

INTRODUÇÃO AOS MODELOS LINEARES EM ECOLOGIA

PROF. ADRIANO SANCHES MELO – asm.adrimelo no gmail.com -DEP. ECOLOGIA – UFG

PÁGINA DO CURSO: www.ecoevol.ufg.br/adrimelo/lm

AULA 3 – REGRESSÃO I

- 1) Usando apenas papel e lápis (computador não!),
- a) Considere os seguintes dados:
 Variável dependente (Y) = riqueza: 12, 14, 17, 20, 19
 Variável independente (X) = hetero: 20, 30, 40, 50, 60
 - b) Obtenha a Soma de Quadrados Total
 - c) Obtenha os coeficientes da regressão usando as formulas
 - d) Obtenha os resíduos
 - e) Obtenha a Soma de Quadrados dos Resíduos
 - f) Obtenha a Soma de Quadrados da Regressão
 - g) Faça em seu caderno um gráfico aproximado contendo i) os valores observados, ii) a reta ajustada e iii) indicação do resíduo.
 - h) Encontre os valores médios de X e de Y. O ponto definido por estas duas médias esta sobre a reta ajustada?
-
- 2) Ainda utilizando apenas papel e lápis:
- a) Considere os seguintes dados:
 Variável dependente (Y) = riqueza: 12, 12, 15, 20, 19
 Variável independente (X) = hetero: 20, 30, 40, 50, 60
 - b) Vamos ajustar uma reta com intercepto passando pela origem ($b_0=0$) por iteração. Para tanto escolha um valor de b_1 entre 0.25 e 0.45 e calcule a Soma de Quadrados do Resíduo. Use vários valores e veja qual deles minimiza a Soma de Quadrados do Resíduo.
-
- 3) No Excel, crie um "programinha" de Regressão Linear:
- a) Crie duas colunas com os dados do exercício 2). [Não esqueça de nomear cada coluna...]
 - b) Em outras colunas, obtenha as diferenças necessárias aos cálculos
 - c) Usando fórmulas, obtenha as estatísticas necessárias ao ajuste de uma reta (e.g. médias de X e Y).
 - d) Usando formulas, obtenha b_1
 - e) Usando formulas, obtenha b_0
 - f) Obtenha, por meio de fórmulas, SST, SSR e SSE e R^2
 - g) Faça um gráfico de dispersão contendo os valores observados.
 - h) Sobreponha ao gráfico os pontos ajustados e una-os com linha.
 - i) Grave em disquete ou pen-drive. Usaremos em aulas seguintes.
 - j) Altere os valores de Y e veja o que acontece com b_0 , b_1 , SST, SSR, SSE, R^2 e com o gráfico. "Brinque" um pouco.
-
- 4) No R,
- a) Crie dois vetores com dados do exercício 1).

```
riqueza<-c(12, 14, 17, 20, 19)
hetero<-c(20, 30, 40, 50, 60)
```
 - b) Faça um gráfico de dispersão e mantenha-o aberto:

```
plot(hetero, riqueza)
```
 - c) Ajuste uma reta e grave como 'ajuste':

```
ajuste<-lm(riqueza~hetero)
```
 - d) Sobreponha a reta ajustada ao gráfico de dispersão:

```
abline(ajuste)
```
 - e) Obtenha informações do ajuste:

```
summary(ajuste) ## As informações coincidem com as obtidas no exercício 1) ?
```
 - f) Obtenha mais informações

```
ajuste$residuals
ajuste$fitted
```
 - g) Veja o que você pode tirar de informações adicionais:

```
names(ajuste)
```

5) No R,

a) Crie um vetor 'riqueza1' e um vetor 'hetero1' com os dados do exercício 2)

b) Ajuste uma reta que passe pela origem:

```
ajuste1<-lm(riqueza1~0+hetero1)
```

c) Examine os resultados com comandos apropriados e veja se coincidem com os obtidos no exercício 2)

6) Obtenha a planilha do Excel feita na aula passada para regressão (dados exercício 2 aula anterior). Procure nomear as colunas e fazer algo organizado. Faça um gráfico de dispersão dos valores observados e sobreponha os valores ajustados unidos por uma linha.

a) inclua uma coluna ao lado da coluna de valores Y. Preencha as células com valores da função aleatório.

b) selecione as células de Y e de valores aleatórios. No menu, escolha –Dados->Classificar. No campo “Classificar por” escolha a coluna de números aleatórios.

c) anote os valores de b_1 neste conjunto aleatorizado. Repita o processo 5 vezes.

d) com dados dos seus colegas, construa uma distribuição dos valores de b_1 .

e) qual a probabilidade de seus dados terem sido obtidos ao acaso?

7) Repita o mesmo procedimento do exercício anterior no R. Inicialmente examine:

```
sample(y)
```

Depois, obtenha:

```
summary(aov(sample(y)~x)) para o teste F
```

```
summary(lm(sample(y)~x)) para o teste t
```

8) Vamos obter um grande número de aleatorizações no R. Usaremos as seguintes funções:

```
numeric(10000)
```

```
sample(y)
```

```
for(i in 1:10000 ) { }
```

9) Diversos fatores podem afetar a riqueza de espécies em diferentes escalas espaciais. Teste a hipótese de que existe relação entre riqueza de espécies e heterogeneidade ambiental com os dados a seguir:

Riqueza	10	16	25	22	27	15	22	19	12	14
Heterogeneidade	0.5	2	4.5	4	5	3	4	3	0.5	1.5

10) A Serra da Mantiqueira possui as maiores altitudes no sudeste do Brasil, entre os estados de SP, MG e RJ. Na face voltada ao Atlântico, a Serra possui em suas porções basais formações florestais típicas de Mata Atlântica (s.s.). Em altitudes intermediárias, elementos de clima frio tornam-se comuns. Nos topos (acima de 2200m), que podem receber neve no inverno, *existe vegetação campestre*. Apesar de estar em região tropical-subtropical, (23-24° latitude), diversas plantas de clima frio são encontradas, entre outras, *Araucaria angustifolia*. Num estudo de mestrado, Felizbino mediu a densidade (por 100 m²) de Araucaria num transecto altitudinal entre as cotas 800 a 1800. Os resultados foram os seguintes:

Altitude	800	900	1000	1100	1200	1300	1400	1500	1600	1700	1800
Densidade	0.2	0.8	0.7	2.1	2.7	2.5	3.1	4.5	3.9	4.3	4.5

a) Construa uma tabela de Análise de Variância.

b) Existe relação entre a densidade de Araucária e altitude?

c) SE houver, isto é causado por altitude?

d) É possível prever a densidade na altitude de 2200 metros? Justifique.

11) Diversos fatores podem afetar a taxa de degradação de folhas em riachos, entre eles a velocidade da água. Os dados abaixo (artificiais) referem-se à quantidade de material restante após 20 dias em 16 riachos. Existe relação entre velocidade e a quantidade de material restante?

Restante	28	35	30	38	32	40	35	41	38	45	39	46	41	48	43	50
Velocidade	38	37	35	34	35	35	32	33	30	29	27	27	25	23	21	22

a) Faça um diagrama de dispersão e depois insira a reta ajustada.

b) Obtenha a partição de variância com o comando `aov`. O valor de probabilidade associado a F ("`aov`") é igual ao valor associado a t ("`lm`")?

c) Construa uma tabela de Análise de Variância.