

INSTITUTO SUPERIOR DE AGRONOMIA
ESTATÍSTICA E DELINEAMENTO - 2022-23
Primeira Chamada de Exame

6 Janeiro 2023

SOLUÇÕES

I

1. A melhor Regressão Linear Simples para modelar a variável NP2(= y) é obtida utilizando a variável preditora NP pois é a que apresenta melhor correlação (maior coeficiente de correlação em módulo) com NP2, $r = 0.8353155$.
2. (a) $NP2 = 0.84735 + 0.61484 NP$
 $R^2 = r^2 = 0.69775$ indica que cerca de 69.8% da variabilidade observada de NP2 é explicada por esta regressão linear simples.
(b) $SQT = (n - 1)s_y^2 = 15.55591$, $SQR = R^2 \times SQT = 10.85417$ e $SQRE = SQT - SQR = 4.701742$.
(c)]0.53838 , 0.69129[.
3. (a) e (b) Demonstrações feitas nas aulas teóricas. Resolução feita na página 7 das Resoluções de exercícios de Regressão Linear Simples.
4. Resolução feita na página 18 das folhas de apoio, assim como nos slides das aulas teóricas.

II

1. Pelo teste F de ajustamento global, verifica-se que o modelo ajustado difere significativamente do modelo nulo, a qualquer um dos habituais níveis de significância.
2. Gráficos para estudo dos resíduos. Explicação da sua interpretação nas páginas 75 a 81 das folhas de apoio, assim como nos slides das aulas teóricas. Os pressupostos de linearidade do modelo, normalidade e homogeneidade das variâncias dos erros aleatórios não estão a ser violados.
3. Com base no teste t de hipóteses $H_0 : \beta_1 - \beta_4 = 0$ vs. $H_1 : \beta_1 - \beta_4 \neq 0$, com um nível de significância $\alpha = 0.05$, é admissível considerar iguais os valores populacionais dos coeficientes dos preditores NP e NL12, β_1 e β_4 , respectivamente.
4. X é uma matriz de dimensão (112×5) , cujas cinco colunas são um vector de uns e quatro vectores com os valores observados dos preditores,

$$X = \begin{bmatrix} 1 & NP_1 & LS_1 & LI_1 & NL12_1 \\ 1 & NP_2 & LS_2 & LI_2 & NL12_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & NP_{112} & LS_{112} & LI_{112} & NL12_{112} \end{bmatrix}$$

5. Demonstração feita nas aulas teóricas. Resolução feita na página 95 das folhas de apoio.
6. (a) A variável LS foi a primeira a ser excluída.
(b) De acordo com um teste F parcial com um nível de significância de 0.05, o ajustamento deste sub-modelo não difere significativamente do modelo original com 4 preditores, pelo que se deve optar pelo submodelo que é mais simples.

III

1. (a) Delineamento fatorial a 2 fatores, variedade (11 níveis) e ano (3 níveis). Existem 6 repetições em cada uma das 33 situações experimentais pelo que se trata de um delineamento equilibrado com um total de 198 observações. Havendo repetições o modelo mais adequado será o modelo ANOVA a 2 fatores com interação, descrito na página 198 das folhas de apoio, assim como nos slides das aulas teóricas. Trata-se de adaptar este modelo ao problema em estudo.
 - (b) Coluna Df, linha ano: 2;
Coluna Df, linha Residuals: 165;
Coluna Mean Sq, linha variedade: 357.7;
Coluna F Value, linha variedade:ano: 3.285714.
 - (c) Ao nível de significância $\alpha = 0.05$, fazendo três testes F, conclui-se que os três tipos de efeitos previstos pelo modelo (variedade, ano e interação variedade:ano) são significativos.
 - (d) Utilizando o termo de comparação de Tukey (12.37761), conclui-se que, para o nível de significância $\alpha = 0.05$, o peso médio do bolbo observado na variedade VL no ano 2009 difere significativamente do peso médio do bolbo observado nesse ano para as restantes variedades.
2. $H_{calc} = 15.493$, logo rejeita-se H_0 para $\alpha = 0.05$. Há pelo menos uma variedade que tem um número de dentes por bolbo superior às restantes variedades.