

MODELOS MATEMÁTICOS E APLICAÇÕES 2015/2016

EXERCÍCIOS DE MODELOS LINEARES MISTOS – PARTE I

1. Num estudo de selecção da videira, com vista a estudar a variabilidade genética do rendimento da casta Touriga Nacional, foi instalado um ensaio de campo em Vila Nova de Fozcoa, com uma amostra aleatória de genótipos da casta (196 genótipos). No campo, cada genótipo foi disposto aleatoriamente em 5 parcelas (ensaio com 5 repetições). Em 1994 foram obtidos os dados de rendimento (kg/planta) que se encontram disponíveis na *data.frame* *touriga*.
 - a) Descreva o modelo que lhe parece ser adequado para estudar a variabilidade genética do rendimento da casta.
 - b) Utilize o comando *aov* do R para obter a tabela de análise de variância do respectivo modelo.
 - (i) Estime as componentes de variância envolvidas no modelo descrito na alínea a).
 - (ii) Com base na informação disponível, efectue um teste à existência de variabilidade genética do rendimento da casta (utilize o nível de significância de 0.05).
 - (iii) Sabendo que $\bar{Y}_{..} = 1.196$ kg/planta e $\bar{Y}_{c0101.} = 1.6044$ kg/planta, qual o melhor preditor empírico linear não enviesado do efeito genotípico do rendimento do clone c0101?
 - c) Ajuste o modelo anteriormente descrito com o método de estimação de máxima verosimilhança restrita, recorrendo
 - (i) ao comando *lme* do pacote “nlme”; ao comando *lmer* do pacote “lme4”; ao comando *varComp* do pacote “varComp”.
 - (ii) Aplique o comando *summary* aos três objectos criados anteriormente e identifique as estimativas de máxima verosimilhança restrita das componentes de variância. Compare os resultados agora obtidos com os da alínea b(i).
 - (iii) Efectue um teste à existência de variabilidade genética do rendimento (teste de razão de verosimilhanças).
 - (iv) Qual o rendimento previsto do clone c0101 na repetição 2?
 - (v) Explore os comandos *ranef* e *fitted* do “nlme” e “lme4”.
 - d) Na verdade, o ensaio de Touriga Nacional atrás descrito foi instalado de acordo com um delineamento experimental em blocos casualizados completos (5 blocos).
 - (i) Ajuste um novo modelo que incorpore o efeito do bloco (considere também que é um factor de efeitos aleatórios). Utilize o pacote *lme4*.
 - (ii) Efectue testes de hipóteses às componentes de variância associadas aos factores clone e bloco.

(iii) Calcule o AIC e o BIC dos dois modelos ajustados (o da alínea a) e o da alínea d)) e seleccione o melhor modelo com base nesses critérios.

2. Considere os dados disponíveis na *data.frame* *terrenos*. Nesse estudo pretende-se comparar o rendimento obtido com quatro variedades de trigo. Identificaram-se 13 terrenos com características de solos diferentes. Assuma que esses terrenos são uma amostra aleatória da população de terrenos onde essas variedades serão cultivadas.

a) Descreva o modelo apropriado a este novo contexto.

b) Complete a tabela de análise de variância para o caso em estudo, diga quais são as estimativas das componentes de variância e teste se existe variabilidade entre terrenos (utilize o nível de significância de 0.05).

	g.l.	SQ	QM	Fcalc
variedade	?	1.799	?	?
terreno	?	2.407	?	?
Resíduos	36	3.513	0.0976	

c) Ajuste esse modelo no R recorrendo ao comando *lme* do pacote “nlme”, utilizando os dois métodos de estimação: máxima verosimilhança e máxima verosimilhança restrita. Compare as estimativas obtidas para as componentes de variância e comente.

d) Execute o comando *plot.design* (terrenos).

e) Considere o resultado referente ao ajustamento do modelo utilizando o método REML.

```
> terrenolme1<-lme(rend~variedade, random=~1|terreno, data=terrenos)
```

```
> summary(terrenolme1)
```

Linear mixed-effects model fit by REML

Data: terrenos

AIC BIC logLik

55.42708 66.65429 -21.71354

Random effects:

Formula: ~1 | terreno

(Intercept) Residual

StdDev: 0.1604919 0.3123811

Fixed effects: rend ~ variedade

	Value	Std.Error	DF	t-value	p-value
(Intercept)	1.5560000	0.09740463	36	15.974600	0.0000
variedadeB	-0.0238462	0.12252595	36	-0.194621	0.8468
variedadeC	-0.3890769	0.12252595	36	-3.175465	0.0031
variedadeD	-0.3778462	0.12252595	36	-3.083805	0.0039

```
> vcov(terrenolme1)
```

(Intercept) variedadeB variedadeC variedadeD

(Intercept) 0.009487662 -0.007506305 -0.007506305 -0.007506305

variedadeB -0.007506305 0.015012610 0.007506305 0.007506305

variedadeC -0.007506305 0.007506305 0.015012610 0.007506305

variedadeD -0.007506305 0.007506305 0.007506305 0.015012610

Recorde os testes de hipóteses a combinações lineares dos efeitos fixos do modelo linear misto dados nas aulas teóricas. Teste os efeitos fixos do modelo (efeitos associados às variedades). Para tal, descreva as hipóteses, a estatística do teste, considere a matriz (assimptótica) das covariâncias estimadas dos estimadores dos efeitos fixos (comando `vcov(terrenolme1)`), defina a matriz L, crie o vector com as estimativas dos efeitos fixos e, com a ajuda do R, calcule a estatística do teste. Para as suas conclusões utilize o nível de significância de 0.05. No final, execute o comando `anova(terrenolme1)`.

f) Existe acréscimo de rendimento da variedade B comparativamente à variedade A (para $\alpha = 0.05$)?

g) No R, com o função `varComp` do pacote “varComp”, fez-se a seguinte sequência de comandos:

```
> terrenosvarcomp1<-varComp(rend~variedade, random=~terreno, data=terrenos)
> terrenosvarcomp1
varComp(fixed = rend ~ variedade, data = terrenos, random = ~terreno)
Fixed effect estimates:
(Intercept)  variedadeB  variedadeC  variedadeD
1.55600000 -0.02384615 -0.38907692 -0.37784615
```

Variance component estimates:

```
terreno  error
0.02575779 0.09758188
Number of observations: 52
```

```
> logLik(terrenosvarcomp1)
'log Lik.' -202.4013 (df=2)
```

```
> terrenosvarcomp0<-varComp(rend~variedade, data=terrenos)
> logLik(terrenosvarcomp0)
'log Lik.' -203.6992 (df=1)
```

Interprete os comandos anteriores e efectue um teste de razão de verosimilhanças para averiguar se existe variabilidade entre terrenos.

h) Execute os comandos e comente os resultados.

```
> terrenos.lme1<-lme(rend~variedade, random=~1|terreno, data=terrenos)
> plot(terrenos.lme1)
> qqnorm(terrenos.lme1, ~resid(.))
> qqnorm(terrenos.lme1, ~ranef(.))
```

3. Numa fase final de selecção de clones de videira da casta Antão Vaz, avaliaram-se 6 clones quanto ao rendimento (kg/planta) em 11 ambientes distintos, tentando que estes últimos fossem representativos da gama de ambientes onde os clones irão ser cultivados. O objectivo é avaliar se existe sensibilidade desses clones à interacção genótipoXambiente. Para tal, em cada ambiente foi instalado um ensaio em que os seis clones foram plantados de acordo com um delineamento experimental totalmente casualizado com 8 repetições. Os dados estão disponíveis na *dataframe* *antaovaz*.

a) Descreva o modelo que lhe parece adequado para o caso em estudo.

b) Complete a tabela de análise de variância para o caso de estudo. Faça os testes de hipóteses adequados para avaliar se existe variabilidade associada à interacção genótipo x ambiente, se existe variabilidade entre ambientes e se existem efeitos principais associados aos clones. Utilize o nível de significância de 0.05.

	GL	SQ	QM	Fcalc
clone	???	???	1.54	???
ambiente	???	1697.2	???	???
clone:ambiente	???	116.2	2.32	???
Resíduos	462	1279.2	???	

d) Ajuste o modelo anteriormente descrito no R com o pacote *lme4*. Proceda de forma adequada para avaliar se existe variabilidade da interacção, se existe variabilidade entre ambientes e se existem efeitos principais associados aos clones.

4. Num estudo sobre variedades tradicionais de tomate, foi avaliado o peso dos frutos (g/fruto) em 41 variedades num ensaio instalado de acordo com um delineamento experimental em blocos casualizados completos (4 blocos). O investigador responsável optou por assumir os efeitos dos blocos e das variedades como aleatórios. O grande objectivo deste estudo é saber se as variedades tradicionais de tomate apresentam variabilidade quanto ao peso do fruto. Os resultados obtidos no R foram os descritos seguidamente:

```
>tomate<-read.table("D:\\ELSA-T\\Aulas\\Modelosmatemáticos-
UCdoutoramento\\dados\\tomate.txt", header=T)
> tomate.lmer1<-lmer(pesofruto~1+(1|variedade)+(1|bloco), data=tomate)
> tomate.lmer1
Linear mixed model fit by REML ['lmerMod']
Formula: pesofruto ~ 1 + (1 | variedade) + (1 | bloco)
Data: tomate
REML criterion at convergence: 1852.448
Random effects:
```

<i>Groups</i>	<i>Name</i>	<i>Std.Dev.</i>
	<i>variedade (Intercept)</i>	72.988
	<i>bloco (Intercept)</i>	9.628
	<i>Residual</i>	53.398

Number of obs: 164, groups: variedade, 41; bloco, 4

Fixed Effects:
(Intercept)
 214.9

```

> logLik(tomate.lmer1)
'log Lik.' -926.2243
> tomate.lmer2<-lmer(pesofruto~1+(1|bloco), data=tomate)
> logLik(tomate.lmer2)
'log Lik.' -967.8129.

```

- a) Descreva o modelo adequado para o caso em estudo.
- b) Efectue o teste de hipóteses adequado para responder ao objectivo do estudo.