# lmerTest R-package for automated mixed ANOVA modelling

Alexandra Kuznetsova [1] Rune H.B. Christensen [1] Per Bruun Brockhoff [1]

[1]DTU Compute, Statistical section, Technical University of Denmark

August 16, 2015

# Outline

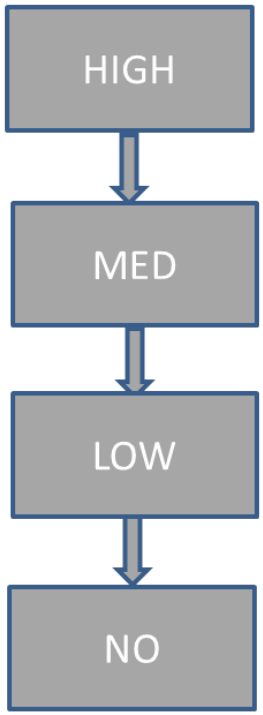
- Simple Mixed Effects models
- Automated analysis/elimination of mixed effects models using lmerTest
- Example 1. Sensory profile with multi-way product structure
- Example 2. External Preference Mapping with Consumer background

# Linear mixed effects models: complexity for a user

lme4

lmerTest

SensMixed


# Use ImerTest to hande even more complex settings

- Unbalanced sensory profile data (e.g. missing observations)
- Incomplete consumer preference data
- 2- (or higher)way product structure in sensory profile data
- 2- (or higher)way product structure in consumer preference data (Conjoint)
- Extending Conjoint to include Consumer background/design factors/covariates
- Complex blocking, product replication, product batch structures in as well sensory as consumer preference data
- A mixed model approach for performing external preference mapping
- Extending mixed model external preference mapping to include product and consumer background/design factors/covariates

# The lmerTest package - functions

- `step` (automated analysis of both random and fixed parts - finds the best simplest model)
- `rand` (analysis of the random part of a mixed model, LRT (likelihood ratio test))
- `anova` (Type I, II and III ANOVA tables with Satterthwaite's approximation to degrees of freedom)
- `summary` ( *t*-tests for fixed effects with Satterthwaite's approximation to degrees of freedom)
- `lsmeans`, `difflsmeans` (least squares means and differences of least square means with confidence intervals) - post-hoc analysis
- `plot` - plots the post-hoc analysis

# Types I - Type III ANOVA tables

Type I ANOVA

- produces sequential sums of squares
- the hypotheses are functions of cell counts
- depends on the order the effects are entered in the model

Type II ANOVA

- the hypotheses are functions of cell counts
- the hypotheses do NOT depend on the order the effects are entered

Type III ANOVA

- the hypotheses are NOT functions of cell counts
- the hypotheses do NOT depend on the order the effects are entered

Types I - III contrast matrices (SAS Technical Report R-101 et. al (1978)) are implemnted in **lmerTest**

In balanced situations all types produce the same output

# Example 1, Sensory profile with multi-way product structure

**TVbo data** (Bang and Olufsen, Sensometrics, As, 2006) *Sensory profile with multi-way product structure*

12 Products

- 3 TV sets
- 4 Pictures

2 replicates

8 Assessors

15 Sensory attributes

# Example 2. Attach TVbo data

Tell **R** to use lmerTest and to use `TVbo` data

```
> library(lmerTest)
```

```
> data(TVbo)
```

# Example 1. TVbo data

Summarize the data

```
> str(TVbo)
```

```
'data.frame':         192 obs. of  5 variables:
 $ Assessor          : Factor w/ 8 levels "1","2","3","4",..: 1 1 1
 $ TVset             : Factor w/ 3 levels "TV1","TV2","TV3": 3 2 1
 $ Repeat            : Factor w/ 2 levels "0","1": 1 1 1 2 2 2 1 1
 $ Picture           : Factor w/ 4 levels "1","2","3","4": 1 1 1 1
 $ Coloursaturation: num  10.4 9.9 7 9.8 10.6 7.5 7.1 9.9 5 10 ..
```

- Response variable: `Sharpnessofmovement`
- Fixed effects: `TVset`, `Picture`
- Random effects: `Assessor`, `Repeat`

# anova, lme4

```
> tv <- lme4::lmer(Sharpnessofmovement ~ TVset*Picture +
+                  (1|Assessor) +(1|Assessor:TVset) +
+                  (1|Assessor:Picture), data = TVbo)
> anova(tv)

Analysis of Variance Table
              Df Sum Sq Mean Sq F value
TVset          2  1.765  0.8825  0.2437
Picture        3 51.857 17.2857  4.7735
TVset:Picture  6 90.767 15.1279  4.1777
```

# anova, lmerTest

```
> library(lmerTest)
> tv <- lmer(Sharpnessofmovement ~ TVset*Picture +
+           (1|Assessor) +(1|Assessor:TVset) +
+           (1|Assessor:Picture), data = TVbo)
> anova(tv)

Analysis of Variance Table of type III  with  Satterthwaite
approximation for degrees of freedom
             Sum Sq Mean Sq NumDF DenDF F.value    Pr(>F)
TVset         1.765  0.8825     2    14  0.2437 0.7869818
Picture      51.857 17.2857     3    21  4.7735 0.0108785 *
TVset:Picture 90.767 15.1279     6   138  4.1777 0.0006845 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Example 1. Specification of the full model. Step 1

Specification of the effects in **R**:

- Fixed effects: `TVset*Picture` will give two main effects and an interaction
- Random effects: `(1|Assessor)`, `1|Assessor:TVset)`, `(1|Assessor:TVset:Picture)` `1|Repeat)`, `1|Repeat:TVset)`, `1|Repeat:Picture)`, `1|Repeat:TVset:Picture)`

Specify the "full" model using `lmer` syntaxis

```
> M <- lmer(Sharpnessofmovement ~ TVset*Picture +
+                 (1|Assessor:TVset) + (1|Assessor:Picture) +
+                 (1|Assessor:Picture:TVset) + (1|Repeat) +
+                 (1|Repeat:Picture) +
+                 (1|Repeat:TVset) + (1|Repeat:TVset:Picture) +
+                 (1|Assessor) , data=TVbo)
```

# Example 1. the step function

Run the step function and plot the post-hoc: performs Step 2 and Step 3 of the automated analysis

```
> s_TV <- step(M)
```

```
> plot(s_TV)
```

# Example 1. Elimination process of random and fixed effects. Step 2 and Step 3

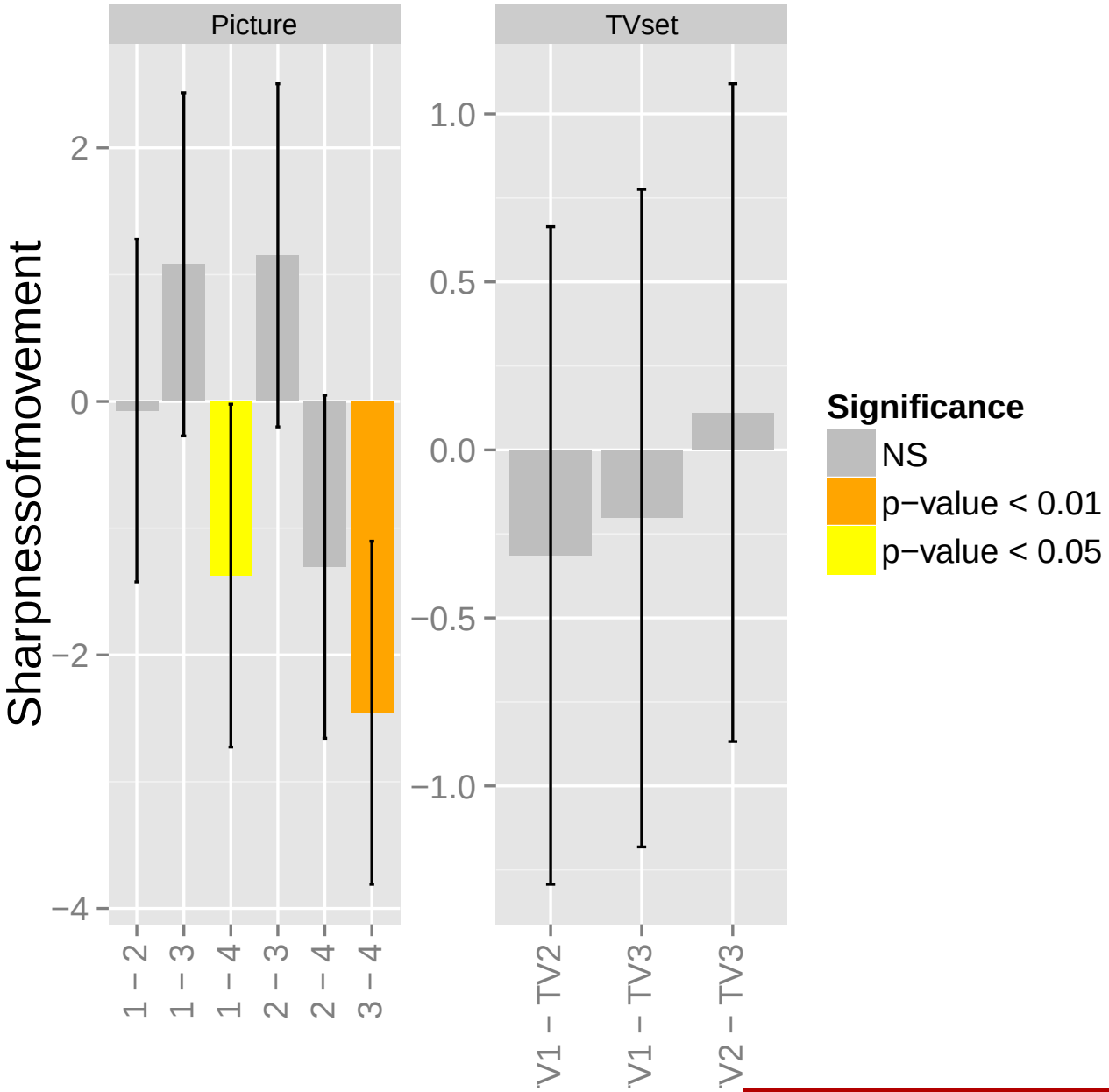
| | Chi.sq | Chi.DF | elim.num | p.value |
|---|---|---|---|---|
| Assessor:Picture:TVset | 0.00 | 1 | 1 | 1.000 |
| Repeat:Picture | 0.00 | 1 | 2 | 1.000 |
| Repeat | 0.00 | 1 | 3 | 1.000 |
| Repeat:TVset | 0.00 | 1 | 4 | 1.000 |
| Repeat:TVset:Picture | 0.00 | 1 | 5 | 1.000 |
| Assessor:TVset | 2.79 | 1 | kept | 0.095 |
| Assessor:Picture | 12.35 | 1 | kept | <0.001 |
| Assessor | 7.47 | 1 | kept | 0.006 |

Table: Automated elimination process of random effects of model M

| | DenDF | F.value | elim.num | Pr(>F) |
|---|---|---|---|---|
| TVset | 14 | 0.24 | kept | 0.787 |
| Picture | 21 | 4.77 | kept | 0.011 |
| TVset:Picture | 138 | 4.18 | kept | <0.001 |

Table: Automated elimination process of fixed effects of model M

# Example 1. Plot post-hoc



# Example 2, External Preference Mapping with consumer background

`carrots` data (Brockhoff, DTU course 02429) *External Preference Mapping with consumer background*

12 carrots varieties

103 Consumers

Consumer background info

- Homesize (2)

Product background info

- Sensory profile PC scores 1 and 2 (`sens1`, `sens2`)

# Example 2. carrots data

Summarize the data

*str(carrots)*

```
'data.frame':        1236 obs. of  5 variables:
 $ Consumer  : Factor w/ 103 levels "168","169","171",..: 1 1 1 1
 $ Homesize  : Factor w/ 2 levels "1","3": 2 2 2 2 2 2 2 2 2 2 ..
 $ Preference: int  4 5 4 7 5 6 5 6 5 6 ...
 $ sens1     : num  6.683 6.441 -1.948 0.481 2.181 ...
 $ sens2     : num  -2.54 3.54 -4.29 -1.33 -4.67 ...
```

- Response variable: `Preference`
- Fixed effects: `sens1`, `sens2`, `Homesize`
- Random effects: `Consumer`, `product` (*For External Preference mapping we consider `product` as random effect!*)

# Example 2. Specification of the model

Specification of the effects in **R**:

- Fixed effects: `sens1 sens2 Homesize` + interactions
- Random effects: `(1 + sens1 + sens2 | product), (1 | Consumer)`

Specify the "full" model using `lmer` syntaxis

```
M <- lmer(Preference ~ sens2*sens1*Homesize +
          (1 | product) + (1 + sens1 + sens2  |Consumer),
          data = carrots)
```

$$(b_0, b_1, b_2) \sim N(0, \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{pmatrix}), \quad c \sim N(0, \sigma^2_{\texttt{product}}), \quad \epsilon_{ijk} \sim N(0, \sigma^2) \tag{1}$$

# Example 2. the step function

Run the step function: performs Step 2 and Step 3 of the automated analysis

```
t <- step(M)
```

# Example 2. Elimination process of random effects. Step 2

| | Chi.sq | Chi.DF | elim.num | p.value |
|---|---|---|---|---|
| sens1:Consumer | 2.08 | 3 | 1 | 0.556 |
| product | 17.23 | 1 | kept | $<$0.001 |
| sens2:Consumer | 8.06 | 2 | kept | 0.018 |

Table: Automated elimination process of random effects of model M

# Example 2. Elimination process of fixed effects. Step 3

| | DenDF | F.value | elim.num | Pr(>F) |
|---|---|---|---|---|
| sens2:sens1:Homesize | 1015 | 1.33 | 1 | 0.249 |
| sens2:sens1 | 7 | 0.11 | 2 | 0.744 |
| sens1:Homesize | 1016 | 0.18 | 3 | 0.673 |
| sens1 | 8 | 0.52 | 4 | 0.489 |
| sens2:Homesize | 101 | 1.04 | 5 | 0.311 |
| sens2 | 12 | 17.48 | kept | 0.001 |
| Homesize | 100 | 5.63 | kept | 0.020 |

Table: Automated elimination process of fixed effects of model M

# Example 2. summary

Reduced model:

$$y_{ijkl} = b_{0j} + \beta_0 + \alpha_k + (b_{2j} + \beta_2)\texttt{sens2}_{ij} + c_k + \epsilon_{ijkl} \qquad (2)$$

$$(b_0, b_1, b_2) \sim N(0, \begin{pmatrix} \sigma_0^2 & \sigma_{02} \\ \sigma_{02} & \sigma_2^2 \end{pmatrix}), \quad c \sim N(0, \sigma_c^2), \quad \epsilon_{ijk} \sim N(0, \sigma^2)$$

```
M_red <- lmer(Preference ~ sens2 + Homesize +
          (1 | product) + (1 + sens2  |Consumer),
          data = carrots)
```

# Example 2. summary

```
summary(M_red)
```

| | Estimate | Std. Error | df | t value | Pr(>\|t\|) |
|---|---|---|---|---|---|
| (Intercept) | 4.91 | 0.09 | 39 | 56.35 | <0.001 |
| sens2 | 0.07 | 0.02 | 12 | 4.18 | 0.001 |
| Homesize3 | -0.25 | 0.11 | 100 | -2.37 | 0.020 |

Table: Automated elimination process of random effects of model M

Consumers prefer products charactarized by the `sens2` (more sweet products)

# To remember!

- Random effects: Consumers, Assessors, Replicates/Sessions
- For the random effects remember to put interaction effects: between random and between random and fixed effects
- For the fixed effects: via "*" you may specify all possible main and interaction effects
- Remember to check whether the variable is factor or not via `str()` function. If not and it should be then use (as.factor()) function , like e.g.:

```
TVbo$Assessor <- as.factor(TVbo$Assessor)
```

# Getting started with lmerTest

1. Install lmerTest from CRAN (by e.g. *Install* button in RStudio)

2. Tell **R** that you want to use it with `library(lmerTest)`
3. Take a look at the **lmerTest-manual.pdf**
4. Run the examples from lecture at lecture_lmerTest_Rcode.R
5. Do the exercises (`lecture_lmerTest_Rcode.R`)

# Bibliography I



Tests of Hypotheses in Fixed-Effects Linear Models *Copyright ©1978 by SAS Institute Inc., Cary, NC, USA*


J.H. Goodnight. General Linear Model Procedure *S.A.S. Institute, Inc.*


Tormod Naes, Per B. Brockhoff and Oliver Tomic. Statistics for sensory and consumer science *©2010 John Wiley and Sons Ltd*


Alain F. Zuur, Elena N. Ieno, Neil Walker, Anatoly A. Saveliev and Graham M. Smith. Mixed effects models and extensions in ecology with R *Statistics for Biology and Health 2009, DOI: 10.1007/978-0-387-87458-6*


G. Bruce Schaalje, Justin B. McBride, Gilbert W. Fellingham. Adequacy of approximations to distributions of test Statistics in complex mixed linear models *©American Statistical Association and the International Biometric Society Journal of Agricultural, Biological, and Environmental Statistics, Volume 7, Number 4, Pages 512-524*

# Bibliography II


Harry T. Lawless, Hildegarde Heymann. Sensory Evaluation of Food
*©2010 Springer Science+Business Media, LLC*