

INSTITUTO SUPERIOR DE AGRONOMIA
ESTATÍSTICA E DELINEAMENTO – 2023-24

3 novembro 2023

Primeiro Teste

Duração: 2h00

Num estudo realizado na variedade de oliveira Cobrançosa, efectuaram-se 152 medições independentes, em oliveiras seleccionadas ao acaso 5 anos após a sua plantação, das seguintes características: rendimento (variável REND, em kg/árvore); teor de gordura na matéria total da azeitona (variável GMT, em %); teor de gordura na matéria seca da azeitona (variável GMS, em %); teor de humidade da azeitona (variável HUM, em %). Eis os seguintes indicadores, bem como a matriz de correlações para todos os pares de variáveis.

Variável	Mínimo	Média	Máximo	Variância
REND	0.0100	0.5062	2.1000	0.09852
GMT	26.14	29.47	33.93	2.2007
GMS	46.56	55.71	70.59	15.6726
HUM	39.02	46.77	63.98	4.9678

```
> cor(oliveira)
      REND      GMT      GMS      HUM
REND 1.0000000 -0.6115455 -0.4835090 0.1063676
GMT -0.6115455 1.0000000 0.6516809 -0.2663570
GMS -0.4835090 0.6516809 1.0000000 0.1804132
HUM 0.1063676 -0.2663570 0.1804132 1.0000000
```

I [9,5 valores]

Foi ajustada a regressão linear múltipla do teor de gordura na matéria total da azeitona (variável GMT) sobre as restantes variáveis observadas. Os resultados foram os seguintes:

```
lm(formula = GMT ~ ., data = oliveira)
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 28.84163    1.75585  16.426 < 2e-16
REND        -1.42688    0.27202  -5.245 5.3e-07
GMS          0.21225    0.02180   9.735 < 2e-16
HUM         -0.22392    0.03409  -6.568 8.1e-10
---
Residual standard error: 0.8949 on 148 degrees of freedom
Multiple R-squared: 0.6434, Adjusted R-squared: 0.6361
F-statistic: 88.99 on 3 and 148 DF, p-value: < 2.2e-16
```

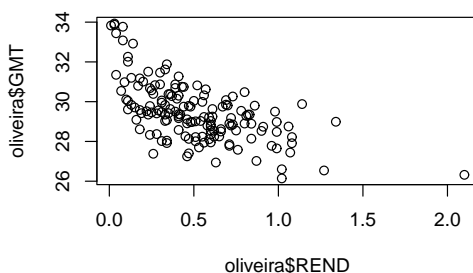
1. Escreva a equação do hiperplano ajustado e interprete o valor do coeficiente de determinação R^2 .
2. Determine o valor das três somas de quadrados associadas a esta regressão linear.
3. Interprete o significado biológico do valor estimado do coeficiente associado à variável preditora rendimento.
4. Diga qual o melhor modelo de regressão linear simples para prever o teor de gordura na matéria total da azeitona. Para a regressão linear simples que escolher determine o declive da recta, bem como o coeficiente de determinação.

5. Dado o Modelo de Regressão Linear:

- (a) indique quais os pressupostos associados aos erros aleatórios e apresente-os usando notação vectorial/matricial;
- (b) prove que a matriz de variâncias-covariâncias do vector dos estimadores dos parâmetros populacionais, $\vec{\beta}$, é dada por $V[\vec{\beta}] = \sigma^2(\mathbf{X}^t \mathbf{X})^{-1}$;
- (c) para o modelo de Regressão Linear Simples, explicita os elementos do vector $\vec{\beta}$ e da matriz $V[\vec{\beta}]$.

II [10,5 valores]

Seguidamente apresenta-se a nuvem de pontos do teor de gordura na matéria total da azeitona (variável GMT) vs. rendimento (variável REND):

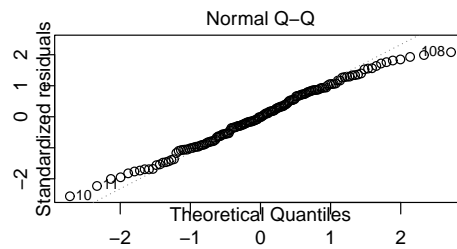
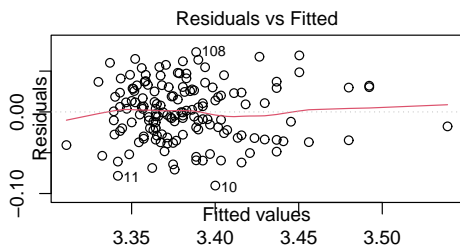


Após a observação deste gráfico, o técnico responsável pelo trabalho em curso decidiu ajustar uma regressão linear usando as transformações logarítmicas de ambas as variáveis. Eis os resultados obtidos.

```
lm(formula = log(GMT) ~ log(REND), data = oliveira)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.342609   0.004338   770.61  <2e-16
log(REND)   -0.042658   0.003511  -12.15  <2e-16
---
Residual standard error: 0.03536 on 150 degrees of freedom
Multiple R-squared:  0.496, Adjusted R-squared:  0.4926
F-statistic: ??? on 1 and 150 DF,  p-value: ???
```

1. Efectue o teste F de ajustamento global e interprete o valor do coeficiente de determinação R^2 .
2. Será admissível afirmar que quando o logaritmo do rendimento aumenta uma unidade, o logaritmo do teor de gordura na matéria total da azeitona diminui, em média, mais de 0.02? Responda através do teste de hipóteses adequado explicando, em pormenor, todos os seus passos.

3. Com 95% de confiança, indique a gama de valores admissíveis para o valor médio do logaritmo do teor de gordura na matéria total da azeitona quando o rendimento é de 1 kg/árvore.
4. Deduza a relação não linear entre o teor de gordura na matéria total da azeitona e o rendimento que corresponde à regressão linear ajustada. Escreva a equação da curva ajustada.
5. Descreva os seguintes gráficos e discuta as suas implicações para o estudo efectuado.



6. Sabendo que para a observação com maior efeito alavanca o valor do teor de gordura na matéria total da azeitona (GMT) é de 26,32%, calcule o resíduo usual dessa observação.