

UC Irvine ISI-BUDS Day 13

Zhaoxia Yu

7/27/2022

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Study Goals

Study Goals

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Introduction: Revisiting LM
- ▶ Correlated Data
 - ▶ Sources of correlation
 - ▶ The consequence of ignoring data dependence: a simulation study
- ▶ Model correlated data using linear mixed-effects model
- ▶ Examples of LME: Example 1
- ▶ The slides are based on my published work:
<https://doi.org/10.1016/j.neuron.2021.10.030>
https://yu-zhaoxia.github.io/MM_in_Neuroscience/

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Introduction

Revisiting LM

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

► Basic assumptions of LM

$$Y_i = \beta_0 + x_{i1} \times \beta_1 + \dots + x_{ip} \times \beta_p + \epsilon_i, i = 1, \dots, n$$

- $E(\epsilon_i) = 0$, which is equivalent to
 $E(Y_i|X_i) = \beta_0 + x_{i1} \times \beta_1 + \dots + x_{ip} \times \beta_p$
- $Var(\epsilon_i) = \sigma^2$. Note, this is equivalent to say
 $Var(Y_i|X_i) = \sigma^2$.
- $(\epsilon_1, \dots, \epsilon_n)$ are mutually **independent**

► Question: what if the observations are dependent?

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Correlated Data

Sources of correlation

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Clustered data: all fifth graders in the Irvine Unified School District
- ▶ Data with spatial correlation: today's highest temperatures of all cities in California
- ▶ Data with temporal correlation: hourly temperatures of Irvine within a day
- ▶ ...

Sources of correlation

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

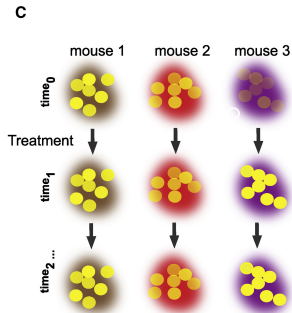
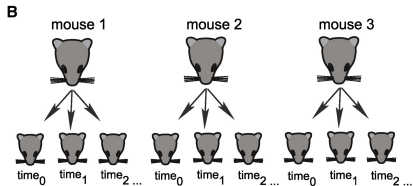
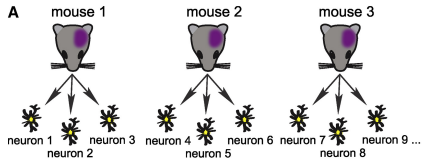
Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons



Example 1: Data

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

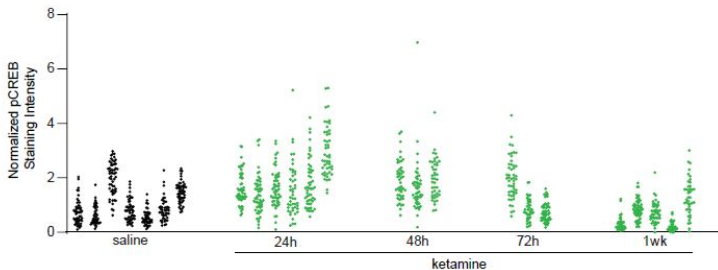


Figure 1: Normalized pCREB staining intensity from 1,200 neurons. The values in each cluster were from one animal. In total, pCREB was measured from 24 mice: saline (7 mice), 24h (6 mice), 48h (3 mice), 72h (3 mice), 1week (5 mice) after treatment.

Example 1: Data

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
Ex1 = read.csv("https://www.ics.uci.edu/~zhaoxia/Data/BeyondT
```

```
#factor the treatment IDs
```

```
Ex1$treatment_idx = as.factor(Ex1$treatment_idx)
```

```
dim(Ex1)# checking the dimensions of the dataset
```

```
## [1] 1200      3
```

```
names(Ex1)# checking the names of each column
```

```
## [1] "res"          "treatment_idx" "midx"
```

Example 1: Data

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
table(Ex1$treatment_idx)
```

```
##  
##      1      2      3      4      5  
## 357 309 139 150 245
```

```
table(Ex1$midx)
```

```
##  
##      1      2      3      4      5      6      7      8      9     10     11     12     13     14     15     16     17     18     19     20  
## 53 49 56 52 46 47 54 52 54 54 47 53 49 47 48 44 50 45 55 4
```

```
#table(Ex1$treatment, Ex1$midx)
```

Example 1: Ignore data dependence (**incorrect** analysis)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ A **common** analysis is to ignore data dependence
- ▶ Representative descriptions of inappropriate analyses
 - ▶ “ $t(28656) = 314$ with $p < 10^{-10}$ over a total of $n = 28657$ neurons pooled across six mice,”
 - ▶ “ $n = 377$ neurons from four mice, two-sided Wilcoxon signed rank test,”
 - ▶ “610 A cells, 987 B cells and 2584 C cells from 10 mice, oneway ANOVA and Kruskal–Wallis test,”
 - ▶ “two-sided paired t test, $n = 1597$ neurons from 11 animals, d.f. = 1596,”

Example 1: Ignore data dependence (**incorrect** analysis)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Example
Example 1

Adjustment for
multiple
comparisons

```
obj.lm=lm(res~treatment_idx, data=Ex1)
coef.table=summary(obj.lm)$coefficients
row.names(coef.table)=c("saline", "24h", "48h", "72h", "1wk")
print(coef.table)
```

##		Estimate	Std. Error	t value	Pr(> t)
##	saline	1.0261907	0.03997259	25.672363	4.064778e-116
##	24h	0.7828564	0.05868406	13.340189	6.040147e-38
##	48h	0.8135287	0.07550847	10.774006	6.760583e-26
##	72h	0.1605790	0.07348870	2.185084	2.907634e-02
##	1wk	-0.3604732	0.06265813	-5.753015	1.112796e-08

Example 1: Ignore data dependence (**incorrect** analysis)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
anova(obj.lm)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: res
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## treatment_idx   4  246.62   61.656   108.09 < 2.2e-16 ***
## Residuals     1195  681.65    0.570
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#The same analysis can be done by the "aov" function

► What is **wrong** with the analysis?

Example 1: Understand the dependence in data

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

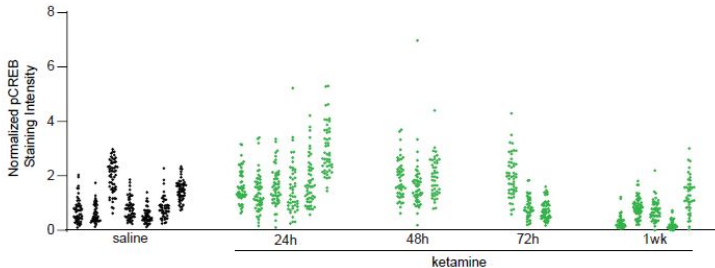
Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ From Figure 1, what can we say about observations within and between animals?
- ▶ Data are clustered in animals (mice)



Intra-class Correlation (ICC)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ We use ICC to quantify the dependence due to clustering. It is defined as

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2},$$

where

- ▶ σ_b^2 denotes the between-class variance
- ▶ σ_e^2 denotes the within-class variance
- ▶ The ICC for naturally occurring clusters is often between 0 and 1
- ▶ $ICC = 0$: the data are uncorrelated
- ▶ $ICC = 1$: all the observations in each cluster are identical

Example 1: Intra-class Correlation (ICC)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
### load the ICC library
library(ICC)
### conduct ICC analysis by organizing all the information in
icc.analysis=data.frame(n=rep(0,5), icc=rep(0,5), design=rep(0,5),
effective.n=rep(0,5), M=rep(0,5), cells=rep(0,5))
# The ICC is computed as follows (code won't show in slides)
```

Intra-class Correlation (ICC): Example 1

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Table 1. ICC, design effect, and effective sample size for the 5 groups in example 1

	Saline (7 mice)	24 h (6 mice)	48 h (3 mice)	72 h (3 mice)	1 week (5 mice)
No. cells	357	209	139	150	245
ICC	0.61	0.33	0.02	0.63	0.54
Design effect	32.0	17.7	1.8	31.8	26.8
Effective sample size	11.1	17.5	76.9	4.7	9.1

ICC and the design effect were the lowest at 48 h, when the data were relatively homogeneous across animals. At baseline and 72 h, the data were noticeably heterogeneous across animals, leading to high ICC.

Intra-class Correlation (ICC): Example 1

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ The dependency due to clustering is substantial
- ▶ the 1,200 neurons should not be treated as 1,200 independent cells
- ▶ **Design effect:** $D_{eff} = 1 + (M - 1)ICC$, where M is the average cluster size
- ▶ **Effective sample size:** $n_{eff} = n/D_{eff}$.

The consequence of ignoring data dependence: a simulation study

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ We generated 1000 data sets
 - ▶ each follows the same ICC structure of Example 1
 - ▶ data are simulated under the assumption of no difference between the five treatments/conditions
 - ▶ each data set is analyzed with the regular LM/ANOVA, without accounting for the data dependence.
 - ▶ significance level is set at $\alpha = 0.05$.
- ▶ Let's guess: What is the type I error rate (proportion of times rejecting the null hypothesis wrongly)?

Type I vs Type II errors

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ A Type I error (false positive): rejecting the null hypothesis when it is true
- ▶ A Type II error (false negative): failing to reject the null hypothesis when it is false
- ▶ The distribution of p values under the null hypothesis
- ▶ Significance level α :
- ▶ If the null hypothesis is true and $\alpha = 0.05$, what is the expected/theoretical type I error rate?

The consequence of ignoring data dependence: a simulation study

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
source("https://www.ics.uci.edu/~zhaoxia/Data/BeyondTandANOVA")
```

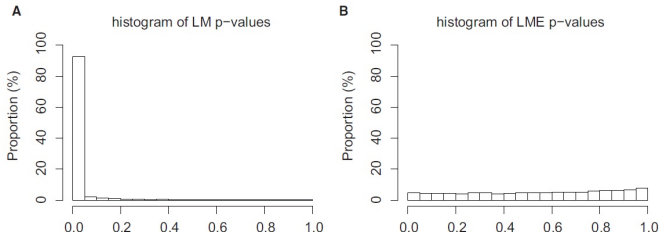


Figure 3. Histograms of p values using simulated data that assume (1) no treatment effects and (2) the same sample sizes and correlation structure with example 1

(A) Histogram of the p values from the inappropriate method (linear model) shows that ignoring the correlation structure of the data led to a surprisingly high type I error rate (90%) at significance level $\alpha = 0.05$.

(B) Histogram of the p values from LME.

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Linear-Mixed Effects Model

A Motivating Example of LME: Example 1 (the wrong model)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ The model **without** accounting for dependence:

$$Y_{ij} = \beta_0 + x_{ij,1}\beta_1 + \cdots + x_{ij,4}\beta_4 + \epsilon_{ij}, i = 1, \dots, 24; j = 1, \dots, n_i$$

where

- ▶ where n_i is the number of observations from the i th mouse and $\sum_{i=1}^{24} n_i = 1200$
- ▶ ϵ_{ij} represents the deviation in pCREB immunoreactivity of observation (cell) j in mouse i from the mean pCREB immunoreactivity of mouse i .

A Motivating Example of LME: Example 1 (the wrong model)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ The model **without** accounting for dependence:

$$Y_{ij} = \beta_0 + x_{ij,1}\beta_1 + \cdots + x_{ij,4}\beta_4 + \epsilon_{ij}, i = 1, \dots, 24; j = 1, \dots, n_i$$

where the coefficients $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ are assumed to be fixed but unknown parameters

- ▶ β_0 is the intercept (the mean of the baseline)
- ▶ β_1 is the effect (change) at 24h, compared to the baseline
- ▶ β_2 is the effect (change) at 48h, ...
- ▶ β_3 is the effect (change) at 72h, ...
- ▶ β_4 is the effect (change) at 1wk, ...

A Motivating Example of LME: Example 1 (the wrong model)

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ The model **without** accounting for dependence:

$$Y_{ij} = \beta_0 + x_{ij,1}\beta_1 + \cdots + x_{ij,4}\beta_4 + \epsilon_{ij}, i = 1, \dots, 24; j = 1, \dots, n_{ij}$$

- ▶ We use four dummy variables to denote treatment/conditions, which are the time after treatment
 - ▶ $x_{ij,1} = 1$ for 24 hours and 0 otherwise
 - ▶ $x_{ij,2} = 1$ for 48 hours and 0 otherwise
 - ▶ $x_{ij,3} = 1$ for 72 hours and 0 otherwise
 - ▶ $x_{ij,4} = 1$ for 1 week and 0 otherwise after ketamine treatments, respectively.

A Motivating Example of LME: Example 1

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Cells from the same animal share the same environment
 - ▶ We add the subject-specific effect u_i into the LM
 - ▶ (u_1, \dots, u_{24}) are assumed to be independent and identically distributed (i.i.d.) from $N(0, \sigma_b^2)$

$$Y_{ij} = \beta_0 + x_{ij,1}\beta_1 + \dots + x_{ij,4}\beta_4 + u_i + \epsilon_{ij}, i = 1, \dots, 24; j = 1, \dots, n_i$$

- ▶ This model is a mixed-effects model because
 - ▶ (u_1, \dots, u_{24}) are random-effect coefficients
 - ▶ $(\beta_0, \dots, \beta_{24})$ are fixed-effect coefficients

A Motivating Example of LME: Example 1

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Including the random effects introduces correlation between two observations from the same cluster:

- ▶ for $j \neq j'$, $\text{cov}(Y_{ij}, Y_{ij'}) =$

- ▶ for $i \neq i'$, $\text{cov}(Y_{ij}, Y_{i'j'}) =$

A Better Representation of the model

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Introduce dummy variables for the animals: let the vector $(z_{ij,1}, \dots, z_{ij,24})$ be the dummy variables for the cluster/animal memberships such that $z_{ij,k} = 1$ for $i = k$ and 0 otherwise.
- ▶ Then the LME can be rewritten to

$$Y_{ij} = \beta_0 + x_{ij,1}\beta_1 + \dots + x_{ij,4}\beta_4 + z_{ij,1}u_1 + \dots + z_{ij,24}u_{24} + \epsilon_{ij},$$

$$i = 1, \dots, 24; j = 1, \dots, n_i;$$

- ▶ A more compact form of the model presented in the previous slide:

$$Y = \mathbf{1}\beta_0 + X\beta + Zu + \epsilon$$

- ▶ Remark 1: Example 1 includes treatment effects as fixed effects. Similar to LM and GLM, when available and sensible, other covariates should be added

LME: fixed vs random effects

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ Remark 2: You might wonder why do we treat u_i 's as random, rather than fixed effects.
- ▶ Fixed-effects: typically, a fixed effect captures a parameter at the population level
- ▶ Random-effects:
 - ▶ A random effect captures cluster-specific effects (e.g., due to cells clustered in animals)
 - ▶ In many situations, the number of clusters (e.g., patients in a longitudinal study) is large. There would be too many parameters to estimate if we treat u_i 's as fixed.
 - ▶ Subject-specific effects are typically of no direct scientific interest
- ▶ Remark 3: Other random effects might also be necessary. e.g., we will discuss random slopes in an example

Fit an LME model

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ The parameters need to be estimated:

$$\beta_0, \beta_1, \dots, \beta_p, \sigma^2, \sigma_b^2$$

- ▶ Two methods to estimated parameters
 - ▶ Maximum likelihood estimation
 - ▶ REML: restricted/residual maximum likelihood estimates
 - ▶ Two main packages:
 - ▶ nlme: lme
 - ▶ lme4: lmer (doesn't provide p-value) and glmm

Fit an LME model: relevant packages

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Table 5. Selected R packages and functions for mixed-effects modeling and statistical inference

Package name	Functions related to mixed-effect modeling
<i>nlme</i>	<i>lme</i> : fit a linear mixed-effects model
<i>lme4</i>	<i>lmer</i> : fit a linear mixed-effects model <i>glmm</i> : fit a generalized linear mixed-effects model
<i>brms</i>	It can conduct Bayesian mixed-effects modeling
<i>lmerTest</i>	It can perform hypothesis testing on fixed and random effects based on models from <i>lme4::lmer</i>
<i>emmeans</i>	It can provide adjusted p values for pairwise and treatments versus control comparisons
<i>pbrktest</i>	It can perform the F-test (Kenward-Roger and Satterthwaite type) and parametric bootstrap test
<i>car</i>	<i>car::Anova</i> provides large-sample Wald test or F-test with Kenward-Roger denominator degrees of freedom
<i>sjPlot</i>	It can provide visualization and create manuscript-style tables

LME Examples: Example 1

Example 1: use nlme::lme

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
##### Linear Mixed-effects Model #####  
#use nlme::lme  
library(nlme) #load the nlme library  
# The nlme:lme function specifies the fixed effects in the fo  
# (first argument) of the function, and the random effects  
# as an optional argument (random=). The vertical bar | denot  
# the cluster is done through the animal id (midx)  
obj.lme=lme(res~treatment_idx, data= Ex1, random = ~ 1|midx)  
#summary(obj.lme)
```

Example 1: evaluate the overall significance of a factor with nlme::lme

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

#Wald F-test from an lme object

```
obj.lme=lme(res~treatment_idx, data= Ex1, random = ~1|midx)  
anova(obj.lme) #Wald F-test
```

##	numDF	denDF	F-value	p-value
## (Intercept)	1	1176	142.8589	<.0001
## treatment_idx	4	19	4.6878	0.0084

Example 1: evaluate the overall significance of a factor with nlme::lme

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjusted for
multiple
comparisons

```
#Likelihood ratio test from lme objects  
# notice the argument of the option "method"  
# which calls for using ML instead of REML  
obj.lme0.ml=lme(res~1, data= Ex1, random = ~ 1|midx, method="ML")  
obj.lme.ml=lme(res~treatment_idx, data= Ex1, random = ~ 1|midx, method="ML")  
anova(obj.lme0.ml, obj.lme.ml)
```

##	Model	df	AIC	BIC	logLik	Test	L
##	obj.lme0.ml	1	3	2281.441	2296.712	-1137.721	
##	obj.lme.ml	2	7	2272.961	2308.592	-1129.481	1 vs 2 16

Example 1: evaluate the overall significance of a factor with nlme::lme

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

#equivalently, one can conduct LRT using drop1
`drop1(obj.lme.ml, test="Chisq")`

```
## Single term deletions
##
## Model:
## res ~ treatment_idx
##           Df      AIC    LRT Pr(>Chi)
## <none>           2273.0
## treatment_idx  4 2281.4 16.48 0.002438 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Example 1: Compare LM and LME

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

► Compare LM and LME

Table 2. p values for comparing pCREB immunoreactivity at each time point (24 h, 48 h, 72 h, and 1 week) after ketamine treatment to the baseline (saline)

	Overall	24 h	48 h	72 h	1 week
Linear model (ANOVA)	1.2×10^{-78}	6.0×10^{-38}	6.8×10^{-26}	0.0291	1.1×10^{-8}
LME	0.0029	0.0049	0.0164	0.5601	0.2525

The "Overall" column corresponds to the null hypothesis of no difference among the 5 groups (example 1). The LME p values are based upon the *lme* function in the *nlme* package, in which the denominator degrees of freedom are determined by the animal grouping level (Pinheiro et al., 2007). The methods for obtaining more accurate p values with adjustments for multiple comparisons can be found in the [supplemental information](#).

Example 1: use lme4::lmer

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
library(lme4) #load the lme4 library  
obj.lmer=lmer(res ~ treatment_idx+(1|midx), data=Ex1)
```


Example 1: use lme4::lmer

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

```
library(lmerTest)
```

```
## Warning: package 'lmerTest' was built under R version 4.1.
```

```
##
```

```
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
```

```
##
```

```
##      lmer
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
##      step
```

```
obj.lmer=lmerTest::lmer(res ~ treatment_idx+(1|midx), data=Ex
```

#when ddf is not specified the F test with Satterthwaite's

Example 1: use lme4::lmer

```
summary(obj.lmer, ddf="Kenward-Roger")
```

```
## Linear mixed model fit by REML. t-tests use Kenward-Roger'
## lmerModLmerTest]
## Formula: res ~ treatment_idx + (1 | midx)
## Data: Ex1
##
## REML criterion at convergence: 2264.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.5388 -0.5761 -0.1129  0.4721  8.8601
##
## Random effects:
## Groups   Name                Variance Std.Dev.
```

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Example 1: use lme4::lmer

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

```
#likelihood ratio test
```

```
obj.lmer.ml=lme4::lmer(res ~ treatment_idx+(1|midx), data=Ex1,  
obj.lmer0.ml=lme4::lmer(res ~ 1+(1|midx), data=Ex1, REML=F),  
anova(obj.lmer0.ml, obj.lmer.ml)
```

```
## Data: Ex1
```

```
## Models:
```

```
## obj.lmer0.ml: res ~ 1 + (1 | midx)
```

```
## obj.lmer.ml: res ~ treatment_idx + (1 | midx)
```

```
##           npar      AIC      BIC  logLik deviance Chisq Df
```

```
## obj.lmer0.ml      3 2281.4 2296.7 -1137.7   2275.4
```

```
## obj.lmer.ml       7 2273.0 2308.6 -1129.5   2259.0 16.48  4
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

Adjustment for multiple comparisons

Adjustment for multiple comparisons

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

- ▶ What is the issue of multiple comparisons
- ▶ Suppose the null hypothesis is true. What is the type I error rate if $\sigma = 0.05$ if we conduct a test?
- ▶ Suppose M null hypotheses are true. What is the probability of making at least one mistake if we use $\sigma = 0.05$ for each test?
 - ▶ This probability is called the family wise error rate
- ▶ Mathematical derivation or simulation will show that the family wise error rate is much greater than 0.05
- ▶ Procedures have been developed to control for the family wise error rate and other types error rate (false discovery rate)

Adjustment for multiple comparisons

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed
Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

##	contrast	estimate	SE	df	t.ratio	p.value
##	1 - 2	-0.8194	0.289	19.0	-2.835	0.0704
##	1 - 3	-0.8429	0.359	19.1	-2.349	0.1727
##	1 - 4	-0.1898	0.359	19.0	-0.529	0.9832
##	1 - 5	0.3200	0.304	19.0	1.051	0.8283
##	2 - 3	-0.0235	0.368	19.0	-0.064	1.0000
##	2 - 4	0.6296	0.367	19.0	1.713	0.4496
##	2 - 5	1.1394	0.315	19.0	3.621	0.0138
##	3 - 4	0.6531	0.425	19.0	1.538	0.5517
##	3 - 5	1.1629	0.380	19.1	3.062	0.0447
##	4 - 5	0.5098	0.380	19.0	1.343	0.6690
##						

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of

Adjustment for multiple comparisons

UC Irvine
ISI-BUDS Day 13

Zhaoxia Yu

Study Goals

Introduction

Correlated Data

Linear-Mixed

Effects Model

LME Examples:
Example 1

Adjustment for
multiple
comparisons

the default method of degrees of freedom is Kenward-Roger's

```
contrast(emmeans(obj.lmer, specs="treatment_idx"), "trt.vs.ct
```

```
## contrast estimate SE df t.ratio p.value
## 2 - 1 0.819 0.289 19.0 2.835 0.0364
## 3 - 1 0.843 0.359 19.1 2.349 0.0965
## 4 - 1 0.190 0.359 19.0 0.529 0.9219
## 5 - 1 -0.320 0.304 19.0 -1.051 0.6613
##
## Degrees-of-freedom method: kenward-roger
## P value adjustment: dunnettx method for 4 tests
```