lmerTest R-package for automated mixed ANOVA modelling

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

¹DTU 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
- ImerTest
- SensMixed

FLOWCHART:

- HIGH
- MED
- LOW
- NO

R-package for automated mixed ANOVA modelling
Use lmerTest to handle 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 implemented 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 \textbf{R} to use \textit{lmerTest} and to use TVbo data

\begin{verbatim}
> library(lmerTest)
> data(TVbo)
\end{verbatim}
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 1 1 1 1 1 ...
  $ TVset : Factor w/ 3 levels "TV1","TV2","TV3": 3 2 1 3 2 1 3 2 1 3 ...
  $ Repeat : Factor w/ 2 levels "0","1": 1 1 1 2 2 2 1 1 1 2 ...
  $ Picture: Factor w/ 4 levels "1","2","3","4": 1 1 1 1 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
> tv <- lme4::lmer(Sharpnessofmovement ~ TVset*Picture +
+   (1|Assessor) + (1|Assessor:TVset) +
+   (1|Assessor:Picture), data = TVbo)
> anova(tv)

Analysis of Variance Table

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVset</td>
<td>2</td>
<td>1.765</td>
<td>0.8825</td>
</tr>
<tr>
<td>Picture</td>
<td>3</td>
<td>51.857</td>
<td>17.2857</td>
</tr>
<tr>
<td>TVset:Picture</td>
<td>6</td>
<td>90.767</td>
<td>15.1279</td>
</tr>
</tbody>
</table>
> 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

<table>
<thead>
<tr>
<th></th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>NumDF</th>
<th>DenDF</th>
<th>F.value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVset</td>
<td>1.765</td>
<td>0.8825</td>
<td>2</td>
<td>14</td>
<td>0.2437</td>
<td>0.7869818</td>
</tr>
<tr>
<td>Picture</td>
<td>51.857</td>
<td>17.2857</td>
<td>3</td>
<td>21</td>
<td>4.7735</td>
<td>0.0108785 *</td>
</tr>
<tr>
<td>TVset:Picture</td>
<td>90.767</td>
<td>15.1279</td>
<td>6</td>
<td>138</td>
<td>4.1777</td>
<td>0.0006845 ***</td>
</tr>
</tbody>
</table>

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Specify the "full" model using lmer syntax

\[
M \leftarrow \text{lmer(Sharpnessofmovement} \sim \text{TVset*Picture} +
+ (1|\text{Assessor:TVset}) + (1|\text{Assessor:Picture}) +
+ (1|\text{Assessor:Picture:TVset}) + (1|\text{Repeat}) +
+ (1|\text{Repeat:Picture}) +
+ (1|\text{Repeat:TVset}) + (1|\text{Repeat:TVset:Picture}) +
+ (1|\text{Assessor}), \text{ data=TVbo})
\]
Example 1. the step function

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

```r
> s_TV <- step(M)
> plot(s_TV)
```
### Example 1. Elimination process of random and fixed effects. Step 2 and Step 3

<table>
<thead>
<tr>
<th></th>
<th>Chi.sq</th>
<th>Chi.DF</th>
<th>elim.num</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessor:Picture:TVset</td>
<td>0.00</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>Repeat:Picture</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
<td>1.000</td>
</tr>
<tr>
<td>Repeat</td>
<td>0.00</td>
<td>1</td>
<td>3</td>
<td>1.000</td>
</tr>
<tr>
<td>Repeat:TVset</td>
<td>0.00</td>
<td>1</td>
<td>4</td>
<td>1.000</td>
</tr>
<tr>
<td>Repeat:TVset:Picture</td>
<td>0.00</td>
<td>1</td>
<td>5</td>
<td>1.000</td>
</tr>
<tr>
<td>Assessor:TVset</td>
<td>2.79</td>
<td>1</td>
<td>kept</td>
<td>0.095</td>
</tr>
<tr>
<td>Assessor:Picture</td>
<td>12.35</td>
<td>1</td>
<td>kept</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Assessor</td>
<td>7.47</td>
<td>1</td>
<td>kept</td>
<td>0.006</td>
</tr>
</tbody>
</table>

**Table:** Automated elimination process of random effects of model M

<table>
<thead>
<tr>
<th></th>
<th>DenDF</th>
<th>F.value</th>
<th>elim.num</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVset</td>
<td>14</td>
<td>0.24</td>
<td>kept</td>
<td>0.787</td>
</tr>
<tr>
<td>Picture</td>
<td>21</td>
<td>4.77</td>
<td>kept</td>
<td>0.011</td>
</tr>
<tr>
<td>TVset:Picture</td>
<td>138</td>
<td>4.18</td>
<td>kept</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table:** Automated elimination process of fixed effects of model M

*ImmerTest R-package for automated mixed ANOVA modelling*
Example 1. Plot post-hoc

```
Sharpness of movement
1 - 2
1 - 3
1 - 4
2 - 3
2 - 4
3 - 4
TV1 - TV2
TV1 - TV3
TV2 - TV3
Significance
NS
p-value < 0.01
p-value < 0.05
```

lmerTest R-package for automated mixed ANOVA modelling
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

\textit{str(carrots)}

'data.frame': 1236 obs. of 5 variables:

$ Consumer : Factor w/ 103 levels "168","169","171",..: 1 1 1 1 1 1 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 (\textit{For External Preference mapping we consider product as random effect!})

\textit{lmerTest R-package for automated mixed ANOVA modelling}
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 syntax:

\[
M \leftarrow \text{lmer}(\text{Preference} \sim \text{sens2*sens1*Homesize} + \\
(1 \mid \text{product}) + (1 + \text{sens1} + \text{sens2}} \mid \text{Consumer}), \\
data = \text{carrots})
\]

\[
(b_0, b_1, b_2) \sim N(0, \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{pmatrix}), \quad c \sim N(0, \sigma_{\text{product}}^2), \quad \epsilon_{ijk} \sim N(0, \sigma^2)
\]

(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

<table>
<thead>
<tr>
<th></th>
<th>Chi.sq</th>
<th>Chi.DF</th>
<th>elim.num</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>sens1:Consumer</td>
<td>2.08</td>
<td>3</td>
<td>1</td>
<td>0.556</td>
</tr>
<tr>
<td>product</td>
<td>17.23</td>
<td>1</td>
<td>kept</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>sens2:Consumer</td>
<td>8.06</td>
<td>2</td>
<td>kept</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Table: Automated elimination process of random effects of model M

ImeRTest R-package for automated mixed ANOVA modelling
Example 2. Elimination process of fixed effects. Step 3

<table>
<thead>
<tr>
<th></th>
<th>DenDF</th>
<th>F.value</th>
<th>elim.num</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sens2:sens1:Homesize</td>
<td>1015</td>
<td>1.33</td>
<td>1</td>
<td>0.249</td>
</tr>
<tr>
<td>sens2:sens1</td>
<td>7</td>
<td>0.11</td>
<td>2</td>
<td>0.744</td>
</tr>
<tr>
<td>sens1:Homesize</td>
<td>1016</td>
<td>0.18</td>
<td>3</td>
<td>0.673</td>
</tr>
<tr>
<td>sens1</td>
<td>8</td>
<td>0.52</td>
<td>4</td>
<td>0.489</td>
</tr>
<tr>
<td>sens2:Homesize</td>
<td>101</td>
<td>1.04</td>
<td>5</td>
<td>0.311</td>
</tr>
<tr>
<td>sens2</td>
<td>12</td>
<td>17.48</td>
<td>kept</td>
<td>0.001</td>
</tr>
<tr>
<td>Homesize</td>
<td>100</td>
<td>5.63</td>
<td>kept</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Table: Automated elimination process of fixed effects of model M
Example 2. summary

Reduced model:

\[ y_{ijkl} = b_{0j} + \beta_0 + \alpha_k + (b_{2j} + \beta_2)\text{sens}_2_{ij} + c_k + \epsilon_{ijkl} \]  

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

\[ M_{\text{red}} \leftarrow \text{lmer}(\text{Preference} \sim \text{sens2} + \text{Homesize} + \\
\text{(1 | product)} + \text{(1 + sens2 |Consumer)}, \text{data = carrots}) \]
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 characterized 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.:

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

**lmerTest R-package for automated mixed ANOVA modelling**
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)
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


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