

Universidade Federal Fluminense (UFF)

Instituto de Matemática e Estatística (IME)

Departamento de Estatística (GET)

Disciplina: Modelos Lineares I

Professor: Dr. José Rodrigo de Moraes

Nome: _____

Data: 27/09/2019 (6ª feira)

Assunto: Regressão Linear Múltipla, Seleção de modelo

Exercício de Modelos Lineares I:

Considere o banco de dados dos pacientes de $n=113$ hospitais, para os quais registrou as seguintes variáveis descritas no Quadro 1:

Quadro 1: Descrição das variáveis relativas aos hospitais

Variável	Descrição
Hospital	Identificador do hospital
Tempinterna	Tempo médio de internação dos pacientes no hospital
Idadem	Idade média dos pacientes no hospital
Pinfectados	Percentual de pacientes com infecção hospitalar
Ncamas	Número de camas no hospital
Nenfermeiros	Número de enfermeiros no hospital

Supondo que o objetivo é estudar a variação dos percentuais de pacientes com infecção hospitalar, pede-se:

- Faça gráficos de dispersão entre cada variável explicativa e a variável resposta do modelo. Faça a análise de cada gráfico avaliando o sentido e a força da relação. Complemente a sua análise utilizando o coeficiente de correlação linear de Pearson.
- Especifique o modelo completo (teórico) e descreva todos os seus termos/variável(eis), no contexto do problema.

- c) Ajuste o modelo especificado na letra (b) pelo método de mínimos quadrados ordinários (MQO), e avalie a significância dos parâmetros do modelo usando o teste de significância individual considerando o nível de 5%.
- d) Especifique a equação do modelo (ajustado) que você escolheu e interprete as estimativas dos parâmetros do modelo (contexto). Obtenha para o modelo selecionado o coeficiente de correlação linear múltipla e o coeficiente de determinação do modelo, e interprete ambos os coeficientes no contexto do problema.
- e) Agora, selecione o modelo utilizando o teste F de comparabilidade de modelos. Mostre todas as etapas até a seleção do modelo final. **OBS:** Especifique as hipóteses de cada teste realizado e utilize o método do p-valor para justificar as suas conclusões.
- f) Existe violação de alguma das hipóteses básicas do modelo? Em caso afirmativo, identifique-as. Use métodos gráficos e, caso necessário, métodos formais, como por exemplo, o teste de normalidade de Shapiro-Wilk considerando o nível de significância de 5%.
- g) Você diria que o modelo é adequado para representar os dados observados? Justifique a sua resposta.