

Introdução aos Modelos Lineares em Ecologia

Prof. Adriano Sanches Melo - Dep. Ecologia – UFG
asm.adrimelo no gmail.com

Página do curso: www.ecologia.ufrgs.br/~adrimelo/lm/

Livro-texto: Crawley, M.J. 2005. Statistics: An Introduction using R.
John Wiley & Sons.

Página do livro na internet:

<http://www3.imperial.ac.uk/naturalsciences/research/statisticsusingr>

AULA 6

1. Introdução

1.1 O que é modelagem ?

“Todos os modelos estão errados, mas alguns são úteis” (Box 1976)

1.2 Modelos devem ser parcimoniosos. Ou seja, preferimos:

Modelos com $n-1$ parâmetros em relação a outro com n parâmetros;

Modelos com $k-1$ variáveis explanatórias em relação a outro com k var.;

Modelos lineares em relação a modelos que sejam curvos;

Modelos sem “corcova” em relação a modelos com “corcova”;

Modelos sem interação em relação a modelos com interação.

2. Tipos de modelos

Modelo Saturado: 1 parâmetro para cada observação

Modelo Máximo: Contém todos parâmetros, interações e covariáveis

Modelo adequado mínimo: Contém apenas variáveis importantes

Modelo nulo: Contém apenas a constante, ou seja, a média geral

3. Simplificação de modelos

3.1 Ajuste o modelo máximo

3.2 Inicie a simplificação do modelo

3.3 Retire o termo que aparentemente tem a menor importância.

Causou uma diferença significativa em “Deviance” ?

SE SIM, retorne o termo ao modelo.

SE NÃO, deixe-o fora e exclua outro termo.

3.4 Repita o procedimentos 3.2 e 3.3 até não conseguir remover termos

4. Fórmulas no R (apenas relembrando..)

$y \sim x$ (onde x é contínuo)

$y \sim \text{sexo}$ (onde sexo é categórico)

$y \sim \text{sexo} * x$
ou

$y \sim \text{sexo} + x + \text{sexo} : x$

$y \sim x + I(x^2)$ (I significa “do jeito que está”)

$y \sim 1$ (modelo com apenas a constante [a média])

$y \sim x - 1$ (modelo sem a constante)

5. Atualização de modelos (útil com modelos muito grandes)

```
modelo1<-lm(y~A*B)
```

```
modelo2<-update(modelo1, ~ . - A:B)
```

6. Comparação de modelos

```
anova(modelo1, modelo2)
```

7. Diagnóstico do modelo: O modelo é válido?

Os mesmos procedimentos vistos em aulas anteriores

8. Modelos Lineares Generalizados (GLMs)

- Extensão dos modelos lineares para variáveis respostas que não sejam:
 - distribuídas segundo uma Distribuição Normal
 - não tenham constância de variância

-Exemplos:

Valores de contagem não podem ser negativos

Valores de proporções não podem ser < 0 ou > 1

-As funções de ligação (link functions)

Tipo de Erro	Link
Normal	identity
Poisson	log
Binomial	logit

-Uso no R:

```
glm(y ~ x, family = binomial)
```

Veremos GLMs em aulas posteriores

Regressão Logística

Regressão de Poisson

Ancova

1. Introdução

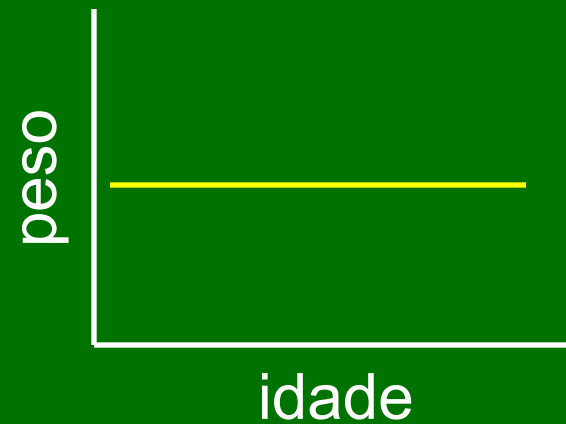
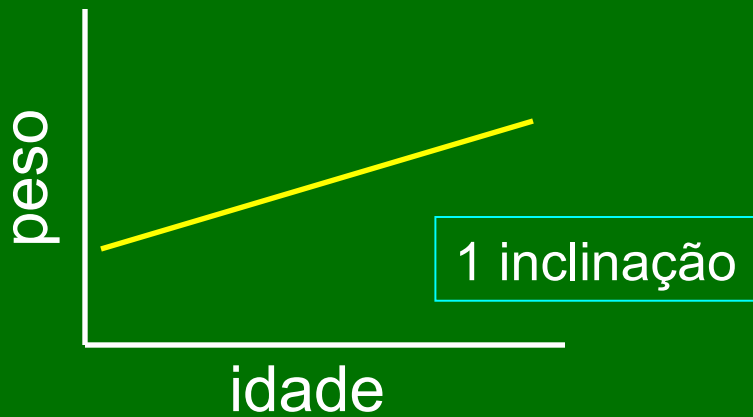
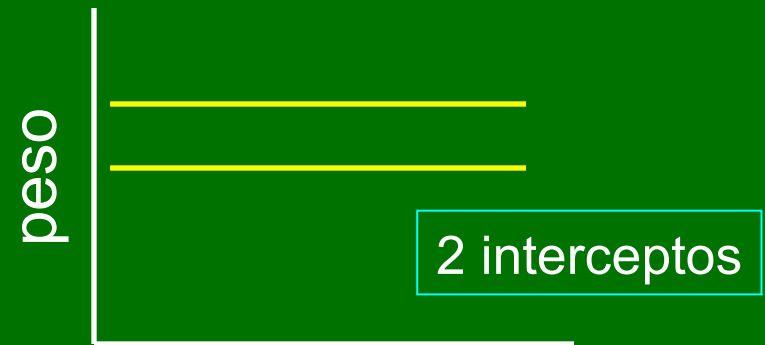
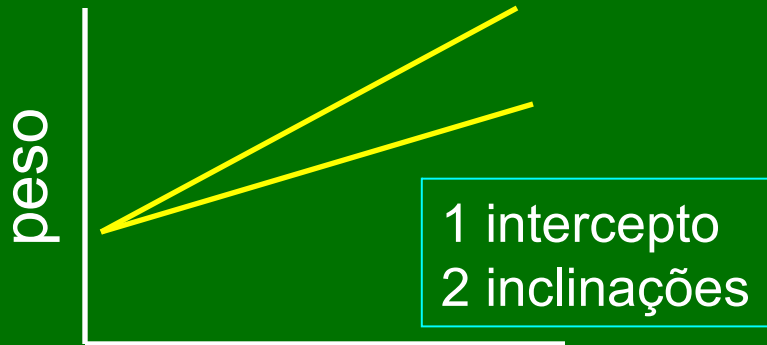
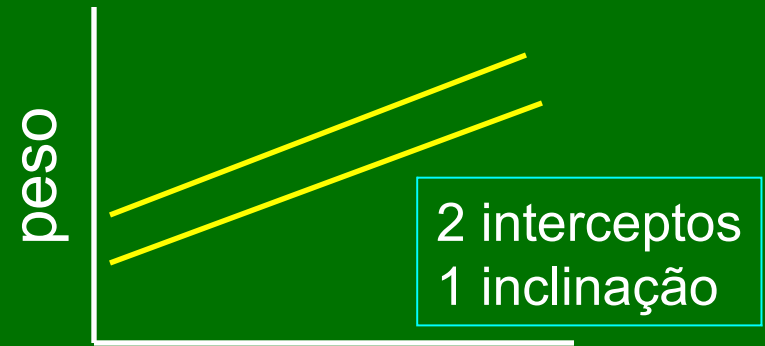
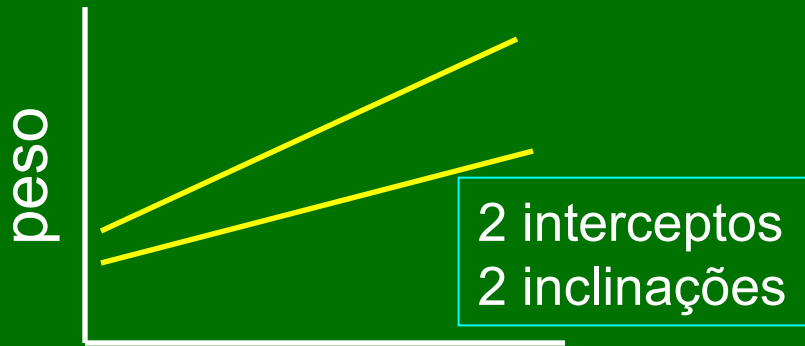
ANCOVA ou Análise de Covariância:

Um modelo misto em termos de preditores

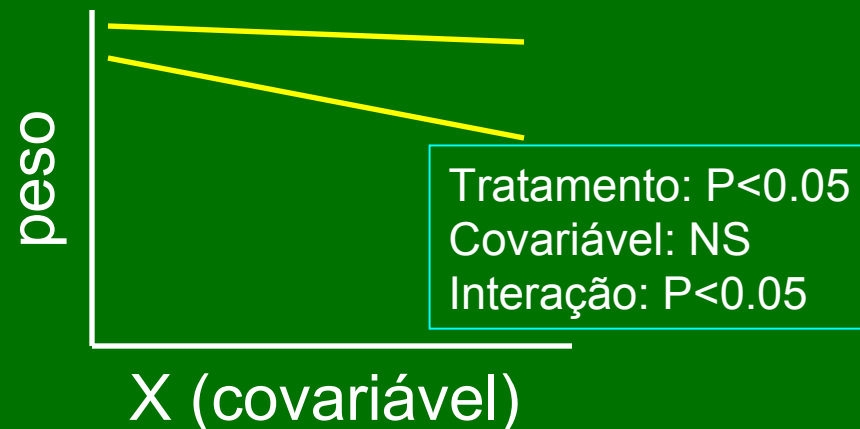
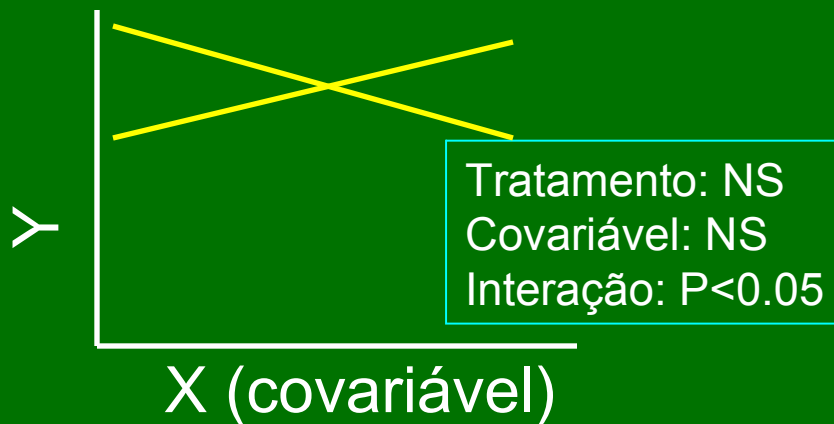
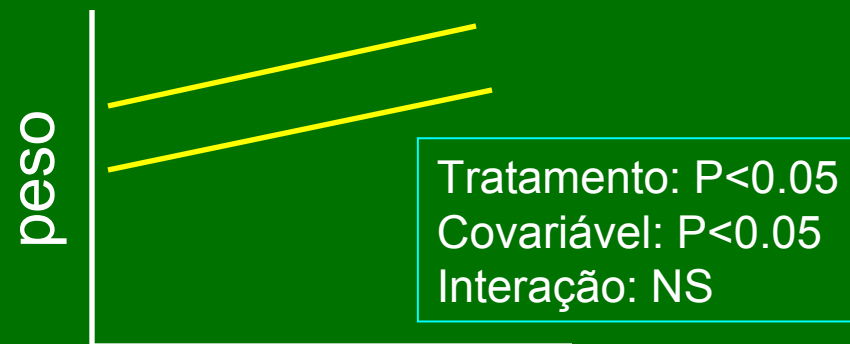
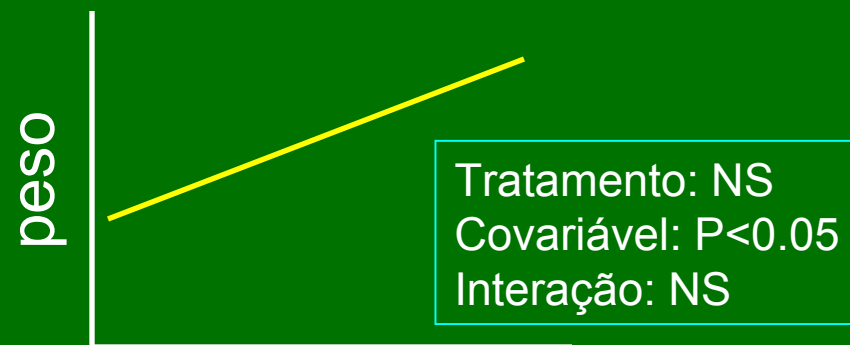
Sendo sexo uma variável categórica e idade uma variável contínua,
o modelo de ANCOVA seria:

peso ~ sexo * idade

2a. Possibilidades de resultado (Crawley p.188)



2b. Possibilidades de resultado (Gotelli & Ellison p.334)



3. Particularidades

3.1 Ordem de entrada dos termos no modelo é importante.
(Dados em geral não são ortogonais)
(O que os pacotes estatísticos fazem então?)

3.2 Interação indica diferença entre inclinações das retas.

3.3 O preditor contínuo é frequentemente chamado de covariável.

Exercícios e estudo individual:

- Lista em sala de aula
- Crawley: Cap. 7 e 10
- Gotelli e Ellison: pp. 314-317
333-334