

ME 613 - Análise de Regressão
Segundo semestre de 2016
Trabalho I - Parte 2
Data de entrega: até 14/12/2016, às 17h00

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 (e.g., desvio, análise residual e análise preditiva) de cada modelo ajustado. Caso nem todos os gráficos caibam no relatório, coloque aqueles que você julgar serem mais importantes.
- Mesmo que o modelo escolhido/final 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).
- Apresente (numérica e/ou graficamente) as estimativas (pontuais e intervalares) das quantidades de interesse (e.g., médias, proporções, razões de chances).
- 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 convenientes, 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. (Extraído de Paula (2013)). No arquivo reg3.dat são descritas as seguintes variáveis referentes a 50 estados norte-americanos: (i) estado (nome do estado), (ii) pop (população estimada em julho de 1975), (iii) percap (renda percapita em 1974 em USD), (iv) analf (proporção de analfabetos em 1970), (v) expvida (expectativa de vida em anos 1969-70), (vi) crime (taxa de criminalidade por 100000 habitantes 1976), (vii) estud (porcentagem de estudantes que concluem o segundo grau 1970), (viii) ndias (número de dias do ano com temperatura abaixo de zero grau Celsius na cidade mais importante do estado) e (ix) area (área do estado em milhas quadradas). O objetivo do estudo é tentar explicar e variável expvida usando um modelo de regressão normal linear dadas as variáveis explicativas percap, analf, crime, estud, ndias e dens, em que $\text{dens} = \text{pop}/\text{area}$. Ajuste um modelo com todas as covariáveis e verifique sua adequabilidade. Utilize algum método de seleção de covariáveis para chegar à um modelo com somente as covariáveis significativas. Além das análises usuais, faça uma análise de influência (para todos os modelos ajustados), conforme visto em sala.