

INSTITUTO SUPERIOR DE AGRONOMIA
ESTATÍSTICA E DELINEAMENTO
ESTATÍSTICA APLICADA AO AMBIENTE
SEGUNDO TESTE – 2011/12

11 de Janeiro, 2012

Duração: 2h00

I

Pretende-se estudar os pesos médios de cinco diferentes variedades de tomate. Havendo quatro terrenos disponíveis, em cada terreno demarcaram-se cinco parcelas, sendo estas cinco parcelas associadas aleatoriamente a cada uma das cinco variedades estudadas. A variável resposta considerada foi o peso médio por planta (em g) dos frutos em cada parcela, num total de $n = 20$ observações. Para cada variedade, e para a totalidade das observações, obtiveram-se os seguintes valores médios e variâncias:

Variedade	40B	40C	40D	41A	ACE	Global
Média	131.80	110.70	189.40	227.85	244.85	180.9205
Variância	520.2600	538.2667	1487.2867	788.5967	739.7967	3529.416

1. Foi inicialmente ajustado um modelo ANOVA a um factor (variedade), ignorando eventuais diferenças entre terrenos.
 - (a) Construa a tabela-resumo da ANOVA associada a este modelo.
 - (b) Admitindo a validade do modelo, é lícito afirmar (para um nível de significância $\alpha = 0.05$) que há rendimentos médios diferentes em diferentes variedades? Justifique formalmente.
 - (c) Independentemente da sua resposta na alínea anterior, utilize o teste de Tukey para identificar variedades cujas médias se devem considerar diferentes das da variedade 40D, a um nível de significância $\alpha = 0.01$.
 - (d) Qual o *efeito* estimado da variedade 40C? Como interpreta esse valor?
2. O modelo ajustado na alínea anterior foi criticado por não prever eventuais efeitos de terrenos.
 - (a) Descreva em pormenor um modelo ANOVA adequado para ultrapassar esta crítica, tendo em conta a descrição da experiência.
 - (b) Construa a tabela-resumo correspondente ao novo modelo, sabendo que o Quadrado Médio Residual é agora 941.
 - (c) Teste formalmente se o ajustamento deste novo modelo difere significativamente do modelo a um factor ajustado na alínea anterior. Comente a sua conclusão, dado o contexto do problema.

II

A casta Arinto é uma das variedades brancas de videira mais cultivada em Portugal. Para estudar o comportamento de 5 clones desta casta (AR36, AR37, AR38, AR39, AR40) relativamente ao rendimento ($kg/planta$) nas suas principais regiões de cultura, foram instalados ensaios em 5 locais: Setúbal, Bucelas, Alenquer, Mealhada e Felgueiras. Em cada local definiram-se 20 parcelas de terreno e cada clone foi disposto aleatoriamente em 4 dessas parcelas.

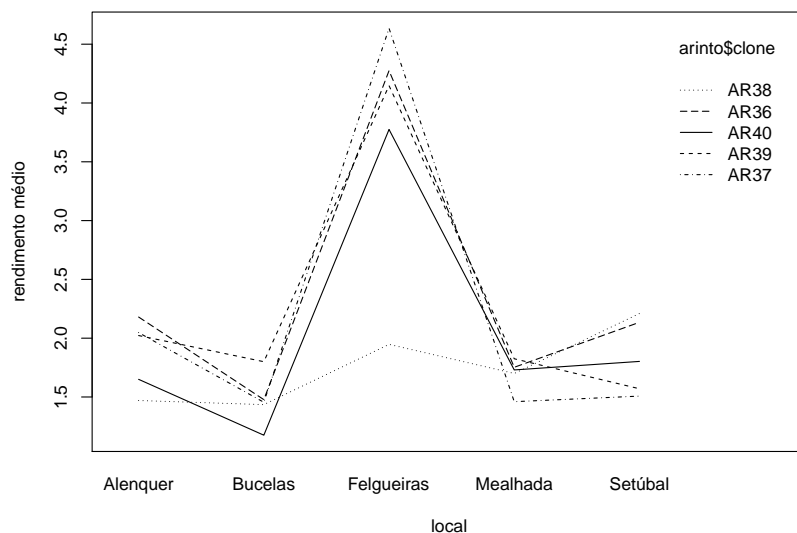
Os rendimentos médios de cada clone em cada local foram os seguintes:

	local				
clone	Alenquer	Bucelas	Felgueiras	Mealhada	Setúbal
AR36	2.180	1.478	4.275	1.753	2.135
AR37	2.047	1.455	4.635	1.460	1.508
AR38	1.470	1.435	1.947	1.700	2.210
AR39	2.023	1.800	4.147	1.825	1.570
AR40	1.650	1.175	3.775	1.730	1.803

1. Foi ajustado um modelo ANOVA que produziu a seguinte tabela-resumo:

	Df	Sum Sq	Mean Sq	F value
clone	??	4.733	1.1834	???
local	??	68.374	17.0936	76.4016
clone:local	??	17.300	???	4.8326
Residuals	??	16.780	0.2237	

- Descreva em pormenor o modelo ANOVA que foi ajustado.
 - Complete a tabela-resumo, indicando os seis valores omissos.
 - Diga qual o valor ajustado pelo modelo para uma observação do clone AR39 em Setúbal.
 - Calcule a variância amostral da totalidade das observações.
2. Quais os tipos de efeitos que devem ser considerados significativos, ao nível de significância $\alpha = 0.05$? Justifique pormenorizadamente num caso, e de forma mais sintética nos restantes.
3. Comente o seguinte gráfico:



4. É admissível afirmar que “com exceção do clone AR38, os rendimentos médios em Felgueiras são significativamente superiores aos de outras localidades”? Justifique a sua resposta com base no teste de Tukey.

NOTA: Na leitura da tabela, utilize o valor disponível mais próximo.

III

1. Considere um modelo ANOVA a 1 factor (k níveis), enquanto caso particular do modelo linear.
 - (a) Descreva a matriz do modelo, \mathbf{X} .
 - (b) Caracterize os vectores do subespaço $\mathcal{C}(\mathbf{X})$, gerado pelas colunas da matriz \mathbf{X} .
 - (c) Mostre que, caso o delineamento seja equilibrado, com n_c observações em cada nível do factor, a estatística do teste de ajustamento global do modelo é dada por n_c vezes a razão entre a variância das médias amostrais de cada nível (ou seja, a variância das médias $\bar{Y}_{i.}$) e a média das variâncias amostrais de cada nível (a média dos S_i^2).
 - (d) Com base na expressão referida no enunciado da alínea anterior, justifique a opção por uma região de rejeição unilateral direita, no teste aos efeitos do factor.

2. Considere um delineamento hierarquizado equilibrado (n_c repetições) com dois factores: A (dominante e com a níveis), e B (subordinado), sempre com b níveis (para qualquer nível do Factor A).
 - (a) Descreva o modelo ANOVA correspondente ao delineamento indicado.
 - (b) Diga qual a decomposição da Soma de Quadrados Total correspondente ao modelo que indicou na alínea anterior, especificando como se obtém cada parcela dessa decomposição.