



## LABORATÓRIO 8

### Atividade 1: Regressão Linear Múltipla - Regressão Polinomial

No arquivo “Lab08q01.txt” constam informações para casais com pelo menos um filho. Os dados são de um estudo demográfico conduzido para determinar o efeito da renda anual do marido (REND) no tempo (TEMPO, em meses) entre o casamento e o nascimento do primeiro filho. Com base nestes dados e considerando-se  $\alpha = 10\%$  objetiva-se ajustar um modelo polinomial de grau três para descrever o efeito da renda anual do marido (REND) no tempo (TEMPO, em meses).

**Observação:** Para avaliar se o modelo cúbico é adequado, faz-se necessário compara-lo com os modelos lineares simples e quadrático. Também a necessidade de verificar o problema de colinearidade. Neste caso, pode ser útil considerar a variável explicativa centrada na média [ $RENDAC = REND - \text{mean}(REND)$ ], ajustar um novo modelo cúbico e refazer a análise de colinearidade. O modelo final escolhido deve ser justificado após essas análises.

### Atividade 2: Regressão Linear Múltipla - Seleção de modelos

A base de dados `prostate` do pacote `faraway` contém informações sobre 97 homens com câncer de próstata. A variável `lpsa` que representa o logaritmo de um antígeno prostático específico deve ser considerada como variável resposta nesta atividade e as demais variáveis como explicativas. Utilize os seguintes métodos de seleção de covariáveis para determinar o “melhor” ajuste:

- (a) Eliminação **backward** baseada no teste  $F$ .
- (b) Seleção **stepwise** baseada no critério de informação de Akaike.
- (c) Seleção **stepwise** baseada no critério de informação Bayesiano.
- (d) Todas as regressões possíveis, baseado no  $R^2$  ajustado.

**Observação:** No AVA consta o script R “SelectionModel.r” que pode auxiliar nas atividades do laboratório. Seguem mais umas dicas:

- Regressão polinomial (cúbica): comando `lm(y ~ x + I(x ^ 2) + I(x ^ 3))` ou `lm(y ~ poly(x, degree = 3, raw = T))`
- Regressão polinomial (cúbica) ortogonal: comando `lm(y ~ poly(x, degree = 3, raw = F))`
- Colinearidade: use função `vif(x)` do pacote `faraway`, em que `x` é a matriz de desenho do modelo.
- Use o comando `model.matrix(modelo)` para obter a matriz de desenho `x`.

### Algumas sugestões para a redação dos relatórios:

- Lembrete: toda modelagem estatística deve ser precedida por uma análise descritiva/exploratória, composta por gráficos e medidas descritivas pertinentes.
- Sejam parcimoniosos quanto aos resultados incluídos no relatório. Obviamente, nem todos os resultados produzidos na análise precisam ser relatados. Algumas representações fundamentais:
  - Gráficos e ou tabelas de análise descritiva/exploratória;
  - Os resultados referentes ao(s) modelo(s) ajustado(s) na forma de quadros, gráficos ou tabelas;
  - Figuras (que podem ser compostas por múltiplos gráficos) referentes ao diagnóstico do ajuste;
- A depender da análise, figuras, quadros ou tabelas para outros tipos de resultados podem ser necessários. Alguns resultados (como medidas e testes de qualidade de ajuste) podem ser inseridos no próprio texto;
- Todos os quadros, tabelas e figuras deverão ter títulos e numeração. Todos eles deverão ser mencionados em algum momento no texto, com a discussão dos respectivos resultados;
- Os resultados deverão ser devidamente editados. Saídas cruas (*outputs*) do *software* R serão desconsideradas;
- Não incluir códigos de programação. O *script* utilizado deverá ser enviado separadamente pelo AVA Moodle ou como apêndice do relatório.
- As páginas do relatório deverão ser numeradas.