

MODELOS MATEMÁTICOS E APLICAÇÕES 2016/2017

Módulo 2: EXERCÍCIOS DE MODELOS LINEARES MISTOS

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:
 - (i) recorrendo 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. 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:
Groups      Name      Std.Dev.
variedade (Intercept) 72.988
bloco      (Intercept)  9.628
Residual                    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.

3. 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

i) 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*).

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

f) 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.

g) 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(.))
```

4. Na *library nlme e lme4* do R está disponível a data.frame *Machines*. Estes dados referem-se a uma experiência referente à comparação de três marcas de máquinas utilizadas num processo industrial. Seis trabalhadores foram escolhidos aleatoriamente entre os funcionários de uma fábrica para operar cada máquina três vezes. A variável resposta é uma pontuação global da produtividade tendo em conta o número e a qualidade das componentes produzidas. No R, com a função *lmer* do pacote *lme4*, executaram-se os seguintes comandos:

```
machines1r<-lmer(score~Machine+(1|Worker)+(1|Worker:Machine), data=Machines)
summary(machines1r)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: score ~ Machine + (1 | Worker) + (1 | Worker:Machine)
```

Random effects:

Groups	Name	Variance	Std.Dev.
Worker:Machine	(Intercept)	13.9095	3.7295
Worker	(Intercept)	22.8585	4.7811
Residual		0.9246	0.9616

Number of obs: 54, groups: Worker:Machine, 18; Worker, 6

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	52.356	2.486	21.062
MachineB	7.967	2.177	3.660
MachineC	13.917	2.177	6.393

```
> anova(machines1r)
```

Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value
Machine	2	38.051	19.025	20.576

```
machines2r<-lmer(score~Machine+(1|Worker), data=Machines)
```

```
summary(machines2r)
```

```
Linear mixed model fit by REML ['lmerMod']
```

```
Formula: score ~ Machine + (1 | Worker)
```

Random effects:

Groups	Name	Variance	Std.Dev.
Worker	(Intercept)	26.487	5.147
Residual		9.996	3.162

Number of obs: 54, groups: Worker, 6

Fixed effects:

	Estimate	Std. Error	t value
--	----------	------------	---------

```

(Intercept)  52.356  2.229 23.485
MachineB     7.967   1.054 7.559
MachineC    13.917   1.054 13.205
lr1<-logLik(machines1r)
lr1
'log Lik.' -107.8438 (df=6)
lr0<-logLik(machines2r)
lr0
'log Lik.' -143.4391 (df=5)
machines3r<-lmer(score~Machine+(1|Worker:Machine), data=Machines)
summary(machines3r)
Linear mixed model fit by REML ['lmerMod']
Formula: score ~ Machine + (1 | Worker:Machine)
Random effects:
Groups      Name          Variance Std.Dev.
Worker:Machine (Intercept) 36.7679  6.0637
Residual                0.9246  0.9616
Number of obs: 54, groups: Worker:Machine, 18
Fixed effects:
              Estimate Std. Error t value
(Intercept)  52.356    2.486 21.062
MachineB     7.967    3.515  2.266
MachineC    13.917    3.515  3.959
> lr30<-logLik(machines3r)
> lr30
'log Lik.' -110.6278 (df=5)

```

- a) Descreva, em pormenor, o modelo que lhe parece ser adequado para o estudo acima descrito.
 - b) Quais os valores das estimativas de máxima verosimilhança restrita das componentes de variância do modelo?
 - c) Os valores das estimativas das componentes de variância obtidas pelo método de máxima verosimilhança seriam superiores ou inferiores aos valores das estimativas enunciados na alínea anterior?
 - d) Efectue o teste de hipóteses adequado para averiguar se existe variabilidade relativamente à interacção trabalhadorXmáquina? Enuncie detalhadamente o teste de hipóteses que está a efectuar. Utilize um nível de significância de 0,01.
 - e) Efectue o teste de hipóteses adequado para averiguar se existe variabilidade relativamente ao trabalhador? Enuncie detalhadamente o teste de hipóteses que está a efectuar. Utilize um nível de significância de 0,01.
 - f) Quais os valores das estimativas dos efeitos fixos do modelo? Explique o seu significado.
 - g) Efectue um teste de hipóteses adequado para avaliar se existem efeitos principais associados às marcas de máquinas. Enuncie detalhadamente o teste de hipóteses que está a efectuar. Utilize um nível de significância de 0,01.
- 5.** 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*.

- Descreva o modelo que lhe parece adequado para o caso em estudo.
- 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 interação 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	???	

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

6. No *package* “*nlme*” do R, existe um conjunto de dados (designado por “*Alfalfa*”).

```
> head(Alfalfa)
Grouped Data: Yield ~ Date | Block/Variety
  Variety Date Block Yield
1 Ladak None  1 2.17
2 Ladak S1  1 1.58
3 Ladak S20  1 2.29
4 Ladak O7  1 2.23
5 Ladak None  2 1.88
6 Ladak S1  2 1.26
...
```

Estes dados foram descritos por Snedecor & Cochran (1980) como exemplo de um delineamento em parcelas divididas (*split-plot design*). Pretende-se estudar se o rendimento (T/acre) da luzerna (*Medicago sativa*) é afectado pela variedade e pela data do terceiro corte. Existem, portanto, dois factores em estudo: factor variedade de luzerna, com 3 níveis (*Cossac*, *Ladak* e *Ranger*) e factor data do corte, com 4 níveis (*none*–sem corte, *S1*– Sep1; *S20* – Sep20; and *O7* – Oct7). As unidades experimentais foram organizadas em 6 blocos, cada bloco com 3 grandes parcelas (*whole plots*), às quais foram aleatoriamente atribuídas as 3 variedades, e cada grande parcela foi dividida em 4 pequenas parcelas (*split plots* ou *subplots*), às quais foram aleatoriamente atribuídas as datas do corte.

- Descreva o modelo que lhe parece adequado para responder ao objectivo do estudo.
- Execute o comando *plot.design* (*Alfalfa*).
- Ajuste o modelo descrito no R recorrendo ao comando *lme* do pacote “*nlme*”.
- Efectue os testes de hipóteses que acha adequados para responder aos objectivos do estudo.
- Compare os resultados obtidos nas alíneas anteriores com o resultados obtidos com o comando “*aov*(*Yield~Date*Variety+Error(Block*Variety)*, *data=Alfalfa*)”.

7. No *package* “*nlme*” do R, existe um conjunto de dados (designado por “*Oats*”).

```
> head(Oats)
Grouped Data: yield ~ nitro | Block
Block  Variety nitro yield
```

```

1 I Victory 0.0 111
2 I Victory 0.2 130
3 I Victory 0.4 157
4 I Victory 0.6 174
5 I Golden Rain 0.0 117
6 I Golden Rain 0.2 114
...

```

Estes dados foram introduzidos por Yates (1935) como o exemplo de um delineamento em parcelas divididas. Pretende-se estudar se o rendimento (bushels/acre) da aveia é afectado pela variedade e concentrações de nitrogénio. Existem, portanto, dois factores em estudo: factor variedade de aveia, com 3 níveis (*Golden Rain*, *Marvellous* e *Victory*) e factor concentração de nitrogénio (cwt/acre), com 4 níveis (0.0, 0.2, 0.4 e 0.6). As unidades experimentais foram organizadas em 6 blocos, cada bloco com 3 grandes parcelas, cada uma dividida em 4 parcelas. As variedades de aveia foram aleatoriamente atribuídas às grandes parcelas e as concentrações de nitrogénio foram aleatoriamente atribuídas às parcelas divididas. Admite-se variedade e concentração de nitrogénio como factores de efeitos fixos e o bloco como factor de efeitos aleatórios.

Considere a concentração de nitrogénio como um factor ((no R utilize *factor(nitro)* na descrição do modelo). Ajuste o modelo que lhe parecer adequado no R, por exemplo, com a função *lmer*. O que conclui sobre a influência da variedade e da concentração de nitrogénio sobre o rendimento?

8. No R está disponível o conjunto de dados *BodyWeight* (Pinheiro e Bates, 2000), referente ao peso do corpo de ratos (em gramas), ao longo de 64 dias. Mais precisamente, cada rato é pesado no dia 1 e a cada sete dias até ao dia 64, com uma pesagem adicional no dia 44. Há três grupos de ratos, cada grupo sujeito a uma dieta diferente.

a) Execute o comando *plot(BodyWeight)* e comente.

b) Ajuste o modelo que lhe parecer adequado com a função *lme* do package *nlme*. Execute os comandos *summary*, *anova*, *ranef* e *fitted*. Explique como se obtém cada valor ajustado.

c) Uma vez no mesmo individuo são realizadas várias pesagens ao longo do tempo, o pressuposto de independência entre observações é posto em causa. Um vez que as observações não são igualmente espaçadas no tempo (note-se que se fez uma observação adicional no dia 44), é mais adequado considerar uma estrutura de correlação espacial para os erros aleatórios. No package *nlme* existem várias estruturas de correlação que podem ser aplicáveis ao caso em estudo, por exemplo, *corExp*, *corGaus*, *corSpher*. Efectue os comandos:

```
bodyw2.lme<-update(bodyw1.lme, corr=corExp(form=~Time))
```

```
bodyw3.lme<-update(bodyw1.lme, corr=corGaus(form=~Time))
```

```
bodyw4.lme<-update(bodyw1.lme, corr=corSpher(form=~Time)).
```

Com base no AIC e no BIC qual é a estrutura de correlação mais adequada?

d) O modelo seleccionado em c) difere significativamente do modelo ajustado em b)?

ANEXO: Scripts no R

EXERCÍCIO 1#

```
touriga<-read.table("touriga.txt", header=T)
touriga
#alínea b)#
touriga.aov<-aov(rend~clone, data=touriga)
summary(touriga.aov)
#teste à componente de variância genotípica#
#valor que numa distribuição F com 195 e 784 graus de liberdade, deixa à direita uma região de
probabilidade 0.05#
qf(1-0.05,195,784)
#p-value do teste#
1-pf(5.22,195,784)
#ajustar o modelo usando o método REML#
#alinea c)#
#lme do package nlme#
tourigalme1<-lme(rend~1, random=~1|clone, data=touriga)
summary(tourigalme1)
#para obter as estimativas das componentes de variância#
VarCorr(tourigalme1)
#lmer do package lme4#
tourigalmer1<-lmer(rend~1+(1|clone), data=touriga)
summary(tourigalmer1)
#varComp do package varComp#
tourigavarcomp1<-varComp(rend~1, random=~clone, data=touriga)
summary(tourigavarcomp1)
#procedimento para realizar um teste de razão de verossimilhanças à componente de variância genotípica#
#REML log-verossimilhança do modelo 1#
logLik(tourigavarcomp1)
#Ajustamento do modelo sem factor de efeitos aleatórios (submodelo)#
tourigavarcomp0<-varComp(rend~1, data=touriga)
#REML log-verossimilhança do modelo sem efeitos aleatórios#
logLik(tourigavarcomp0)
h1<-logLik(tourigavarcomp1)
h0<-logLik(tourigavarcomp0)
estatisticacalculada<-2*(h1-h0)
estatisticacalculada
'log Lik.' 275.5467 (df=2)
qchisq(1-0.05,1)
[1] 3.841459
1-pchisq(275.5467,1)
2#EBLUPS dos efeitos aleatórios#
ranef(tourigalme1)
#valores ajustados#
fitted(tourigalme1)
#alínea d)#
#ajustamento do modelo com dois factores de efeitos aleatórios: clone e bloco#
#com o pacote lme4#
tourigalmer2<-lmer(rend~1+(1|clone)+(1|bloco), data=touriga)
summary(tourigalmer2)
```



```

#procedimento para realizar de testes de razão de verosimilhanças à componente de variância genotípica e
à componente de variância do bloco#
#REML log-verosimilhança do modelo 2#
logLik(tourigalmer2)
#ajustamento do modelo só com o factor de efeitos aleatórios (bloco)#
tourigalmer3<-lmer(rend~1+(1|bloco), data=touriga)
#REML log-verosimilhança do modelo 3#
logLik(tourigalmer3)
#ajustamento do modelo só com o factor de efeitos aleatórios (clone)#
tourigalmer4<-lmer(rend~1+(1|clone), data=touriga)
#REML log-verosimilhança do modelo 4#
logLik(tourigalmer4)

```

#EXERCÍCIO 2 #

```

tomate<-read.table("tomate.txt", header=T)
#No package lme4, ajustamento do modelo a dois factores de efeitos aleatórios, sem interacção: factor
variedade de tomate, com 41 níveis, e factor bloco, com 4 níveis#
tomate.lmer1<-lmer(pesofruto~1+(1|variedade)+(1|bloco), data=tomate)
summary(tomate.lmer1)

#o teste de hipóteses adequado para responder ao objectivo do estudo: trata-se de um teste de hipóteses à
componente de variância associada ao factor variedade (ver aulas para descrição completa do teste de
razão de verosimilhanças)#
#Para obter o valor da log-verosimilhança restrita do modelo ajustado#
loglikH1<-logLik(tomate.lmer1)
loglikH1
#Para obter o valor da log-verosimilhança restrita do modelo sem o factor variedade#
tomate.lmer0<-lmer(pesofruto~1+(1|bloco), data=tomate)
loglikH0<-logLik(tomate.lmer0)
loglikH0
#cálculo da estatística do teste#
estcalc<-2*(loglikH1-loglikH0)
estcalc
#p-value do teste#
1-pchisq(estcalc,1)

```

#EXERCÍCIO 3#

```

terrenos<-read.table("terrenos.txt", header=T)
head(terrenos)
#ajustamento do modelo com 1 factor de efeitos fixos (variedade) e 1 factor de efeitos aleatórios
(terreno), método REML#
terrenoslme1<-lme(rend~variedade, random=~1|terreno, data=terrenos)
summary(terrenoslme1)

#cálculo da estatística do teste aos efeitos fixos#

#matriz de covariâncias estimadas dos estimadores dos efeitos fixos#
vcov(terrenoslme1)

```

```

c<-vcov(terrenoslme1)
#Construção da matriz L#
l<-matrix(c(0,1,0,0,0,0,1,0,0,0,0,1),4,3)
l
#vector com as estimativas dos efeitos fixos#
b<-c(1.556,-0.0238462, -0.3890769,-0.3778462)
#cálculo do numerador da estatística F do teste aos efeitos fixos#
numerador<-t(b)%*%l%*%solve(t(l)%*%c%*%l)%*%t(l)%*%b
#informação sobre a matriz L, nomeadamente sobre a característica da matriz#
qr(l)
#cálculo da estatística do teste aos efeitos fixos#
F<-numerador/3
F
#teste aos efeitos fixos directamente com o comando anova#
anova(terrenoslme1)

```

#EXERCICIO 4#

```

#com o lme4, função lmer#
#para um estudo descritivo preliminar#
plot.design(Machines)
attach(Machines)
interaction.plot(Machine, Worker, score)
detach()
#Ajustamento do modelo descrito na alínea a): modelo linear misto, 1 factor de efeitos fixos (máquina,
com 3 níveis, 1 factor de efeitos aleatórios (trabalhador, com 6 níveis), com interacção#
machines1r<-lmer(score~Machine+(1|Worker)+(1|Worker:Machine), data=Machines)
#para ver os resultados do ajustamento#
summary(machines1r)
#Para ver as estimativas de máxima verosimilhança#
summary(lmer(score~Machine+(1|Worker)+(1|Worker:Machine), REML=FALSE, data=Machines))
#comandos para efectuar o teste de hipóteses à componente de variância da interacção#
#ajustamento do modelo sem interacção#
machines2r<-lmer(score~Machine+(1|Worker), data=Machines)
summary(machines2r)
#log-verosimilhança restrita do modelo com interacção"
lr1<-logLik(machines1r)
lr1
#log-verosimilhança restrita do modelo sem interacção"
lr0<-logLik(machines2r)
lr0
#cálculo da estatística do teste de razão de verosimilhanças#
estcal<-2*(lr1-lr0)
estcal
#p-value#
1-pchisq(estcal,1)

#comandos para efectuar o teste de hipóteses à componente de variância associada factor trabalhador#
#ajustamento do modelo sem factor trabalhador#
machines3r<-lmer(score~Machine+(1|Worker:Machine), data=Machines)

```

```

#log-verosimilhança restrita do modelo sem factor trabalhador"
lr30<-logLik(machines3r)
lr30
#cálculo da estatística do teste de razão de verosimilhanças#
estcal2<-2*(lr1-lr30)
estcal2
#p-value#
1-pchisq(estcal2,1)
#p-value corrigido de acordo com a verdadeira distribuição da estatística do teste#
0.5*(1-pchisq(estcal2,1))

#Para efectuar o teste de hipóteses aos efeitos fixos do modelo, consultar o valor da estatística do teste
calculada#
anova(machines1r)
#valor que numa distribuição F com 2 e 10 graus de liberdade deixa à direita uma região de probabilidade
0.05#
qf(1-0.05,2,10)
#p-value#
1-pf(20.576,2,10)

#gráficos de diagnóstico para validação de alguns dos pressupostos do modelo#
#gráfico dos resíduos vs valores ajustados"
plot(machines1r)
"qqnorm dos resíduos#
residuos<-resid(machines1r)
qqnorm(residuos)
#qqnorm dos preditores dos efeitos dos trabalhadores#
eblupsworker<-ranef(machines1r)$Worker
qqnorm(eblupsworker[,1])
#qqnorm dos preditores dos efeitos da interacção#
eblupsinteracao<-ranef(machines1r)$`Worker:Machine`
qqnorm(eblupsinteracao[,1])

#EXERCÍCIO 5#
antaovaz<-read.table("antaovaz.txt", header=T)
#no package lme4, ajustamento do modelo linear misto com 1 factor de efeitos fixos (clone, com 6
níveis), 1 factor de efeitos aleatórios (ambiente, com 11 níveis), com interacção (por cada combinação
clonexambiente fizeram-se 8 observações), método REML#
antaovazlmer1<-lmer(rend~clone+(1|ambiente)+(1|ambiente:clone), data=antaovaz)
summary(antaovazlmer1)
#Proceda de forma adequada para avaliar se existe variabilidade da interacção: teste de razão de
verosimilhanças à componente de variância associada à interacção cloneXambiente#
#nota: descrever o teste detalhadamente por escrito, hipóteses, estatística do teste, distribuição da
estatística do teste e região crítica##
loglikH1<-logLik(antaovazlmer1)
#ajustamento do modelo sem interacção#
antaovazlmer0<-lmer(rend~clone+(1|ambiente), data=antaovaz)
loglikH0<-logLik(antaovazlmer0)
#cálculo da estatística do teste#
estcalc<-2*(loglikH1-loglikH0)

```

```
estcalc
#p-value do teste#
1-pchisq(estcalc,1)
```

```
#Proceda de forma adequada para avaliar se existe variabilidade entre ambientes: teste de razão de
verosimilhanças à componente de variância associada ao factor ambiente# #
```

```
#nota: descrever o teste detalhadamente por escrito, hipóteses, estatística do teste, distribuição da
estatística do teste e região crítica##
```

```
#ajustamento do modelo sem factor ambiente#
```

```
antaovazlmer2<-lmer(rend~clone+(1|ambiente:clone), data=antaovaz)
```

```
loglikH02<-logLik(antaovazlmer2)
```

```
#cálculo da estatística do teste#
```

```
estcalc2<-2*(loglikH1-loglikH02)
```

```
estcalc2
```

```
#p-value do teste#
```

```
1-pchisq(estcalc2,1)
```

```
#Proceda de forma adequada para avaliar se existem efeitos principais associados aos clones: teste de
hipóteses aos efeitos fixos do modelo#
```

```
#nota: descrever o teste detalhadamente por escrito, hipóteses, estatística do teste, distribuição da
estatística do teste e região crítica##
```

```
#consultar o valor da estatística do teste calculada, estatística  $F=QMA/QMAB$ #
```

```
anova(antaovazlmer1)
```

```
#valor que numa distribuição F com 5 e 50 graus de liberdade deixa à direita uma região de probabilidade
0.05#
```

```
qf(1-0.05,5,50)
```

```
#p-value#
```

```
1-pf(0.5659,5,50)
```

#EXERCICIO 6#

```
#alínea b)#
```

```
plot.design(Alfalfa)
```

```
attach(Alfalfa)
```

```
interaction.plot(Date, Variety, Yield)
```

```
detach(Alfalfa)
```

```
#alínea c), ajustamento do modelo descrito na alínea a)#
```

```
alfalfalmer1<-lmer(Yield~Variety+Date+Variety*Date+(1|Block)+(1|Block:Variety)+(1|Block:Date),
data=Alfalfa)
```

```
summary(alfalfalmer1)
```

```
logLik(alfalfalmer1)
```

```
#ajustamento do modelo descrito na alínea a), mas sem a interacção Block*Date#
```

```
alfalfalmer2<-lmer(Yield~Variety+Date+Variety*Date+(1|Block)+(1|Block:Variety), data=Alfalfa)
```

```
summary(alfalfalmer2)
```

```
logLik(alfalfalmer2)
```

```
#comparação dos dois modelos#
```

```
estcal1<-2*(-4.4728+4.978184)
```

```
estcal1
```

```
1-pchisq(estcal1,1)
```

```
0.5*(1-pchisq(estcal1,1))
```

```
#Foi escolhido o modelo 2, o que se segue é com base neste modelo#  
summary(alfalfalmer2)
```

```
#alínea d), teste aos efeitos fixos#  
anova(alfalfalmer2)  
#p-value do teste aos efeitos da interacção#  
1-pf(1.2548,6,45)  
#p-value do teste aos efeitos principais da variedade#  
1-pf(0.6534,2,10)  
#p-value do teste aos efeitos principais da data de corte#  
1-pf(23.3897,3,45)
```

```
#o mesmo modelo ajustado com a função lme do package nlme#  
alfafalme<-lme(Yield~Variety*Date, random=~1|Block/Variety, data=Alfalfa)  
summary(alfafalme)  
anova(alfafalme)
```

```
#alínea e)#  
aov(Yield~Variety*Date+Error(Block*Variety), data=Alfalfa)  
summary(aov(Yield~Variety*Date+Error(Block*Variety), data=Alfalfa))
```

#EXERCÍCIO 7#

```
#No package “nlme” do R, existe um conjunto de dados (designado por “Oats”)#
```

```
library(nlme)
```

```
Oats
```

```
#utilizar o comando factor(nitro) para que esta variável seja considerada como factor#
```

```
nitrof<-as.factor(Oats$nitro)
```

```
#Ajustamento do modelo no R com a função lmer: análise de um split-plot, factor variedade de efeitos  
fixos, factor nitro como de efeitos fixos, interacção variedade*nitro de efeitos fixos, factor bloco de  
efeitos aleatórios, interacção bloco*variedade de efeitos aleatórios, interacção bloco*nitro como efeitos  
aleatórios#
```

```
oatslmer1<-lmer(yield~Variety+nitrof+nitrof*Variety+(1|Block)+(1|Block:Variety)+(1|Block:nitrof),  
data=Oats)
```

```
summary(oatslmer1)
```

```
#Pela análise do resultado do ajustamento, conclui-se que se pode retirar a interacção bloco*nitrof#
```

```
#ajustamento do modelo descrito anteriormente, mas sem a interacção Block*nitrof#
```

```
oatslmer2<-lmer(yield~Variety+nitrof+nitrof*Variety+(1|Block)+(1|Block:Variety), data=Oats)  
summary(oatslmer2)
```

```
#Conclusão sobre a influência da variedade e da concentração de nitrogénio sobre o rendimento#
```

```
#testes aos efeitos fixos da interacção, aos efeitos fixos da variedade e aos efeitos fixos da concentração  
de nitrogénio#
```

```
#valor calculado da estatística F para cada caso)
```

```
anova(oatslmer2)
```

```
#p-value do teste aos efeitos da interacção#
```

```
1-pf(0.3028,6,45)
```

```
#p-value do teste aos efeitos principais da variedade#  
1-pf(1.4853,2,10)  
#p-value do teste aos efeitos principais da concentração de nitrogénio#  
1-pf(37.6856,3,45)
```

#EXERCICIO 8#

```
#alínea a)#  
plot(BodyWeight)  
#alínea b), ajustamento do modelo linear misto com a função lme#  
bodyw1.lme<-(lme(weight~Time*Diet,random=~Time|Rat, data=BodyWeight))  
summary(bodyw1.lme)  
anova(bodyw1.lme)  
#blups dos efeitos aleatórios#  
ranef(bodyw1.lme)  
#valores ajustados#  
fitted(bodyw1.lme)  
  
#alínea d)#  
bodyw2.lme<-update(bodyw1.lme, corr=corExp(form=~Time))  
bodyw3.lme<-update(bodyw1.lme, corr=corGaus(form=~Time))  
bodyw4.lme<-update(bodyw1.lme, corr=corSpher(form=~Time))  
#tratam-se de modelos não encaixados, só podem ser comparados com base nos critérios de informação#  
anova(bodyw2.lme,bodyw3.lme, bodyw4.lme)  
summary(bodyw2.lme)  
  
#alínea e) - são modelos encaixados, pode-se fazer um teste de razão de verosimilhanças##  
anova(bodyw1.lme,bodyw2.lme)
```