

Exercícios - Modelos Matemáticos e Aplicações - 2022-23

3 Análise de Variância de Efeitos Fixos

AVISO: Os conjuntos de dados de vários exercícios estão disponíveis na página *web* da disciplina.

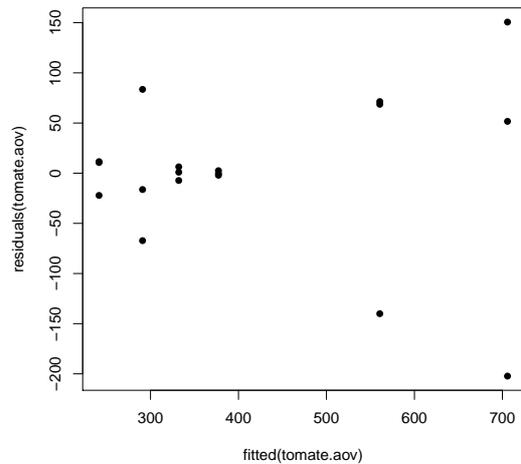
EXERCÍCIOS

1. No melhoramento de variedades tradicionais de tomate, uma característica importante é a resistência da película. Esta característica foi avaliada em 6 variedades de tomate, num ensaio envolvendo 18 parcelas de tomateiros. Concretamente, cada variedade foi associada aleatoriamente a 3 parcelas de tomateiros. Cada observação individual é constituída pela resistência média dos frutos de cada parcela (medida num texturómetro, em grama força, *gf*). Os dados estão no ficheiro `tomateresistencia.txt`, disponível na página *web* da disciplina. Eis os valores obtidos em cada parcela, bem como as médias e variâncias para cada variedade e para a totalidade das observações:

Variedade	Observações			Média	Variância
18	632.04	629.30	420.59	560.6433	14 713.08
28	253.00	219.34	252.11	241.4833	367.9434
29	223.71	374.48	274.66	290.9500	5881.921
40C	503.51	757.44	856.39	705.7800	33 132.64
Ace	375.18	376.81	379.77	377.2533	5.414433
Roma	333.05	324.82	338.45	332.1067	47.11163

- A média global das observações é $\bar{y}_{..} = 418.0361$;
- a variância amostral da totalidade das observações é $s_y^2 = 34 517.82$.

- Indique qual o tipo de delineamento experimental em causa. Explícite o modelo correspondente e todas as hipóteses adicionais que sejam necessárias à consideração do problema em estudo.
- Construa a tabela-resumo da análise de variância correspondente a este caso,
 - utilizando apenas uma máquina de calcular e a informação disponível neste enunciado;
 - utilizando, no R, o comando `summary(aov(res.pel ~ variedade , data=tomateresistencia))`.
- Formalize e efectue o teste F adequado ao problema acima referido, com um nível de significância de 5%. Pode afirmar-se que a resistência média da película não é sempre igual, em todas as variedades?
- Qual o maior nível de significância α para o qual mudaria a sua resposta na alínea anterior? Como se designa esse valor?
- Utilize o comando `model.matrix` do R para inspeccionar a natureza da matriz do modelo, \mathbf{X} , neste contexto.
- Utilize o comando `fitted` do R para identificar os valores ajustados da variável resposta, nesta Análise de Variância.
- Quais as variedades cuja resistência difere significativamente da resistência da variedade 40C? Justifique com um teste adequado, ao nível $\alpha = 0.05$.
No final, execute o comando `library(agricolae)` do R e utilize os comandos:
`tomate.aov<-aov(res.pel ~ variedade , data=tomate)`
`HSD.test(tomate.aov, "variedade", console=TRUE)`.
- O gráfico dos resíduos (usuais) das observações, contra os valores ajustados pelo modelo de análise de variância, é apresentado a seguir. Comente o gráfico e as suas possíveis implicações. Identifique a observação cujo resíduo é, em módulo, mais elevado.



2. Em 2015 foi efectuado um ensaio em Vila Real envolvendo 6 genótipos da casta Aragonez, também conhecida no Douro por Tinta Roriz, num trabalho do ISA e da PORVID (Associação Portuguesa para a Diversidade da Videira). Pretendeu-se medir o chamado Índice de Fertilidade Potencial (variável `if`, definida como o número de cachos por número de olhos abrolhados) e saber se este difere consoante os genótipos estudados. Uma vez que se sabe que a fertilidade está também dependente do tipo de solo, foram escolhidos para o ensaio seis terrenos com solos diferentes. Cada terreno foi dividido em seis parcelas, sendo cada uma dessas parcelas associada, aleatoriamente, a um diferente genótipo. Os dados estão no ficheiro `tintarorizif.txt`, disponível na página *web* da disciplina. Foi ajustado um modelo ANOVA com os seguintes resultados:

	Df	Sum Sq	Mean Sq	F value
genotipo	??	??	??	4.204
terreno	??	??	??	2.691
Residuals	??	??	0.3660	

- Indique o tipo de delineamento experimental utilizado, descrevendo em pormenor o modelo ANOVA correspondente à tabela-resumo acima indicada.
 - Complete a tabela-resumo, indicando como obtém os valores em falta.
 - Diga, justificando formalmente, que tipo de efeitos são significativos, ao nível $\alpha = 0.01$.
 - Um analista que apenas conhecia as ANOVAs a um factor, decidiu estudar os efeitos de genótipo, com os mesmos dados. Construa a tabela-resumo resultante e indique a que conclusão chegaria (ao mesmo nível $\alpha = 0.01$). Comente.
 - Outro analista afirma que há genótipos que se dão particularmente bem em determinados terrenos estudados, e sugere que essa possibilidade deve ser tida em conta pelo modelo. Com base na experiência realizada é possível prever essa eventualidade?
3. Num estudo sobre características de crescimento de pinheiro manso, conduzido em Sines e em Tavira pelo Instituto Nacional de Investigação Agrária e Veterinária (INIAV), avaliou-se a altura média de pinheiros de cinco diferentes proveniências (Marrocos, Grécia, Portugal e duas proveniências de Itália), dois anos após a plantação. Quer em Sines, quer em Tavira, foram plantados seis talhões com árvores de cada proveniência, gerando assim $n = 60$ valores de alturas (variável `alt2`, em cm), cuja variância amostral é $s^2 = 34.49584$. Eis algumas médias resultantes.

prov

local

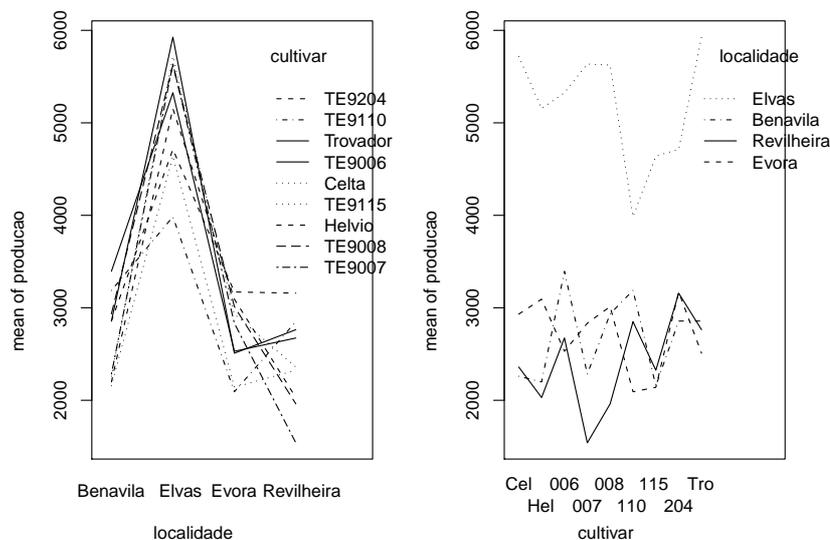
Grecia	Italia-1	Italia-2	Marrocos	Portugal	Sines	Tavira
28.81	32.75	30.23	35.13	31.90	28.14	35.38

prov:local			
	local		Grand mean
prov	Sines	Tavira	31.76298
Grecia	22.52	35.10	
Italia-1	31.03	34.46	
Italia-2	26.91	33.56	
Marrocos	31.16	39.09	
Portugal	29.09	34.70	

- (a) Identifique o delineamento experimental utilizado e o modelo ANOVA adequado. Descreva pormenorizadamente o modelo.
- (b) Sabendo que o Quadrado Médio Residual é 16.59 e que a Soma de Quadrados associada às cinco diferentes proveniências é 280.61, construa a tabela-resumo do modelo ANOVA adequado.
- (c) Use um teste F para avaliar a existência de efeitos de proveniência dos pinheiros. Comente as suas conclusões. Indique brevemente que outros tipos de efeitos devem ser considerados significativos. Considere $\alpha = 0.05$.
- (d) Na amostra, a maior altura média em Sines é inferior à menor altura média em Tavira. Independentemente das suas respostas nas alíneas anteriores, use o teste de Tukey para indicar se igual afirmação se pode estender à população. Comente.
4. Uma engenheira agrónoma pretende seleccionar cultivares de trigo para as quatro explorações agrícolas pelas quais é responsável, que se localizam em Elvas, Évora, Benavila e Revilheira. Em cada uma destas explorações, definem-se 36 parcelas de terra, associando aleatoriamente quatro parcelas a cada uma de nove cultivares: Celta, Helvio, TE9006, TE9007, TE9008, TE9110, TE9115, TE9204 e Trovador. Em cada parcela foi medido o rendimento, em kg/ha. A variância da totalidade dos rendimentos observados é $s^2 = 1714242$.
- (a) Especifique o delineamento experimental utilizado, e descreva em pormenor o modelo ANOVA adequado a esta experiência.
- (b) Foi ajustado um modelo ANOVA, com o programa R. Resultados parciais desse ajustamento são dados de seguida.

	Df	Sum Sq	Mean Sq	F value
localidade	???	183759916	???	234.9531
cultivar	???	???	964060	???
localidade:cultivar	???	???	???	4.0768
Residuals	???	28156076	260704	

- i. Complete a tabela, indicando como obtém cada um dos valores omissos.
- ii. Qual o valor estimado da variância dos erros aleatórios do modelo, e quais as suas unidades de medida?
- iii. Teste formalmente (a um nível de significância $\alpha = 0.01$) quais os tipos de efeitos do modelo que devem ser considerados significativos. Descreva um teste em pormenor e discuta os restantes de forma sintética.
- iv. Os gráficos de interação associados a esta experiência são os seguintes. Comente-os, relacionando as suas conclusões das alíneas anteriores com os gráficos.



- v. Discuta o efeito de mudar as unidades de medida da variável resposta de kg/ha para toneladas por hectare. Quais os valores da tabela que se alteram, e quais os que ficam iguais? Quais os efeitos da mudança de unidades nas conclusões dos testes F ?
- (c) Os rendimentos médios das quatro parcelas correspondentes a cada combinação de localidade e cultivar foram os seguintes.

localidade	cultivar								
	Celta	Helvio	TE9006	TE9007	TE9008	TE9110	TE9115	TE9204	Trovador
Benavila	2260	2198	3395	2283	2935	3188	2157	2857	2857
Elvas	5719	5154	5326	5636	5624	3984	4640	4713	5927
Evora	2932	3094	2530	2833	3013	2092	2140	3171	2509
Revilheira	2360	2030	2673	1540	1962	2850	2325	3158	2763

Utilizando o teste de Tukey, indique quais as cultivares que, em Elvas, não diferem significativamente do melhor rendimento observado nessa localidade ($\alpha = 0.10$). Comente.

5. Num estudo sobre avaliação da interação genótipoxambiente, estudaram-se cinco clones da variedade de videira Arinto quanto ao rendimento ($kg/planta$), em cinco locais diferentes de Portugal (Setúbal, Bucelas, Alenquer, Mealhada e Felgueiras). Em cada local, foram escolhidas 20 parcelas de terreno e, aleatoriamente, cada clone foi associado a quatro dessas 20 parcelas. Os dados estão no ficheiro `arinto.txt`, disponível na página *web* da disciplina.
- (a) Especifique o delineamento experimental utilizado e descreva em pormenor o modelo ANOVA adequado a esta experiência.
- (b) Ajuste o modelo descrito na alínea a) com o programa R, executando o comando
- ```
arinto.aov<-aov(yield~clone*site, data=arinto)
```
- (c) Teste formalmente (a um nível de significância  $\alpha = 0.05$ ) quais os tipos de efeitos do modelo que devem ser considerados significativos. Descreva um teste em pormenor e discuta os restantes de forma sintética.
- (d) Obtenha e comente os gráficos de interação associados a esta experiência. Execute os seguintes comandos:

```
interaction.plot(response=arinto$yield, x.factor=arinto$clone, trace.factor=arinto$site)
interaction.plot(response=arinto$yield, trace.factor=arinto$clone, x.factor=arinto$site)
```

- (e) Identifique o clone com maior média amostral de rendimento em Felgueiras. Execute o comando:

```
model.tables(arinto.aov, type="mean")
```

Utilizando o teste de Tukey, indique quais os clones que, em Felgueiras, diferem significativamente do melhor rendimento observado nessa localidade ( $\alpha = 0.05$ ). No fim, execute o comando:

```
TukeyHSD(arinto.aov)
```

também pode usar a função `HSD.test()` do package `agricolae`:

```
library(agricolae)
```

```
HSD.test(arinto.aov, c("clone","site"), console=TRUE)
```