

Ficha de Trabalho 4

Modelos Lineares e Aplicações

Exercício 1:

Considere o conjunto de dados iris, disponível na distribuição padrão do R. Trata-se de observações em quatro variáveis morfométricas: largura e comprimento de pétalas e sépalas, para $n = 150$ lírios. Para o conjunto das $n = 150$ observações:

- (a) Construa as nuvens de pontos para cada possível par de variáveis. Comente.
- (b) Calcule a matriz de correlações entre as 4 variáveis observadas. Comente.
- (c) Ajuste uma regressão linear múltipla da largura das pétalas sobre as restantes três variáveis predictoras. Interprete os valores das estimativas dos coeficientes de cada uma das variáveis predictoras.
- (d) Discuta a qualidade de ajustamento do modelo.
- (e) Construa os intervalos a 95% de confiança para os parâmetros do modelo.
- (f) Determine a matriz de variâncias-covariâncias para os parâmetros estimados do modelo.
- (g) Efectue o teste de ajustamento global do modelo. Discuta o resultado.

Exercício 2:

Considere uma base de dados económicos *savings* relativos a 50 países. O conjunto de dados é composto pelas médias entre os anos de 1960 e 1970 das seguintes variáveis:

'sr' savings rate - personal saving divided by disposable income

'pop15' - percent population under age of 15

'pop75' - percent population over age of 75

'dpi' - per-capita disposable income in dollars

'ddpi' - percent growth rate of dpi

Investigue brevemente esta base de dados:

```
> library(faraway)
```

```
> data(savings)
```

```
> savings
```

- (a) Determina a matriz das correlações.
- (b) Represente graficamente os dados.

- (c) Ajuste um modelo de regressão com *sr* como variável resposta e as restantes 4 variáveis como preditoras. Apresente o *output* respectivo.

Todos os testes seguintes devem ser realizados a 5% nível de significância.

- (d) Qual o valor da estimativa do coeficiente de regressão de *dpi*?
- (e) Escreva e realize o teste de hipóteses relativo ao facto de que nenhum dos preditores tem um efeito linear na variável resposta.
- (f) Construa um IC para o coeficiente de regressão de *pop75*.
- (g) Estime a variância dos erros.
- (h) Prever *sr* cujo valores das variáveis explicativas no modelo anterior são (25, 4, 1000, 4). Obtenha o intervalo de predição a 95% para essa observação.