

## Lista de Exercícios aulas 8 e 9: Modelos lineares

1) Considere o data.frame “vespas”:

a) Usando as operações matriciais,

$$Y \sim X b$$

$$b \sim [X'X]^{-1} [X'Y]$$

Estime os valores de  $a$  e  $b$  dos modelos:



$$\text{Vespas} = a + b \cdot \text{Flores} + e$$

$$\text{Sementes} = a + b \cdot \text{Flores} + e$$

$$\text{Vespas} = a + b \cdot \text{Sementes} + e$$

- Dica: crie as matrizes  $Y$  e  $X$  conforme exemplo em aula

b) Usando as operações matriciais acima, estime  $a$ ,  $b_1$  e  $b_2$  do modelo:  $\text{Vespas} = a + b_1 \cdot \text{Flores} + b_2 \cdot \text{Sementes} + e$



2.a) Extraia do data frame ‘caixeta.csv’ o subconjunto de dados referente a espécie *Tabebuia cassinoides* e salve num objeto chamado ‘caixeta.Tc’.

b) Teste os modelos que circunferência na altura do peito (cap) e altura (h) são funções do local.

c) Se houve diferenças significativas, qual(is) local(is) provavelmente tiveram maiores efeitos? Use boxplots para avaliar.

d) Faça análises diagnósticas dos modelos. Os dados atendem aos pressupostos teóricos dos modelos lineares? Por que?



3.a) Repita o modelo  $\text{cap} \sim \text{local}$  para *Tabebuia cassinoides*, mas usando o objeto **caixeta** (original).

b) Ajuste o modelo  $\text{h} \sim \text{cap}$  para *Tabebuia cassinoides* na **jureira**, usando o objeto **caixeta** (original).



4) Use os dados ‘vespas.completo’

a) Ajuste o modelo Polinizador  $\sim$  Diametro + Sementes + NP1 + NP2 + NP3

b) Use a função **anova(modelo)** para avaliar a existência de possíveis variáveis colineares. Mude a ordem das variáveis e observe as **somas dos quadrados** associados a elas.

c) Quais as relações colineares mais fortes que você detectou com a anova()?



5) Considere os dados ‘vespas’. Ajuste o modelo com interações  $\text{Vespas} \sim \text{Arvore} * \text{Sementes} * \text{Flores}$  e proceda a seleção do modelo mais adequado.



6) Repita os exemplos dados em aula (no script), principalmente referentes às funções update(), drop1 e add1.



7) Considere os dados caixeta.Tc. Faça **uma** linha de comando que execute **automaticamente** o modelo  $\text{h} \sim \text{cap}$  por **local**. Esse comando deve retornar **3 análises lm**, uma para cada local.

Dica: use a função split(). Outras soluções são bem-vindas!!!