CC0291- Estatística Não Paramétrica

Hoffmann - 20/06/2023

Prof. Maurício Mota

1. Um ensaio de adubação forneceu os seguintes resultados

   Admite-se que as variáveis X e Y estão relacionadas de acordo com o modelo

X=dose de adubo por hectare	Y+produção por hectare
0	6;8
1	16;18
2	18;20
3	12;14

   Pode-se verificar que:

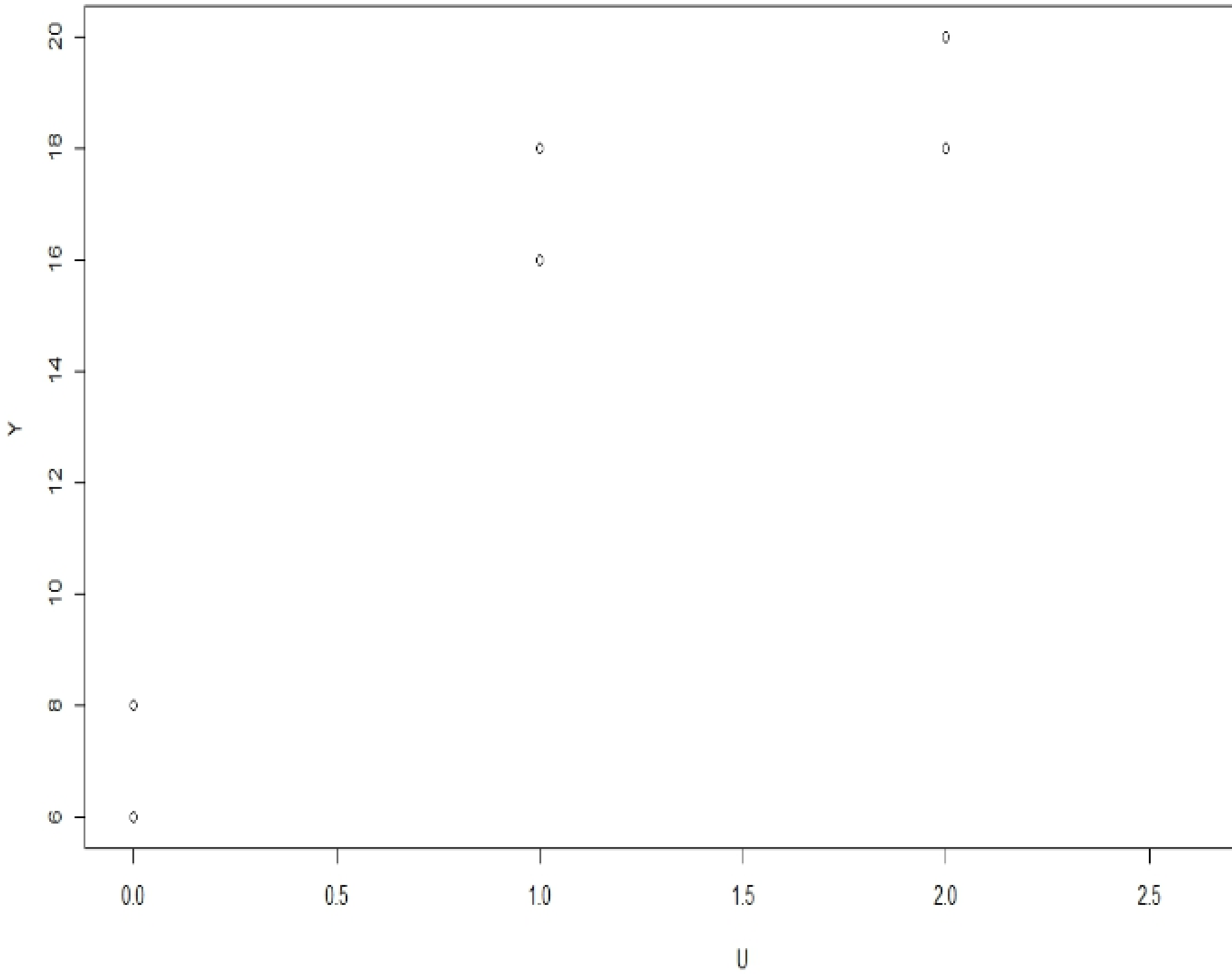
$$\sum_{i=1}^{8} Y_i = 112 \ , \ \ \bar{Y} = 12 \ \ \sum_{i=1}^{8}(Y_i - \bar{Y})^2 = 176.$$

   a. Admitindo que a função de produção seja uma parábola ,determine as estimativas dos parâmetros dessa função de acordo com o método de mínimos quadrados.
   b. Faça a análise de variância e calcule o coeficiente de determinação da regressão. Considere um nível de significância de 1%.
   c. Sabendo que a relação entre o preço do adubo e o preço do produto seja igual a, determine a quantidade economicamente ótima de adubo a ser aplicada.
   d. Verifique se o coeficiente do termo quadrático é estatisticamente diferente de zero. Pressupõe-se que a lei dos rendimentos marginais decrescentes seja válida.
   e. Teste a hipótese de que a produção máxima é obtida aplicando-se 2 doses de adubo por hectare.

      Considere na questão $\alpha = 0,01$..

Solução:

Inicialmente vamos construir um diagrama de dispersão que nos diz que o modelo proposto é bem razoável.



$$Y_i = \beta_0 + \beta_1\, U_i + \beta_2\, U_i^2 + \epsilon_i,,\ \ i = 1, 2, \ldots, n,$$

onde os $\epsilon's$ são variáveis aleatórias independentes com distribuição normal de média zero e variância σ^2.

Note que

$$E(Y_i) = \beta_0 + \beta_1\, U_i + \beta_2\, U_i^2$$

que é estimada por:

$$\bar{Y} = b_0 + b_1 U + b_2 U^2.$$

Vamos analisar a saída do R:

```
>
> Y=c(6,8,16,18,18,20,12,14)
> n=length(Y);n
[1] 8
>
> mod1=lm(Y ~U+ U2);mod1

Call:
lm(formula = Y ~ U + U2)

Coefficients:
(Intercept)            U           U2
7           14           -4

> anova(mod1)
Analysis of Variance Table

Response: Y
Df Sum Sq Mean Sq F value    Pr(>F)
U          1     40    40.0      25 0.0041047 **
U2         1    128   128.0      80 0.0002911 ***
Residuals  5      8     1.6
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1
> summary(mod1)

Call:
lm(formula = Y ~ U + U2)

Residuals:
1  2  3  4  5  6  7  8
-1  1 -1  1 -1  1 -1  1

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.0000     0.8718   8.030 0.000484 ***
U            14.0000     1.4000  10.000 0.000171 ***
U2           -4.0000     0.4472  -8.944 0.000291 ***


Residual standard error: 1.265 on 5 degrees of freedom
Multiple R-squared:  0.9545,    Adjusted R-squared:  0.9364
F-statistic:  52.5 on 2 and 5 DF,  p-value: 0.0004405
```

Nossa função de produção é estimada por:

$$\bar{Y} = 7 + 14U - 4U^2.$$

A soma de quadrados total é dada por:

$$SQTot = 176 = 40 + 128 + 8.$$

A soma de quadrados de regressão l é dada por:

$$SQReg = 40 + 128 = 168.$$

O coeficiente de determinação é dada por:

$$R^2 = \frac{SQReg}{SQTot} = \frac{168}{178} = 0,9545.$$

Vamos resolver o item c:

Sejam a o preço da dose e b o preço do produto . De acordo com o enunciado temos:

$$\frac{a}{b} = 2.$$

A receita obtida é dada por

$$Receita = a * E(Y) = b * (\beta_0 + \beta_1 \, U_i + \beta_2 \, U_i^2).$$

Seja U_0 a dose ótima . A despesa é dada por

$$Despesa = aU$$

O lucro é dado por:

$$L = Receita - Despesa = b * (\beta_0 + \beta_1 \, U + \beta_2 \, U^2) - aU$$

Derivando em relação a U temos:

$$L' = b \; \beta_1 + 2\beta_2 bU - a = 0$$

$$U = \frac{a - b\beta_1}{2b\beta_2} = \frac{a/b - \beta_1}{2 \; \beta_2} = \frac{2 - \beta_1}{2\beta_2}.$$

$$U_o = \frac{2 - \beta_1}{2\beta_2},$$

que é estimada por:

$$\hat{U}_o = \frac{2 - b1}{2b2} = \frac{-12}{-8} = 1,5.$$

Queremos testar:

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 < 0.$$

Sabemos que se H_0 é verdade temos:

$$t = \frac{b2}{s_b 2} \sim t(n-3) = t(5).$$

$$t_{cal} = \frac{-4}{0.4472} = -8.944$$

```
>
> alfa=0.01
> t_tab=qt(alfa,n-3);t_tab
[1] -3.36493
>
> t_cal=-4/0.4472;t_cal
[1] -8.944544
>
> t_tab<t_cal
[1] FALSE
>
> nd=pt(t_cal,n-3);nd
[1] 0.00
```

O ponto de máximo é dado por:

$$U_M = -\frac{-\beta_1}{2\beta_2} = 2,$$

que é equivalente a

$$-\beta_1 = 4\beta_2,$$

ou

$$\theta = \beta_1 + 4\beta_2 = 0.$$

Vamos estimar θ:

$$\hat{\theta} = b_1 + 4b_2 = 14 - 16 = -2.$$

A variância estimada de $\hat{\theta}$ é

$$S2_e = \hat{V}(b_1) + 16\hat{V}(b_1) + 8C\hat{O}V(b_1, b_2) =$$

Queremos testar:

$$H_0 : \beta_1 + 4\beta_2 = 0 \quad vs \quad H_1 : \beta_1 + 4\beta_2 \neq 0.$$

```
 vcov(mod1)
(Intercept)     U   U2
(Intercept)        0.76 -0.84  0.2
U                 -0.84  1.96 -0.6
U2                 0.20 -0.60  0.2
```

$$S2_e = 1,96 + 16 \times 0,2 - 8 \times 0,6 = 1,96 + 3,2 - 4,8 = 0,36.$$

Logo

$$t_{cal} = \frac{b1 + 4b2}{s_e} = -\frac{2}{0.6} = -\frac{10}{3} = -3,33.$$

Pela tabela temos:

$$P(T(5) \leq 4,032) = 0,995.$$

Rejeitar H_0 : se

$$|t_{cal}| > 4,032.$$

Não podemos rejeitar H_0.

```
alfa=0.01
>
> t_tab=qt(1-alfa/2,n-3);t_tab
[1] 4.032143
>
> t_cal=-2/0.6;t_cal
[1] -3.333333
>
> t_tab^2
[1] 16.25818
> t_cal^2
[1] 11.11111
>
```

2. Resolva a questão matricialmente. Apresente todas as envolvidas bem como sua solução usando o R.

```
 X=cbind(1,U,U2);X
       U U2
[1,] 1 0  0
[2,] 1 0  0
[3,] 1 1  1
[4,] 1 1  1
[5,] 1 2  4
[6,] 1 2  4
[7,] 1 3  9
[8,] 1 3  9
>
> XlX=t(X)%*%X;XlX
      U   U2
    8 12  28
U  12 28  72
U2 28 72 196

> det(XlX)
[1] 640
> IXlX=solve(XlX);IXlX
```

```
              U     U2
    0.475 -0.525  0.125
U  -0.525  1.225 -0.375
U2  0.125 -0.375  0.125

> Y=as.matrix(Y)
> X1Y=t(X)%*%Y;X1Y
    [,1]
    112
U   188
U2  420


> b=IX1X%*%X1Y;b
      [,1]
       7
U     14
U2    -4
>
>
> Jn=matrix(1,n,n);Jn
       [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,]     1    1    1    1    1    1    1    1
[2,]     1    1    1    1    1    1    1    1
[3,]     1    1    1    1    1    1    1    1
[4,]     1    1    1    1    1    1    1    1
[5,]     1    1    1    1    1    1    1    1
[6,]     1    1    1    1    1    1    1    1
[7,]     1    1    1    1    1    1    1    1
[8,]     1    1    1    1    1    1    1    1
>
> A=diag(n)-(1/n)*Jn
>


 SQT=t(Y)%*%A%*%Y;SQT
     [,1]
[1,]  176

>
> H=X%*%IX1X%*%t(X)
>
> SQRES=t(Y)%*%( diag(n)-H)%*%Y;SQRES
       [,1]
[1,]    8
>
>
> SQREG=t(Y)%*%( H-(1/n)*Jn )%*%Y;SQREG
     [,1]
[1,]  168
>
>
```

Vamos terminar matricialmente.