3ª Lista de Exercícios - SME0812 Modelos Lineares - 1º/2014

1 Considere o modelo $Y_i = \beta x_i + \epsilon_i$, com $\epsilon_i \sim N(0, \sigma_i^2)$ variáveis aleatórias independentes e $\sigma_i^2 = \sigma^2 x_i^2$, $i = 1, \ldots, n$.

(a) Obtenha o estimador não viesado de variância uniformemente mínima de β.

(b) Determine a variância deste estimador.

(c) Obtenha um intervalo com coeficiente de confiança de $100(1-\alpha)\%$ para β, $\alpha \in (0, 1)$.

2 Deseja-se comparar dois métodos de ensino A e B e com esse objetivo, aplicou-se uma prova a n_1 alunos sujeitos ao método de ensino A e n_2 alunos sujeitos ao método de ensino B. Admitindo o modelo

$$Y_i = \beta_0 + \beta_1 W_i + \epsilon_i,$$

com ϵ_i e ϵ_j variáveis aleatórias independentes, $\epsilon_i \sim N(0, \sigma^2)$ e

$$W_i = \begin{cases} 0, & \text{para o método A e} \\ 1, & \text{para o método B} \end{cases}$$

e Y_i a nota da prova do i-ésimo aluno, $i = 1, \ldots, n$.

(a) Qual o significado de β_0 e β_1?

(b) Obtenha o estimador não viesado de variância uniformemente mínima da diferença das notas médias para alunos sujeitos aos métodos A e B, justificando sua resposta.

(c) Construa um teste de hipóteses para a igualdade das notas médias seguindo os dois métodos, contra a hipótese alternativa de que as médias são diferentes. Utilize a forma geral $\boldsymbol{C\beta} = \mathbf{m}$.

(d) Prove que o teste do item (c) é equivalente ao teste t de igualdade de médias para amostras independentes.

3 Considere o modelo $Y_i = \alpha + \beta x_i + \epsilon_i$ com $\epsilon_i \sim N(0, \sigma^2 x_i^2)$, ϵ_i e ϵ_j variáveis aleatórias independentes para $i \neq j$. Considere os valores de X e Y para uma amostra de cinco observações abaixo.

X	1	2	5	5	10
Y	13	10	20	15	50

(a) Obtenha os estimadores não viciados de variância uniformemente mínima para α e β e as correspondentes estimativas.

(b) Teste, ao nível de significância $\alpha = 0,10$, a hipótese $H_0 : \beta = 1$ contra $H_1 : \beta > 1$.

4 Obtenha os estimadores de mínimos quadrados generalizados para os parâmetros do modelo

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \; i = 1, \ldots, n,$$

em que com ϵ_i e ϵ_j variáveis aleatórias independentes para $i \neq j$ e sabendo que

$$\text{Var}(Y|X) = \begin{cases} \sigma^2, & \text{para } X = 1, \; 2 \text{ ou } 3, \\ 2\sigma^2, & \text{para } X = 4 \text{ ou } 5 \end{cases}.$$

Os dados são os seguintes:

X	1	1	2	2	3	3	4	4	5	5
Y	2,5	3,5	3,5	4,5	4,5	5,5	5,0	7,0	6,0	8,0

Admitindo normalidade dos erros, teste a hipótese $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 \neq 0$.

5 Considere o modelo

$$Y_1 = \theta_1 + \epsilon_1,\ Y_2 = 2\theta_1 - \theta_2 + \epsilon_2 \quad \text{e} \quad Y_3 = \theta_1 + 2\theta_2 + \epsilon_3.$$

com $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3)' \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I}_3)$. Obtenha o teste para avaliar se $H_0 : \theta_1 = \theta_2$ contra $H_1 : \theta_1 \neq \theta_2$.

6 Os dados a seguir correspondem a uma amostra de $n = 20$ terrenos localizados em três diferentes regiões de uma cidade, sendo X a área do terreno (em m^2) e Y o preço de venda (em u.v.).

X	200	200	250	300	300	400	400	450	500	500
Y	16	12	20	24	20	22	33	34	40	41
Região	C	A	B	B	A	A	C	B	C	C
X	550	550	600	600	600	700	700	800	800	800
Y	43	42	50	50	48	55	54	60	61	62
Região	B	A	C	C	B	B	A	A	A	A

(a) Ajuste um modelo de regressão linear a estes dados.

(b) Construa o gráfico dos resíduos contra a região e comente.

(c) Admitindo um modelo de três retas paralelas para as regiões A, B e C, obtenha as equações para as três retas ajustadas. Que restrição está sendo imposta quando assumimos o paralelismo das três retas?

(d) Teste a hipótese de que as retas para as regiões B e C são coincidentes.

(e) Teste a hipótese de que o preço independe da região.

7 Obtenha o teste da razão de verossimilhanças generalizada para $H_0 : \boldsymbol{\beta} = \mathbf{b}$ contra $H_1 : \boldsymbol{\beta} \neq \mathbf{b}$ em que $\mathbf{b}$ é um vetor de constantes conhecido no modelo com intercepto

$$\mathbf{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon} \text{ com } \boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

Como realizar o teste usando a linguagem R? E em SAS?