

ME 714 A - Análise de dados discretos

Primeiro semestre de 2017

Trabalho (Parte II). Data da entrega: até o dia 20/06/2017, às 18h00.

Observações

- Sempre apresente as estimativas pontuais e intervalares, bem como os testes de nulidade individuais dos parâmetros de cada modelo ajustado.
- Sempre realize análises de diagnóstico apropriadas (desvio, análise residual e análise preditiva). Pelo menos o gráfico de envelopes para os RDC's (com exceção da Questão 1) e uma análise preditiva, têm de ser apresentados. Caso não se utilize outras técnicas, deve-se apresentar as devidas justificativas.
- Mesmo que o modelo não se ajuste bem ao conjunto de dados, continue a análise com ele, fazendo as devidas ressalvas (de que o modelo não está bem ajustado e de que alguma alternativa teria de ser considerada, sugerindo ao menos uma alternativa).
- Sempre deixe o intercepto no modelo, ainda que ele não seja significativo.
- Apresentar (numérica e/ou graficamente) as estimativas (pontuais e intervalares) das quantidades de interesse (médias e proporções).
- Para se testar hipóteses de interesse, utilize quantos testes (análise do desvio, teste da RV, teste $C\beta = M$ etc) julgar necessário, apresentando sempre as devidas justificativas, a menos que se solicite testes específicos.
- Utilize as técnicas de seleção de modelo que julgar conveniente, apresentando sempre as devidas justificativas, a menos que se solicite técnicas específicas.
- Sempre apresente as interpretações dos resultados.
- As metodologias utilizadas (estimação, resíduos etc) não precisam ser explicadas, somente comentadas com as respectivas referências. Entretanto, todos os modelos utilizados devem constar no relatório, com as respectivas interpretações dos parâmetros.

Questões

1. No arquivo `recrutas.dat` são descritos os resultados de um estudo desenvolvido em 1990 com recrutas americanos referente a associação entre o número de infecções de ouvido (variável resposta) e alguns fatores (covariáveis). Os dados são apresentados na seguinte ordem: hábito de nadar (ocasional ou frequente), local onde costuma nadar (piscina ou praia), faixa etária (15-19, 20-25 ou 25-29), sexo (masculino ou feminino) e número de infecções de ouvido diagnosticadas pelo próprio recruta. Para responder as questões de interesse, considere um modelo log-linear de Poisson.
 - a) Faça uma análise descritiva apropriada.
 - b) Selecione, de forma apropriada, as covariáveis de interesse, não se esquecendo de realizar as análises de diagnóstico pertinentes, pelo menos, para o modelo completo e o final.
 - c) Determine quais covariáveis mais impactam no número de infecção do ouvidos.
 - d) Caso o modelo log-linear não tenha se ajustado bem aos dados, ajuste (somente o modelo completo) o respectivo modelo log-linear binomial-negativo e comente, apenas, sobre o ajuste desse novo modelo.