2, ..., n$, em que $\xi_i \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. Encontre a forma escalar dos estimadores de mínimos quadrados ordinários (MQO) de $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ e as respectivas distribuições marginais. Compare esses estimadores com aqueles obtibos no modelo $Y_i = \alpha_0 + \alpha_1 x_i + \xi_i$. Calcule $\text{Cov}(\widehat{\beta}_0, \widehat{\beta}_1)$ e $\text{Cov}(\widehat{\alpha}_0, \widehat{\alpha}_1)$. O que você pode dizer sobre a independência entre $\widehat{\beta}_0$ e $\widehat{\beta}_1$ e entre $\widehat{\alpha}_0$ e $\widehat{\alpha}_1$?

4. Nas questões de 2 e 3, proponha um teste para testar $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, em que θ_0 é um valor conhecido e θ reresenta cada um dos parâmetros do modelo em questão. Por exemplo, no modelo da questão 2), θ corresponde à β_0 ou β_1. Propor um teste significa: constuir uma estatística do teste que faça sentido em termos do problema, apresentar sua distribuição (exata ou assintótica) sob H_0, apresentar as regiões crítica e de aceitação, bem como o p-valor associado. Considere um nível de significância de α. Considere σ^2 conhecido e depois desconhecido.

5. Seja $\boldsymbol{A}_{(m \times m)}$ uma matriz real e simétrica (sugestão: utilize a decomposição apresentada na página 19 dos slides sobre álgebra de matrizes). Mostre que:

   a) $tr(\boldsymbol{A}) = \sum_{i=1}^{m} \lambda_i$, em que $\lambda_i, i = 1, ..., m$ são os autovalores de $\boldsymbol{A}$.
   b) $det(\boldsymbol{A}) = \prod_{i=1}^{m} \lambda_i$.
   c) Se $\boldsymbol{P}$ for uma matriz ortogonal, os autovalores de $\boldsymbol{PAP}'$ são os mesmos de $\boldsymbol{A}$.

6. Seja $\boldsymbol{X} = (X_1, X_2, X_3)'$ uma vetor aleatório tal que

$$Cov(\boldsymbol{X}) = \begin{bmatrix} 7 & 1 & 3 \\ 1 & 3 & 0 \\ 3 & 0 & 5 \end{bmatrix}$$

Responda os itens:

a) Calcule a variância de $X_1 - 3X_2 + 2X_3$.

b) Calcule $Cov(\boldsymbol{Z})$ em que $\boldsymbol{Z} = (Z_1, Z_2)$ com $Z_1 = X_1 + X_2$ e $Z_2 = X_1 + X_2 + X_3$.

7. Considere o modelo de regressão linear normal homocedástico usual em sua forma matricial, ou seja, $\boldsymbol{Y}_{(n\times 1)} = \boldsymbol{X}_{(n\times p)}\boldsymbol{\beta}_{(p\times 1)} + \boldsymbol{\xi}_{(n\times 1)}$, em que $\boldsymbol{\xi} \sim N_n(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. Defina $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$, $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, $\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$, $\boldsymbol{R} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}}$ e $\widehat{\boldsymbol{\theta}}_i = \boldsymbol{C}_i\widehat{\boldsymbol{\beta}} - \boldsymbol{M}_i$ em que $\boldsymbol{C}_{i(a_i\times p)}$ são matrizes não-aleatórias tais que $r(\boldsymbol{C}_i) = a_i < p$ e $\boldsymbol{M}_{i(a_i\times 1)}$ são vetores não aleatórios, $i = 1, 2...$. Obtenha as distribuições conjuntas de i) $\widehat{\boldsymbol{\beta}}$ ii) $\widehat{\boldsymbol{\theta}}_i$ iii) $(\widehat{\boldsymbol{Y}}, \boldsymbol{R})'$, iv) $(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)$ e v) $(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\theta}}_1)$.

8. Os dados apresentados no arquivo “Macchione1995” são oriundos de um estudo realizado na Faculdade de Medicina da Universidade de São Paulo para avaliar os efeitos de agentes oxidantes no sistema respiratório. Espera-se que a exposição à maiores concentrações desses agentes causem danos crescentes aos cílios e às células secretoras de muco, os principais meios de defesa do sistema respiratório. Com essa finalidade, 56 palatos de rãs foram aleatoriamente alocados a seis grupos com 10 palatos em cinco deles e 6 em outro. Os palatos de cada um desses grupos foram imersos por 35 minutos numa solução de peróxido de hidrogênio com uma concentração especificada, a saber, 0, 1, 8, 16, 32 ou 64 microM (que correspondem aos grupos). A velocidade de transporte mucociliar (mm/s) (variável resposta) foi observada a cada 5 minutos após a imersão (considere que, quanto maior, melhor). A concentração de peróxido deve ser considerada como um fator entre unidades e o tempo como a condição de avaliação. Responda aos itens:

   a) Classifique o estudo longitudinal com relação aos aspectos visto em sala de aula.

   b) Faça uma análise descritiva conforme visto em sala de aula, desconsiderando a concentração de peróxido e o último tempo.

   c) Repita o item b) desconsiderando a concentração mas considerando o último tempo.

   d) Repita o item b) considerando todo o conjunto de dados.