MAE 5725 Modelos Lineares 1ª. Prova 27/05/2021

A prova totaliza 12,0 pontos. Notas acima de 10,0 serão arredondadas para 10,0.

Enviar uma cópia digitalizada para o email selian@ime.usp.br até as 2:00h do dia 28/05.

(1) Seja $\underset{\sim}{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$, $\underset{\sim}{X} \sim N_3\left(\underset{\sim}{\mu}, \sum\right)$ com $\underset{\sim}{\mu}' = [2 \ \ -3 \ \ 1]$ e $\sum = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$.

Determine:

a) A distribuição conjunta de X_1 e X_2.

b) A distribuição condicional de X_3 dado $X_1 = x_1$ e $X_2 = x_2$.

c) O coeficiente de correlação parcial entre X_1 e X_2 dado X_3, $\rho_{12\cdot 3}$.

d) Um vetor $\underset{\sim}{a}$, 2×1, tal que X_2 e $X_2 - \underset{\sim}{a}' \begin{bmatrix} X_1 \\ X_3 \end{bmatrix}$ sejam independentes. Justifique todos os fatos utilizados. (2,0 pontos)

(2) Os dados a seguir são provenientes de um estudo onde desejava-se analisar o efeito da temperatura na potência de um antibiótico.Para uma amostra de 10 observações foram fixados os valores da Temperatura (X) e observou-se o valor da correspondente variável resposta Potência do antibiótico(Y):

X	30^0	30^0	50^0	50^0	50^0	70^0	70^0	70^0	90^0	90^0
Y	38	43	32	26	33	19	27	23	14	21

Admitindo o modelo de regressão simples $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$, variáveis aleatórias independentes, i = 1,2,...10,

a) Obtenha as estimativas de mínimos quadrados de β_0 e β_1.
b) Teste, ao nível de significância de 0,05, a hipótese H_0: $\beta_0 = 30$, $\beta_1 = 1$ contra a alternativa de que pelo menos uma das igualdades não se verifica.
c) Construa um intervalo de confiança com coeficiente de confiança de 0,95 para β_1.
d) Caso desejássemos calcular o poder do teste do item b) sob H_a : $\beta_0 = 20$, $\beta_1 = 2$, supondo que $\sigma^2 = 10$, especifique a distribuição de probabilidades a ser utilizada, seus graus de liberdade e parâmetro de não centralidade. (3,0 pontos)

(3) Considere o modelo linear geral $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$ com $\underset{\sim}{\epsilon} \sim N_n(\underset{\sim}{O}, \sigma^2 I)$.

a) Determine W, o estimador ótimo (não viciado uniformemente de variância mínima) de $E\left(\underset{\sim}{\omega}'\underset{\sim}{Y}\right)$, onde $\underset{\sim}{\omega}$ é um vetor $n \times 1$ de constantes conhecidas.
b) Mostre que se $\underset{\sim}{b}'\underset{\sim}{Y}$ é tal que $\underset{\sim}{b}'X = 0$, então $\mathsf{Cov}(W, \underset{\sim}{b}'\underset{\sim}{Y}) = 0$.
(1,5 pontos)

(4) Seja $\underset{\sim}{Y} \sim N_n(\underset{\sim}{\mu}, I_n)$ e $\underset{\sim}{Z} = A\underset{\sim}{Y}$ onde A é uma matriz $p \times n$, de posto p, com $p < n$.
a) Qual é a distribuição de probabilidades de $(\underset{\sim}{Z} - A\underset{\sim}{\mu})'(AA')^{-1}(\underset{\sim}{Z} - A\underset{\sim}{\mu})$?
b) Qual é a distribuição de probabilidades de $\underset{\sim}{Z}'(AA')^{-1}\underset{\sim}{Z}$? (1,5 pontos)

(5) Seja

$$\begin{aligned} Y_1 &= \theta_1 + \theta_2 + \epsilon_1 \\ Y_2 &= 2\theta_2 + \epsilon_2 \\ Y_3 &= -\theta_1 + \theta_2 + \epsilon_3 \end{aligned}$$

com $\epsilon_i \sim N(0, \sigma^2)$, $i = 1, 2, 3$, independentes. Determine o estimador de mínimos quadrados de (θ_1, θ_2) sob a hipótese $H_o : \theta_1 + \theta_2 = 3$. (2,0 pontos)

---

(6) Com base no artigo

The Hat Matrix and ANOVA – David C. Hoaglin and Roy E. Welsch

The American Statistician, February 1978, vol. 32, No. 1, 17-22.

Responda às questões:

1) Qual é o objetivo do artigo?

2) Quais os resultados utilizados que foram vistos no curso?

3) Quais propriedades da matriz "hat" são consequência do fato dela ser uma matriz de projeção?

4) Qual é a regra para determinação de pontos de alavanca com base na matriz "hat" ? Como essa regra foi obtida?

5) Por que a matrix "hat" é útil na determinação de "outliers" multidimensionais?

6) Qual é o valor de h_{ii}, no modelo de regressão linear simples? Quais pontos são de alta alavanca nesse caso?

7) Qual é a utilidade da fórmula (5.1) ?

8) Qual é a expressão do resíduo "studentizado" e qual é a sua utilidade? (2,0 pontos)