

MI 416 -Introdução aos Modelos Lineares

Primeiro semestre de 2013

Lista de Exercícios III

Data da entrega: até o dia 03/06/2013, às 10h00 (durante a aula).

Exercícios selecionados para a entrega: 3, 5 e 6).

1. Resolva TODOS os exercícios deixados em sala.
2. Considere um modelo (sob a parametrização casela de referência) para uma estrutura (possivelmente desbalanceada) com três fatores (A,B,C) com dois níveis cada. Responda os itens:
 - a) Escreva (em função das médias e dos parâmetros do modelo) as hipóteses (nula e alternativa) relacionada à existência de interação de primeira ordem entre os fatores A e B (na presença e na ausência da interação de segunda ordem).
 - b) Escreva (em função das médias e dos parâmetros do modelo) as hipóteses (nula e alternativa) relacionada à existência de efeito do fator A e do efeito geral do fator A (na presença e na ausência das interações de primeira ordem que envolvem o FATOR A). Considere ausência de interação de segunda ordem nas duas situações anteriores.
3. Considere o modelo $Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \xi_i, i = 1, 2, \dots, n$, em que $\xi \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. (Você pode utilizar os resultados obtidos na Lista I). Responda os itens abaixo:
 - a) Encontre o mínimo (máximo) de $E(Y_i)$. Chame-o de θ .
 - b) Proponha, com base nos estimadores de mínimos quadrados, um estimador para θ .
 - c) Utilizando o método Delta, encontre a distribuição assintótica do estimador proposto no item b).
 - d) Obtenha um intervalo de confiança assintótico (com $\gamma\%$ de confiança) para θ e proponha um teste (assintótico) para testar $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ (estatística do teste, regiões crítica e de aceitação), em que θ_0 é um número conhecido, à um nível de aproximadamente α .
 - e) Encontre o teste da RV para testar $H_0 : \beta_2 = 0$ vs $H_1 : \beta_2 \neq 0$ (estatística do teste, regiões crítica e de aceitação), à um nível de significância de α . Simplifique a estatística do teste o máximo possível. Sugestão, utilize os resultados que você obteve na Questão 4 da Lista II.
4. Considere o modelo linear $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$ sob as suposições usuais mas com $\boldsymbol{\xi} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ desconhecido. Obtenha os testes de Wald e Escore de Rao (estatística do teste, regiões crítica e de aceitação), para testar $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{M}$ vs $H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{M}$ à um nível de significância de aproximadamente α . Você pode utilizar os resultados obtidos na Lista II.
5. Os dados do arquivo fosforo.txt referem-se à um estudo onde se tem o interesse em saber se a quantidade de fósforo administrada no solo afeta a produção de milho (de uma certa variedade). O fator de interesse é a quantidade de fósforo (cinco níveis), com quatro repetições por nível. A resposta é a produção de milho (quanto maior, melhor). Responda os itens abaixo na forma de relatório (tal como na lista 2). Não se esqueça, também, de controlar o nível de significância global dos testes. Certamente, com $n = 20$ observações, a análise residual torna-se menos útil mas, ainda sim, faça-a.
 - a) Realize uma análise descritiva nos dados, tendo em mente os objetivos do problema.

- b) Ajuste um modelo ANOVA para comparar as cinco médias em questão (faça a análise de variância). Tendo em mente que o fator é quantitativo, teste exatamente três hipóteses do tipo $C\beta$ para ver se existe uma tendência linear na média da resposta em função dos níveis do fator. Ajuste, se for o caso, um modelo reduzido. Não se esqueça de fazer uma análise residual para o(s) modelo(s).
- c) Ajuste um modelo de regressão linear, considerando a quantidade de fósforo como variável explicativa (sem transformá-la), interpretando os resultados e ajustando um modelo reduzido, se for o caso. Não se esqueça de fazer uma análise residual para o(s) modelo(s).
- d) Ajuste um modelo de regressão quadrático, considerando a quantidade de fósforo como variável explicativa (sem transformá-la), interpretando os resultados e ajustando um modelo reduzido, se for o caso. Não se esqueça de fazer uma análise residual para o(s) modelo(s).
- e) Compare os modelos ajustados nos itens a) e b), através do teste da RV desenvolvido na questão 3 e), e conclua qual dos dois modelos é mais apropriado para o conjunto de dados em questão.
6. Considere os dados do arquivo Sef199.xls, já discutidos em sala de aula. Utilize apenas os dados relativos ao DIA 1, considerando o índice de placa bacteriana “depois” como a variável resposta e as variáveis tipo de escova, dentrificação e índice de placa bacteriana “antes” como variáveis explicativas. Considere o grupo 1 (escova convencional com dentrificação) como referência (utilize a parametrização casela de referência). Utilize testes $C\beta = M$. Responda os itens abaixo na forma de relatório (tal como na lista 2). Para todo modelo ajustado, antes de realizar os testes, faça uma análise de resíduos. Não se esqueça, também, de controlar o nível de significância global dos testes. Quanto maior for a remoção do índice de placa bacteriana, melhor o tratamento (tipo de escova + dentrificação).

- a) Ajuste o seguinte modelo aos dados

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{ij}(x_{ijk} - \bar{x}) + \xi_{ijk},$$

$$i = 1, 2; j = 1, 2; k = 1, 2, \dots, 8; \bar{x} = \frac{1}{32} \sum_{i,j,k} x_{ijk}, \xi_{ijk} \stackrel{ind.}{\sim} N(0, \sigma^2)$$

o índice i se refere ao tipo de escova, o índice j se refere à utilização ou não do dentrificação e o índice k às unidades experimentais, x_{ijk} : índice de placa bacteriana “antes”.

- a) Faça uma análise descritiva, tendo em mente os objetivos do problema. Faça, inclusive, um gráfico de perfis médios (em função dos fatores tipo de escova e dentrificação).
- b) Ajuste o modelo acima e teste, considerando que os diferentes grupos apresentaram os mesmos índices de placa bacteriana “antes”, ausência de interação, ausência de efeitos dos fatores (se for o caso) (e possíveis igualdades entre as médias), de modo parcimonioso.
- c) Teste a igualdade dos incrementos no índice de placa bacteriana “depois” entre os quatro grupos, em relação ao aumento nos índices de placa bacteriana “antes”? Há incrementos nulos?
- d) Ajuste um modelo reduzido (se for o caso), com base nos resultados obtidos anteriormente. Qual(is) dos quatro grupos é melhor, utilizando os “critérios” do item b) e do item c)? Justifique, adequadamente, suas respostas.