

MI 416 -Introdução aos Modelos Lineares

Primeiro semestre de 2013

Lista de Exercícios II

Data da entrega: até o dia 16/04/2013, às 10h00, pessoalmente, na sala 210.

Exercícios selecionados para a entrega: TODOS, com exceção da Questão 1).

1. Resolva TODOS os exercícios deixados em sala.
2. Seja $X = (X_1, \dots, X_n)'$ um vetor de variáveis aleatórias independentes e identicamente distribuídas com $E(X_i) = \mu$ e $Var(X_i) = \sigma^2$, $i = 1, \dots, n$ e considere a forma quadrática $Q = \sum_{i=1}^n (X_i - \bar{X})^2$. Utilize as propriedades das formas quadráticas para responder o exercício.
 - a) Mostre que a matriz da forma quadrática Q é idempotente.
 - b) Mostre que $Q/[n(n-1)]$ é um estimador não viciado de $Var(\bar{X})$.
 - c) Calcule a variancia de $Q/[n(n-1)]$ sob a suposição de que X tem distribuição Normal.
3. Considere o modelo linear $Y = X\beta + \xi$ cujos componentes $Y_{(n \times 1)}$, $X = [\mathbf{1}, Z]_{(n \times p)}$, com $\mathbf{1} = (1, \dots, 1)_{(n \times 1)}$ e $r(X) = p$, $\beta = [\alpha, \theta^t]_{(p \times 1)}$ e $\xi_{(n \times 1)}$ com $\mathcal{E}(\xi) = \mathbf{0}$ e $Cov(\xi) = \sigma^2 I_n$, têm a interpretação usual.
 - a) Obtenha $\hat{\theta}$, o estimador de mínimos quadrados de θ .
 - b) Mostre que $Cov(\hat{\theta}) = \sigma^2 (Z^t A Z)^{-1}$ com $A = I_n - n^{-1} \mathbf{1} \mathbf{1}^t$.
4. Considere o modelo linear $Y = X\beta + \xi$ com as suposições usuais. Mostre que o teste da razão de verossimilhanças para a hipótese $C\beta = M$ pode ser baseado na estatística F definida como quociente entre os quadrados médios devidos às hipótese e ao resíduo.

Para as questões relativas à análise de dados acesse

http://www.ime.unicamp.br/~cnaber/Material_PlaPes_2012.htm

no link “modelo de relatório” para ver uma sugestão de relatório. Não é necessário seguir integralmente o modelo, o mais importante é responder as perguntas de interesse de modo acessível para leitores que não têm conhecimento de Estatística, interpretar os resultados em termos do problema, descrever claramente os modelos considerados, apresentar os resultados em forma de tabelas, gráficos ou no próprio texto (quando for o caso) e seguir uma sequência lógica. Também, não se esqueça de, se for apropriado, fazer análises residuais para todos os modelos ajustados. Conclusões finais devem ser apresentadas.

5. Em relação ao problema da Questão 9 da Lista I, utilize o modelo proposto na lista I para responder às perguntas de interesse. Ou seja, inicie a análise com um modelo completo e reduza-o (se for o caso) até encontrar o modelo mais simples compatível com os dados. Em cada passo defina claramente os parâmetros e as hipóteses de interesse, estime os parâmetros e teste as hipóteses, e apresente os resultados de forma acessível para os pesquisadores (que não têm conhecimento de Estatística). Não se esqueça de realizar análises de resíduos para cada modelo ajustado.

6. Em relação ao problema da Questão 10 da Lista I, utilize o modelo proposto na lista I para responder às perguntas de interesse. Ou seja, inicie a análise com um modelo completo e reduza-o (se for o caso) até encontrar o modelo mais simples compatível com os dados. Em cada passo defina claramente os parâmetros e as hipóteses de interesse, estime os parâmetros e teste as hipóteses e apresente os resultados de forma acessível para os pesquisadores (que não têm conhecimento de Estatística). Não se esqueça de realizar análises de resíduos para cada modelo ajustado. Não se esqueça, também, de fazer gráficos de perfis médios observados e ajustados (pelo modelo final).
7. Os dados do arquivo Ortodontia.xls são provenientes de um estudo cujo objetivo é comparar as distribuições das distâncias nácio-sela (NS) em indivíduos brasileiros e americanos com e sem mordida aberta (MA). Os dados correspondentes aos indivíduos americanos foram obtidos da literatura científica pertinente e estão resumidos na tabela abaixo (não se tem acesso aos dados individuais). A distância NS é importante para o planejamento de cirurgias buco-faciais. Utilize um modelo linear com dois fatores (Tipo de mordida: aberta ou fechada e Procedência: Brasil ou EUA) para analisar os dados. Inicie a análise com um modelo saturado e reduza-o até encontrar o modelo mais simples compatível com os dados. Em cada passo defina claramente os parâmetros e as hipóteses de interesse, estime os parâmetros e teste as hipóteses e apresente os resultados de forma acessível para os pesquisadores (que não têm conhecimento de Estatística).

Tabela 1: Distância Nácio-sela para indivíduos americanos (mm)

Grupo	n	Média	Desvio-padrão
Com mordida aberta	88	75,6	5,0
Sem mordida aberta	88	74,7	4,6