



Modelos Mistos - Lista 2

```
## importando pacote e dados necessários
```

```
library(nlme)  
data("Orthodont")  
data = Orthodont
```

Exercício 1

Considere a base Orthodont, ajuste o modelo misto da distância em função da idade e do sexo, com o efeito aleatório sujeito, considerando o método ML e REML. Observe se existe diferenças.

Resposta:

Primeiramente ajustamos o modelo pelo método da máxima verossimilhança.

```
## ajustando um modelo de efeitos fixos pelo método da máxima verossimilhança  
lmeML1=lme(fixed= distance~age + Sex, random=~ 1 |Subject, method="ML",data=data)
```

```
## resultados  
lmeML1
```

```
## Linear mixed-effects model fit by maximum likelihood  
##   Data: data  
##   Log-likelihood: -217.4282  
##   Fixed: distance ~ age + Sex  
## (Intercept)      age  SexFemale  
## 17.7067130    0.6601852 -2.3210227  
##  
## Random effects:  
##   Formula: ~1 | Subject  
##      (Intercept) Residual  
## StdDev:      1.730079 1.422728  
##  
## Number of Observations: 108  
## Number of Groups: 27
```

As estimativas dos parâmetros são:

- intercepto: 17,7
- idade: 0,66
- sexo: caso
- desvio padrão dos resíduos: 1.4
- desvio padrão do efeito aleatório: 1,7

Ajustando agora o modelo com o método da máxima verossimilhança restrita vemos que as estimativas dos coeficientes de efeito fixo se mantiveram. As estimativas dos desvios sofreram mudança.

```
## ajustando um modelo de efeitos fixos pelo método da máxima verossimilhança restrita
lmeREML1=lme(fixed= distance~age + Sex, random=~ 1 |Subject, method="REML",data=data)
```

```
## resultados
summary(lmeREML1)
```

```
## Linear mixed-effects model fit by REML
## Data: data
##      AIC      BIC    logLik
##  447.5125 460.7823 -218.7563
##
## Random effects:
## Formula: ~1 | Subject
##      (Intercept) Residual
## StdDev:      1.807425 1.431592
##
## Fixed effects: distance ~ age + Sex
##              Value Std.Error DF   t-value p-value
## (Intercept) 17.706713 0.8339225 80 21.233044 0.0000
## age          0.660185 0.0616059 80 10.716263 0.0000
## SexFemale    -2.321023 0.7614168 25 -3.048294 0.0054
## Correlation:
##      (Intr) age
## age      -0.813
## SexFemale -0.372 0.000
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -3.74889609 -0.55034466 -0.02516628 0.45341781 3.65746539
##
## Number of Observations: 108
## Number of Groups: 27
```

Exercício 2

Refaça o item anterior considerando outra parametrização.

Resposta:

A parametrização padrão para variáveis categóricas no pacote nlme é a casela de referência, onde crianças do sexo masculino foram marcadas com '0' e crianças do sexo feminino como '1'.

```
lmeML1$contrasts
```

```
## $Sex
##      Female
## Male      0
## Female    1
```

Agora iremos alterar a parametrização para contraste da soma, onde o sexo masculino aparece como '1' e o sexo feminino como '-1'.

Ajustando o modelo com estimativa de máxima verossimilhança para os parâmetros vemos que apenas as estimativas do intercepto e do coeficiente associado ao fator fixo de idade se alteram. As demais estimativas se mantêm iguais ao modelo ajustado pela máxima verossimilhança com a parametrização casela de referência.

```
lmeML2=lme(fixed= distance~age + Sex, random=~ 1 |Subject, method="ML",data=data,
           contrasts = list(Sex = "contr.sum"))

summary(lmeML2)
```

```
## Linear mixed-effects model fit by maximum likelihood
## Data: data
##      AIC      BIC    logLik
##  444.8565 458.2671 -217.4282
##
## Random effects:
## Formula: ~1 | Subject
##      (Intercept) Residual
## StdDev:      1.730079 1.422728
##
## Fixed effects: distance ~ age + Sex
##              Value Std.Error DF   t-value p-value
## (Intercept) 16.546202 0.7775322 80 21.280406  0.0000
## age          0.660185 0.0620929 80 10.632212  0.0000
## Sex1         1.160511 0.3715334 25  3.123572  0.0045
## Correlation:
##      (Intr) age
## age  -0.878
## Sex1 -0.088  0.000
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -3.77682007 -0.55426744 -0.01578248  0.45835495  3.68124620
##
## Number of Observations: 108
## Number of Groups: 27
```

```
lmeML2$contrasts
```

```
## $Sex
##      [,1]
## Male      1
## Female   -1
```

Similar a situação anterior, quando comparamos as estimativas do modelo estimado com máxima verossimilhança restrita, e diferença na parametrização de sexo, apenas intercepto e estimativa de coeficiente associado ao sexo são diferentes. Além disso, as variâncias de ambos os modelos permaneceram iguais.

```
lmeREML2=lme(fixed= distance~age + Sex, random=~ 1 |Subject, method="REML",data=data,
             contrasts = list(Sex = "contr.sum"))

summary(lmeREML2)
```

```
## Linear mixed-effects model fit by REML
## Data: data
##      AIC      BIC    logLik
##  448.8988 462.1686 -219.4494
##
## Random effects:
## Formula: ~1 | Subject
##      (Intercept) Residual
## StdDev:      1.807425 1.431592
##
## Fixed effects: distance ~ age + Sex
##              Value Std.Error DF   t-value p-value
## (Intercept) 16.546202 0.7772830 80 21.287229 0.0000
## age          0.660185 0.0616059 80 10.716263 0.0000
## Sex1         1.160511 0.3807084 25  3.048294 0.0054
## Correlation:
##      (Intr) age
## age -0.872
## Sex1 -0.091  0.000
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -3.74889609 -0.55034466 -0.02516628  0.45341781  3.65746539
##
## Number of Observations: 108
## Number of Groups: 27
```

Exercício 3

Considerando valores próximos as estimativas de ML para os parâmetros fixos e de REML para as componentes de variância, simule duas novas bases de dados, uma em que o número de meninos é 32 e o de meninas é 22, e outra em que o número de meninos é 96 e o de meninas 33.

Resposta:

```
age = 0.6
sex = -2.5
intercepto = 17
sigmab = 1.85^2 #o summary solta o desvio padrão
sigma = 1.45^2
set.seed(2109)
#Para 32 meninos e 22 meninas
head(data)
```

```
## Grouped Data: distance ~ age | Subject
## distance age Subject Sex
## 1      26.0   8      M01 Male
## 2      25.0  10      M01 Male
## 3      29.0  12      M01 Male
## 4      31.0  14      M01 Male
## 5      21.5   8      M02 Male
## 6      22.5  10      M02 Male
```

```

X = model.matrix(~age + Sex, data)
X2 = rbind(X,X) #Dobrei a matrix X
beta1 = c(intercepto, age, sex)
muv = X2%*%beta1
Zi = matrix(1,4,1) #matriz dos efeitos aleatorios de cada individuo i (são iguais para todo i)
Vi = sigma*(diag(4)) + sigmab*(Zi%*%(t(Zi))) # Matriz de variância para indiv i
V = kronecker(diag(27*2), Vi) #matriz de variancia geral, essa multiplicação cria uma matriz quadrada em que cada elemento da diagonal é Vi
library(MASS)
y = mvrnorm(1, muv, V) #simulando da normal multivariada
individuo = rep(seq(1:54),each = 4) #gerando o indice dos individuos
datasim2 = data.frame(cbind(y, X2, individuo))
head(datasim2)

```

```

##          y X.Intercept. age SexFemale individuo
## X1 21.59705          1   8          0          1
## X2 25.20810          1  10          0          1
## X3 25.51742          1  12          0          1
## X4 25.88901          1  14          0          1
## X5 24.60869          1   8          0          2
## X6 24.49393          1  10          0          2

```

```

names(datasim2) = c("y","Intercept", "age", "Sex","Subject")
table(datasim2$Sex)/4 # 4 coletas por indivíduo

```

```

##
##  0  1
## 32 22

```

#Para 96 meninos e 32 meninas

```

X = model.matrix(~age + Sex, data)
Xmale = X[X[,3] == 0,]
Xmale = matrix(rep(t(Xmale), 6), ncol = ncol(Xmale), byrow = TRUE)
nrow(Xmale)/4

```

```
## [1] 96
```

```

Xfemale = X[X[,3] == 1,]
Xfemale = matrix(rep(t(Xfemale), 3), ncol = ncol(Xfemale), byrow = TRUE)
nrow(Xfemale)/4

```

```
## [1] 33
```

```

X3 = rbind(Xmale, Xfemale)
beta1 = c(intercepto, age, sex)
muv = X3%*%beta1
Zi = matrix(1,4,1)
Z = kronecker(diag(129), Zi)
Vi = sigma*(diag(4)) + sigmab*(Zi%*%(t(Zi)))

```

```
V = kronecker(diag(129), Vi)
library(MASS)
y = mvrnorm(1, muv, V)
indivduo = rep(seq(1:129),each = 4)
datasim3 = data.frame(cbind(y, X3, indivduo))
head(datasim3)
```

```
##           y V2 V3 V4 indivduo
## 1 22.62393  1  8  0          1
## 2 21.54880  1 10  0          1
## 3 21.22283  1 12  0          1
## 4 24.52932  1 14  0          1
## 5 19.16168  1  8  0          2
## 6 20.16732  1 10  0          2
```

```
names(datasim3) = c("y", "Intercept", "age", "Sex", "Subject")
```

Exercício 4

Realize o solicitado nos itens anteriores para a nova base.

Resposta:

Base simulada 1: 32 meninos e 22 meninas.

```
lmeML_sim1=lme(fixed= y~age + Sex, random=~ 1 |Subject, method="ML",data=datasim2)
lmeML_sim1

## Linear mixed-effects model fit by maximum likelihood
##   Data: datasim2
##   Log-likelihood: -428.5442
##   Fixed: y ~ age + Sex
##   (Intercept)          age          Sex
##    17.052539     0.605882    -3.426249
##
## Random effects:
##   Formula: ~1 | Subject
##           (Intercept) Residual
## StdDev:    1.924368 1.330241
##
## Number of Observations: 216
## Number of Groups: 54

lmeREML_sim1=lme(fixed= y~age + Sex, random=~ 1 |Subject, method="REML",data=datasim2)
lmeREML_sim1

## Linear mixed-effects model fit by REML
##   Data: datasim2
##   Log-restricted-likelihood: -430.8304
```

```
## Fixed: y ~ age + Sex
## (Intercept)      age      Sex
## 17.052539    0.605882   -3.426249
##
## Random effects:
## Formula: ~1 | Subject
## (Intercept) Residual
## StdDev:    1.96466 1.334366
##
## Number of Observations: 216
## Number of Groups: 54
```

Base simulada 2: 96 meninos e 33 meninas.

```
lmeML_sim2=lme(fixed= y~age + Sex, random=~ 1 |Subject, method="ML",data=datasim3)
lmeML_sim2
```

```
## Linear mixed-effects model fit by maximum likelihood
## Data: datasim3
## Log-likelihood: -1059.951
## Fixed: y ~ age + Sex
## (Intercept)      age      Sex
## 17.0602429    0.5663241   -2.1310935
##
## Random effects:
## Formula: ~1 | Subject
## (Intercept) Residual
## StdDev:    1.833312 1.475319
##
## Number of Observations: 516
## Number of Groups: 129
```

```
lmeREML_sim2=lme(fixed= y~age + Sex, random=~ 1 |Subject, method="REML",data=datasim3)
lmeREML_sim2
```

```
## Linear mixed-effects model fit by REML
## Data: datasim3
## Log-restricted-likelihood: -1063.393
## Fixed: y ~ age + Sex
## (Intercept)      age      Sex
## 17.0602429    0.5663241   -2.1310935
##
## Random effects:
## Formula: ~1 | Subject
## (Intercept) Residual
## StdDev:    1.849627 1.477229
##
## Number of Observations: 516
## Number of Groups: 129
```

Exercício 5

Compare os resultados

Resposta:

Modelo	Intercepto	Idade	Sexo	$\hat{\sigma}$	$\hat{\sigma}_d$	$\hat{\sigma}_d/\hat{\sigma}$
ML	17.70671	0.6601852	-2.321023	1.730079	1.422728	0.8223486
ML reparametrizado	16.54620	0.6601852	1.160511	1.730079	1.422728	0.8223486
ML Simulado 1	17.05254	0.6058820	-3.426249	1.924368	1.330241	0.6912612
ML Simulado 2	17.06024	0.5663241	-2.131093	1.833312	1.475319	0.8047288
REML	17.70671	0.6601852	-2.321023	1.807425	1.431592	0.7920616
REML reparametrizado	16.54620	0.6601852	1.160511	1.807425	1.431592	0.7920616
REML Simulado 1	17.05254	0.6058820	-3.426249	1.964660	1.334366	0.6791842
REML Simulado 2	17.06024	0.5663241	-2.131093	1.849627	1.477229	0.7986632

Pela tabela acima percebe-se que as estimativas dos efeitos fixos permanecem iguais quando ajustado no mesmo banco de dados, independente se foi utilizado o método de Máxima verossimilhança ou máxima verossimilhança restrita. A principal diferença entre os métodos se encontra nas estimativas dos desvios padrão, em que o método de máxima verossimilhança subestima σ . Utilizando um banco de dados maior, que são as bases simuladas, vemos que a diferença entre os desvios é menor, o que leva a considerar que não existe diferença entre os métodos para uma amostra suficientemente grande.