

# Mixed Models Exam

## Question 1: contrasts for dataset “needle\_litter\_fall\_rbd.txt”

This script is already included in the Answer for Homework Exercise 1

```
#rm(list = ls()) # Clean the workspace
#install.packages("lme4") # If necessary install the library
#install.packages("nlme") # If necessary install the library
library(lme4) # Call the library to the workspace
library(car) # Call the library to the workspace
library(lsmmeans) # Call the library to the workspace
library(multcomp) # Call the library to the workspace
library(nlme) # Call the library to the workspace
# Read the data
# "PUT HERE THE ADDRESS OF THE FILE " #
myfile<-"C:/Users/toledo/Dropbox/UNIPD/Biostatistics Course R Spring 2018/corso STAT PhD 2018 Mixed Models/dat
rbd_litter<-read.table(file=myfile,stringsAsFactors = TRUE,header = TRUE,sep = "\t")
options(contrasts = c("contr.SAS","contr.poly")) # Change contrast options
rbd_litter$block<-as.factor(rbd_litter$block) # Set block as factor
#rbd_litter$treatment<-as.factor(rbd_litter$treatment) # Set treatment as factor: not necessary because treatm

# Fit the mixed model
model.1<-lmer(needle_litter_fall ~ treatment + (1 | block) ,data = rbd_litter, REML = TRUE)
lsmeans(model.1,"treatment") # LSM

## treatment lsmean SE df lower.CL upper.CL
## C 7.5 0.2198484 3.92 6.884785 8.115215
## I 7.4 0.2198484 3.92 6.784785 8.015215
## IF 9.1 0.2198484 3.92 8.484785 9.715215
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95

#Adding the contrasts to model.1
coef(model.1) # See the coefficients of your model

## $block
## (Intercept) treatmentC treatmentI
## 1 8.853304 -1.6 -1.7
## 2 9.130837 -1.6 -1.7
## 3 9.624229 -1.6 -1.7
## 4 8.791630 -1.6 -1.7
##
## attr(,"class")
## [1] "coef.mer"

library(multcomp)
K <- matrix(c(0, 1, -1), 1) # Contrast treatment1 vs treatment2
t <- glht(model.1, linfct = K) # fit a general linear hypothesis test
summary(t) # See the results: no differences among treatments

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = needle_litter_fall ~ treatment + (1 | block),
## data = rbd_litter, REML = TRUE)
##
## Linear Hypotheses:
## Estimate Std. Error z value Pr(>|z|)
## 1 == 0 0.1000 0.1374 0.728 0.467
```

```
## (Adjusted p values reported -- single-step method)
K <- matrix(c(0, 1, 0), 1)      # Contrast treatment1 vs treatment3
t <- glht(model.1, linfct = K)   # fit a general linear hypothesis test
summary(t)                      # See the results: signif. differences among treatments

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = needle_litter_fall ~ treatment + (1 | block),
## data = rbd_litter, REML = TRUE)
##
## Linear Hypotheses:
## Estimate Std. Error z value Pr(>|z|)
## 1 == 0 -1.6000 0.1374 -11.64 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

K <- matrix(c(0, 0, 1), 1)      # Contrast treatment2 vs treatment3
t <- glht(model.1, linfct = K)   # fit a general linear hypothesis test
summary(t)                      # See the results: signif. differences among treatments

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = needle_litter_fall ~ treatment + (1 | block),
## data = rbd_litter, REML = TRUE)
##
## Linear Hypotheses:
## Estimate Std. Error z value Pr(>|z|)
## 1 == 0 -1.7000 0.1374 -12.37 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

K <- matrix(c(0, 1, 1), 1)      # Contrast treatment1+2 vs treatment3
t <- glht(model.1, linfct = K)   # fit a general linear hypothesis test
summary(t)                      # See the results: signif. differences among t1 vs. t2+t3

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = needle_litter_fall ~ treatment + (1 | block),
## data = rbd_litter, REML = TRUE)
##
## Linear Hypotheses:
## Estimate Std. Error z value Pr(>|z|)
## 1 == 0 -3.300 0.238 -13.86 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

#the first contrast tells you that there are no differences between t1 and t2,
#the further contrasts tell you that t3 is significantly different from 1 and 2
```

## Question 2: linear model vs mixed model for dataset “needle\_litter\_fall\_rbd.txt”

```
# Fit the linear model
model.2<-lm(needle_litter_fall ~ treatment + block ,data = rbd_litter)
Anova(model.2, type=3,test.statistic = "F") # ANOVA table SS type III
```

```
## Anova Table (Type III tests)
##
## Response: needle_litter_fall
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 153.709  1 4068.765 9.982e-10 ***
## treatment    7.280  2   96.353 2.753e-05 ***
## block        1.513  3   13.353 0.004593 **
## Residuals    0.227  6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans(model.2,"treatment") # LSM
```

```
## treatment lsmean      SE df lower.CL upper.CL
## C          7.5 0.09718253  6 7.262203 7.737797
## I          7.4 0.09718253  6 7.162203 7.637797
## IF         9.1 0.09718253  6 8.862203 9.337797
##
## Results are averaged over the levels of: block
## Confidence level used: 0.95
```

```
#comparison with model 1:
Anova(model.1, type=3,test.statistic = "F") # ANOVA table SS type III
```

```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: needle_litter_fall
##              F Df Df.res    Pr(>F)
## (Intercept) 1713.310  1  3.922 2.503e-06 ***
## treatment    96.353  2  6.000 2.753e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans(model.1,"treatment") # LSM
```

```
## treatment lsmean      SE  df lower.CL upper.CL
## C          7.5 0.2198484 3.92 6.884785 8.115215
## I          7.4 0.2198484 3.92 6.784785 8.015215
## IF         9.1 0.2198484 3.92 8.484785 9.715215
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
```

```
#no differences for the treatment effect and no differences in lsmeans
```

### Question 3: contrasts for dataset "Yield\_rbd"

```
rbd<-read.table(file="C:/Users/toledo/Dropbox/UNIPD/Biostatistics Course R Spring 2018/corso STAT PhD 2018 Mix
               stringsAsFactors = TRUE,header = TRUE,sep = "\t")

rbd$block<-as.factor(rbd$block)           # Set block as factor
rbd$fertilizer<-as.factor(rbd$fertilizer) # Set fertilizer as factor
contrasts(rbd$block)<-contr.SAS            # Change the reference grid to SAS
contrasts(rbd$fertilizer)<-contr.SAS       # Change the reference grid to SAS

# Fit the mixed model (here called model.3, in the lesson's exercises it is model.2)
model.3<-lmer(yield ~ fertilizer + (1 | block), data = rbd, REML = TRUE)
lsmeans(model.3,"fertilizer")             # LSM

## fertilizer    lsmean      SE    df  lower.CL upper.CL
## 1             9.50000 0.6680541 6.24   7.880260 11.11974
## 2            10.66667 0.6680541 6.24   9.046926 12.28641
## 3            12.33333 0.6680541 6.24  10.713593 13.95307
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95

#Adding the contrasts to model.3
coef(model.3)                             # See the coefficients of your model

## $block
## (Intercept) fertilizer1 fertilizer2
## 1    10.92403   -2.833333   -1.666667
## 2    14.36899   -2.833333   -1.666667
## 3    13.11628   -2.833333   -1.666667
## 4    10.61085   -2.833333   -1.666667
## 5    11.86357   -2.833333   -1.666667
## 6    13.11628   -2.833333   -1.666667
##
## attr("class")
## [1] "coef.mer"

library(multcomp)
K <- matrix(c(0, 1, -1), 1)  # Contrast Fertilizer1 vs Fertilizer2
t <- glht(model.3, linfct = K) # fit a general linear hypothesis test
summary(t)                   # See the results: signif. differences among levels

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = yield ~ fertilizer + (1 | block), data = rbd,
##       REML = TRUE)
##
## Linear Hypotheses:
##       Estimate Std. Error z value Pr(>|z|)
## 1 == 0   -1.1667     0.3801   -3.07  0.00214 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

K <- matrix(c(0, 1, 0), 1)  # Contrast Fertilizer1 vs Fertilizer3
t <- glht(model.3, linfct = K) # fit a general linear hypothesis test
summary(t)                   # See the results: signif. differences among levels

##
## Simultaneous Tests for General Linear Hypotheses
##
```

```

## Fit: lmer(formula = yield ~ fertilizer + (1 | block), data = rbd,
##       REML = TRUE)
##
## Linear Hypotheses:
##       Estimate Std. Error z value Pr(>|z|)
## 1 == 0  -2.8333      0.3801  -7.455 8.99e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

K <- matrix(c(0, 0, 1), 1)      # Contrast Fertilizer2 vs Fertilizer3
t <- glht(model.3, linfct = K)   # fit a general linear hypothesis test
summary(t)                      # See the results: signif. differences among levels

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lmer(formula = yield ~ fertilizer + (1 | block), data = rbd,
##       REML = TRUE)
##
## Linear Hypotheses:
##       Estimate Std. Error z value Pr(>|z|)
## 1 == 0  -1.6667      0.3801  -4.385 1.16e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

```

## Question 4-7

```
nested_trees<-read.table(file="C:/Users/toledo/Dropbox/UNIPD/Biostatistics Curse R Spring 2018/corso STAT PhD
                        stringsAsFactors = TRUE,header = TRUE,sep = "\t")
options(contrasts = c("contr.SAS","contr.poly")) # Change contrast options
nested_trees$Plot<-as.factor(nested_trees$Plot) # Set the variable as factor
nested_trees$Soil<-as.factor(nested_trees$Soil) # Set the variable as factor

#### Random Model
model.4<-lmer(Height_growth ~ Soil + (1|Soil:Plot), data = nested_trees ) # Fit a model with Random effect
summary(model.4) # See the results

## Linear mixed model fit by REML ['lmerMod']
## Formula: Height_growth ~ Soil + (1 | Soil:Plot)
## Data: nested_trees
##
## REML criterion at convergence: 8.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.88231 -0.58865 -0.08226  0.80333  2.32247
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Soil:Plot (Intercept) 0.04221  0.2055
## Residual              0.04783  0.2187
## Number of obs: 48, groups:  Soil:Plot, 8
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   2.4500     0.1120  21.873
## SoilA        -0.7500     0.1584  -4.735
##
## Correlation of Fixed Effects:
##      (Intr)
## SoilA -0.707

Anova(model.4, type=3,test.statistic = "F") # ANOVA

## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: Height_growth
##              F Df Df.res    Pr(>F)
## (Intercept) 478.428  1      6 5.965e-07 ***
## Soil        22.417  1      6 0.003209 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans(model.4,"Soil") # lsmeans

## Soil lsmean      SE df lower.CL upper.CL
## A      1.70 0.1120103  6 1.425921 1.974079
## B      2.45 0.1120103  6 2.175921 2.724079
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
```

## Question 8-9

```
# Read the data
myfile<-"C:/Users/toledo/Dropbox/UNIPD/Biostatistics Course R Spring 2018/curso STAT PhD 2018 Mixed Models/dat
mydata<-read.table(file=myfile,stringsAsFactors = TRUE,header = TRUE,sep = "\t")
options(contrasts = c("contr.SAS","contr.poly"))      # Change contrast options
mydata$alluminium<-as.factor(mydata$alluminium)      # Set as factor
mydata$seedling<-as.factor(mydata$seedling)          # Set as factor
```

```
# Fit the same model with different types of variance and covariance
#model.a and model.b already included in homework exercises
model.a<-lme(height ~ alluminium + time + alluminium*time, data = mydata, # Model
              random = ~ 1|seedling,                                     # Random Effect
              weights = varIdent(form = ~ 1|time),                      # Constant variance(s)
              correlation = NULL)                                       # No correlation

model.b<-lme(height ~ alluminium + time + alluminium*time, data = mydata, # Model
              random = ~ 1|seedling,                                     # Random Effect
              weights = varIdent(form = ~ 1|time),                      # Constant variance(s)
              correlation = corAR1())                                   # autocorrelation structure of order 1
```

```
#model.c included following question 8
model.c<-lme(height ~ alluminium + time + alluminium*time, data = mydata, # Model
              random = ~ 1|seedling,                                     # Random Effect
              weights = varIdent(form = ~ 1|time),                      # Constant variance(s)
              correlation = corCompSymm())                             # compound symmetry
```

```
anova(model.a,model.b,model.c)
```

##	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
##	model.a	1 21	2479.600	2553.718	-1218.800			
##	model.b	2 22	2392.086	2469.733	-1174.043	1 vs 2	89.51447	<.0001
##	model.c	3 22	2481.117	2558.764	-1218.558			

```
#the best fitted model is model.b (question 9)
```

```
anova(model.b)
```

##	numDF	denDF	F-value	p-value
## (Intercept)	1	189	270.38679	<.0001
## alluminium	3	63	1.02365	0.3882
## time	3	189	13.85646	<.0001
## alluminium:time	9	189	2.14919	0.0274