

Universidade Estadual de Maringá

Graduação em Estatística - Dep. de Estatística

Disciplina: **DES8076 - Modelos Mistos**

Professor: **Prof. Vanderly Janeiro (DES/UEM)**

2ª Lista de exercícios

1. A Tabela 1 fornece medições derivadas de um experimento em microbiologia (MOHAMED, 2015)¹. O objetivo principal foi determinar o número de colônias do patógeno *Escherichia coli* 0157:H7 em amostras fecais contaminadas coletadas de 12 carcaças de bovinos. Para uma determinada amostra fecal, o número de colônias foi determinado por um novo teste (Petrifilm HEC) e por um “teste padrão” em duas subamostras; os resultados são registrados como o logaritmo do número de colônias (Tabela 1). As duas primeiras linhas correspondem às determinações repetidas com base no uso do teste padrão e as duas segundas linhas correspondem às determinações repetidas do novo teste.

Tabela 1: Logaritmo do número de colônias de *E. coli* 157:H7 em amostras retiradas de 12 carcaças de Bovinos.

Teste (subamostra)		1	2	3	4	5	6	7	8	9	10	11	12
Standard	1	2.356	2.149	2.452	2.255	2.694	2.43	2.322	2.322	2.491	2.322	2.322	2.491
	2	2.384	2.263	2.417	2.299	2.684	2.44	2.491	2.041	2.322	2.322	2.491	2.785
New test	1	2.283	2.061	2.322	2.162	2.068	2.322	2.491	2.041	2.322	2.491	2.041	2.785
	2	2.265	1.987	2.316	2.127	2.111	2.28	2.491	2.041	2.041	2.71	2.322	2.322

- (a) Descreva o modelo que lhe parece ser adequado para estudar a e decidir entre os métodos.
 - (b) Obtenha o coeficiente de correlação intraclasse para cada um dos métodos de medição e conclua sobre eles. (sugestões: <https://stats.stackexchange.com/questions/15768/computing-repeatability-of-effects-from-an-lmer-model>)
2. Seja um experimento em um Delineamento Inteiramente Casualizado (DIC) no qual avaliou-se a produção em volume de madeira aos 3 anos de idade de três clones ($I=3$), tomados de forma aleatória, em 3 parcelas experimentais ($J=3$). Considerando os dados abaixo:

¹Shoukri, Mohamed. (2015). Measurement of Agreement. 10.1002/9781118445112.stat05301.pub2.

Repetição	Clone	A	Clone	B	Clone	C
1	31		35		13	
2	36		19		19	
3	39		31		31	

- (a) Descreva o modelo que lhe parece ser adequado para estudar a variabilidade.
 - (b) Estime os parâmetros fixos envolvidas no modelo descrito na alínea a).
 - (c) Considere a forma matricial do modelo da alínea a), $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$, em que $\mathbf{u} \sim \mathbf{N}(0, \mathbf{G})$ e $\boldsymbol{\varepsilon} \sim \mathbf{N}(0, \mathbf{R})$, apresente as matrizes \mathbf{Z} , \mathbf{G} , \mathbf{R} e obtenha $\mathbf{V} = \text{Var}(\mathbf{Y})$
 - (d) Repita as alíneas a), b) e c) considerando que cada repetição ocorrem em um bloco.
3. Num estudo de seleção de videira, com vista a estudar a variabilidade genética do rendimento da casta Touriga Nacional, foi instalado um ensaio de campo, 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 no `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, efetue um teste à existência de variabilidade genética do rendimento da casta (utilize o nível de significância de 0.05).
 - (c) Ajuste o modelo descrito na alínea a) com o método de estimação de máxima verosimilhança restrita, recorrendo
 - i. ao comando `lme` do pacote `nlme`; ao comando `lmer` do pacote `lme4`;
 - ii. aplique o comando `summary` aos objetos criados anteriormente e identifique que as estimativas de máxima verosimilhança restrita das componentes de variância. Compare os resultados agora obtidos com os da alínea b(i).
 - iii. Efetue um teste à existência de variabilidade genética do rendimento (teste de razão de verosimilhanças).
 - iv. 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).
- Ajuste um novo modelo que incorpore o efeito do bloco (considere também que é um fator de efeitos aleatórios). Utilize o pacote `lme4`.
 - Efetue testes de hipóteses às componentes de variância associadas aos fatores clone e bloco.
 - Calcule o AIC e o BIC dos dois modelos ajustados (o da alínea a) e o da alínea d)) e selecione o melhor modelo com base nesses critérios.
4. Considere os dados disponíveis no `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 formam uma amostra aleatória da população de terrenos onde essas variedades foram cultivadas.
- Descreva o modelo apropriado a este contexto.
 - 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).

	<i>gl</i>	<i>SQ</i>	<i>QM</i>	<i>Fcalc</i>
<i>Variedade</i>	___	1.799	___	___
<i>Terreno</i>	___	2.407	___	___
<i>Resíduos</i>	36	3.513	0.0976	

- 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.
- Execute o comando `plot.design` (`terrenos`).
- Considere o resultado referente ao ajustamento do modelo utilizando o método REML.
`> terrenolme1 <- lme(rend ~ variedade, random = 1|terreno, data = terrenos)`

Recorde os testes de hipóteses a combinações lineares dos efeitos fixos do um modelo linear misto. 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`).

- 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
```

```
> logLik(terrenosvarcomp1)
```

```
> terrenosvarcomp0 < -varComp(rend ~ variedade, data = terrenos)
```

```
> logLik(terrenosvarcomp0)
```

Interprete os comandos anteriores e efetue um teste de razão de verossimilhanças 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, ~ rane(.))
```