INSTITUTO SUPERIOR DE AGRONOMIA

ESTATÍSTICA E DELINEAMENTO EXPERIMENTAL

29 de outubro, 2025 **PRIMEIRO TESTE 2025-26** **Soluções**

I

1. O modelo de RLM e os seus pressupostos foram múltiplas vezes descritos nas aulas teóricas e práticas, estão descritos nos slides das aulas teóricas e nos apontamentos do modelo linear (páginas 100 e 101). Trata-se, portanto, de adaptar a sua descrição ao caso de estudo apresentado. O modelo em estudo tem 5 parâmetros.

2. Os valores em falta A, B e C:
A - o valor da estimativa do desvio padrão do estimador do parâmetro β_4 é a raíz quadrada do elemento (5,5) da matriz de (co-) variâncias estimadas dos estimadores dos parâmetros do modelo (matriz *vcov* do R, dada no enunciado), ou seja, $\hat{\sigma}_{\hat{\beta}_4} = \sqrt{0.002259} = 0.04753$;
B - o valor calculado da estatística do teste ao parâmetro β_4 ser zero, $t - value = b_4/\hat{\sigma}_{\hat{\beta}_4} = 0.074280/0.047530 = 1.563$;
C - graus de liberdade associados à Soma dos Quadrados dos Resíduos, $n - (p+1) = 349$.

3. Os valores das 3 somas de quadrados:
$SQT = (n-1)\,s_y^2 = 353 \times 0.2889 = 101.9817$;
$SQR = R^2 \times SQT = 72.2846$;
$SQRE = 101.9817 - 72.2846 = 29.6971$.

4. Prova feita nas aulas teóricas e na página 13 dos apontamentos de apoio.

5. Prova feita nas aulas teóricas e na página 110 dos apontamentos de de apoio.

6. A estimativa do coeficiente associado ao preditor Rs, $b_2 = 0.025346$; significa que quando a radiação solar aumenta 1 MJ/m^2/dia, a transpiração das árvores aumenta, em média, 0.025346 mm H_2O/dia, mantendo os valores dos restantes preditores constantes. Com 95% de confiança, os valores possíveis do coeficiente na população variam de 0.018769799 a 0.03192254 (dado no enunciado no comando $confint$ do R).

7. Trata-se de um teste de hipóteses ao parâmetro β_3 :

   Hipóteses: $H_0 : \beta_3 \leq 0.01$ vs. $H_1 : \beta_3 > 0.01$

   Estatística do Teste: $T = \frac{\hat{\beta}_3 - 1}{\hat{\sigma}_{\hat{\beta}_3}} \cap t_{(n-(p+1))}$, sob H_0.

   Nível de significância: $\alpha = 0.05$.

   Região Crítica: (Unilateral direita) Rejeitar H_0 se $T_{\text{calc}} > t_{0.05(349)} \approx 1.646$.

   Conclusões: Tem-se $T_{\text{calc}} = \frac{b_3 - 1}{\hat{\sigma}_{\hat{\beta}_3}} = \frac{0.010751 - 0.01}{0.004529} = 0.1658203$. $T_{\text{calc}} = 0.1658203 < 1.646$, assim, não se rejeita a hipótese nula, ao nível de significância de 0.05, pelo que não é admissível considerar que, quando a temperatura média do ar (`Ta`) aumenta 1 °C, mantendo os restantes preditores constantes, a transpiração das árvores aumenta, em média, mais de 0.01 mm H_2O/dia.

8. Trata-se de um teste de hipóteses a uma combinação linear de parâmetros $\vec{\mathbf{a}}^T\vec{\boldsymbol{\beta}}$, com $\vec{a}^T = [0\ \ 0\ \ 1\ \ -1\ \ 0]$, ficando, $\beta_2 - \beta_3$:

   Hipóteses: $H_0 : \beta_2 - \beta_3 = 0$ vs. $H_1 : \beta_2 - \beta_3 \neq 0$

Estatística do Teste: $T = \frac{(\hat{\beta}_2-\hat{\beta}_3)-0}{\hat{\sigma}_{(\hat{\beta}_2-\hat{\beta}_3)}} \cap t_{(n-(p+1))}$, sob H_0.

Nível de significância: $\alpha = 0.05$.

Região Crítica: (Bilateral) Rejeitar H_0 se $|T_{\text{calc}}| > t_{0.025(349)} \approx 1.96$.

Conclusões: Tem-se $\hat{\sigma}_{(\hat{\beta}_2-\hat{\beta}_3)} = \sqrt{0.000011 + 0.000021 - 2 \times (-0.000005)} = 0.006481$
$T_{\text{calc}} = \frac{(b_2-b_3)-1}{\hat{\sigma}_{\hat{\beta}_2-\hat{\beta}_3}} = \frac{(0.025346-0.010751)-0}{0.006481} = 2.252057$
$|T_{\text{calc}}| = 2.252 > 1.96$, assim, rejeita-se a hipótese nula, ao nível de significância de 0.05, pelo que não é admissível considerar que os coeficientes associados às variáveis preditoras temperatura média do ar (`Ta`) e radiação solar (`Rs`) são iguais.

9. (a) Trata-se de uma combinação linear do tipo $\vec{\mathbf{a}}^T\vec{\boldsymbol{b}}$, com $\vec{a}^T = [1 \;\; 3 \;\; 19 \;\; 16 \;\; 0.9]$, ficando: $\hat{\mu}_{\ Et|PPB=3,\ Rs=19,\ Ta=16,\ DPV=0.9} = 0.119291 + 3\times0.266494 + 19\times0.025346 + 16\times0.010751 + 0.9 \times 0.074280 = 1.639215\ mmH_2O/dia$.

   (b) Trata-se de um intervalo de confiança (95%) para $\mu_{\ Et|PPB=3,\ Rs=19,\ Ta=16,\ DPV=0.9}$. A formulação do intervalo encontra-se no formulário, $\hat{\sigma}_{\vec{\mathbf{a}}^T\hat{\vec{\beta}}} = 0.02115743$ (dado no enunciado), $t_{0.025(349)} \approx 1.96$ e, com base no resultado obtido na alínea anterior, obtém-se: $]1.597746, 1.680684[$.

10. O valor do R^2 modificado está próximo do valor do coeficiente de determinação porque o número de observações é muito elevado face ao número de preditores do modelo. A explicação do significado do R^2 modificado encontra-se no slide 153 e nas páginas 145-146 dos apontamentos de apoio.

11. Ao eliminar o preditor `DPV`, o intervalo de valores possível para o coeficiente de determinação desse submodelo é $[0.6669^2, 0.7088]$.

12. Pode ver-se na matriz de correlações apresentada no enunciado que o maior coeficiente de correlação, em módulo, com transpiração das árvores (variável `Et`) é obtido com a radiação solar (`Rs`), $r_{Et,Rs} = 0.6669$. Assim, o coeficiente de determinação desta regressão linear simples é $R^2 = 0.6669^2 = 0.44476$.

II

1. A equação da recta ajustada: $ln(DPV) = -2.246028 + 0.106755Ta$. O valor obtido para o coeficiente de determinação $R^2 = 0.6134$ significa que 61.34% da variabilidade observada para o logaritmo do défice médio de pressão de vapor do ar é explicada pela regressão sobre o preditor temperatura média do ar.

2. Responde-se à questão através do teste de F de ajustamento global do modelo.

   Hipóteses: $H_0 : \mathcal{R}^2 = 0$ vs. $H_1 : \mathcal{R}^2 > 0$.

   Estatística do teste: $F = \frac{QMR}{QMRE} = \frac{n-(p+1)}{p} \frac{R^2}{1-R^2} \cap F_{(p,n-(p+1))}$, sob H_0.

   Nível de significância: $\alpha = 0.05$.

   Região Crítica (Unilateral direita): Rej. H_0 se $F_{calc} > f_{\alpha\,(p,n-(p+1))} = f_{0.05(1,352)} \approx 3.84$.

   Conclusões: O valor calculado da estatística é $F_{calc} = 558.5$. Logo, rejeita-se a hipótese nula, que corresponde à hipótese dum modelo inútil. Esta conclusão também resulta directamente da análise do valor de prova (*p-value*) associado à estatística de teste calculada: $p-value < 2.2 \times 10^{-16} \approx 0 << \alpha$, o que corresponde a uma rejeição de H_0 aos níveis usuais de α, entre os quais 0.05.

3. Uma vez que foi apenas a variável resposta nesta regressão linear que foi logaritmizada, a relação não linear entre as variáveis originais é uma relação exponencial. Concretamente, e exponenciando a recta ajustada, tem-se:

$$\begin{array}{rrcl} & \ln(DPV) & = & -2.246028 + 0.106755 \cdot Ta \\ \Leftrightarrow & DPV & = & \mathrm{e}^{-2.246028+0.106755 \cdot Ta} \;=\; \mathrm{e}^{-2.246028} \cdot \mathrm{e}^{0.106755 \cdot Ta} \\ \Leftrightarrow & DPV & = & 0.1058187 \cdot \mathrm{e}^{0.106755 \cdot Ta} \end{array}$$

4. Prova feita nas aulas teóricas e na página 68 dos apontamentos de apoio.