Food Quality and Preference 40 (2015) 31–38

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Automated mixed ANOVA modeling of sensory and consumer data

Alexandra Kuznetsova [a,*], Rune H.B. Christensen [a], Cecile Bavay [b], Per Bruun Brockhoff [a]

[a] DTU Compute, Statistical section, Technical University of Denmark, Richard Petersens Plads, Building 324, DK-2800 Kongens Lyngby, Denmark
[b] Groupe ESA, UPSP GRAPPE 55, rue Rabelais BP30748, 49007 Angers, Cedex 01, France

ARTICLE INFO

Article history:
Received 30 January 2014
Received in revised form 24 June 2014
Accepted 15 August 2014
Available online 27 August 2014

Keywords:
Mixed-effects models
Automated model building
R program
Conjoint
Consumer preference
ANOVA

ABSTRACT

Mixed effects models have become increasingly prominent in sensory and consumer science. Still applying such models may be challenging for a sensory practitioner due the challenges associated with the choosing the random effects, selecting an appropriate model, interpreting the results. In this paper we introduce an approach for automated mixed ANOVA/ANCOVA modeling together with the open source **R** package lmerTest developed by the authors that can perform automated complex mixed-effects modeling. The package can in an automated way investigate and incorporate the necessary random-effects by sequentially removing non-significant random terms in the mixed model, and similarly test and remove fixed effects. Tables and figures provide an overview of the structure and present post hoc analysis. With this approach, complex error structures can be investigated, identified and incorporated whenever necessary. The package provides type-3 ANOVA output with degrees of freedom corrected *F*-tests for fixed-effects, which makes the package unique in open source implementations of mixed models. The approach together with the user-friendliness of the package allow to analyze a broad range of mixed effects models in a fast and efficient way. The benefits of the approach and the package are illustrated on four data sets coming from consumer/sensory studies.

Introduction

Mixed models are used extensively for analyzing sensory and consumer data. Sensory quantitative descriptive analysis (QDA) data are typically analyzed attribute by attribute using analysis of variance (ANOVA) techniques to extract the important attribute-wise product difference information (Lawless & Heymann, 2010). The proper analysis will typically evaluate the statistical significance of product differences by using the assessor-by-product interaction as error structure (Lawless & Heymann, 2010). This is, what generally in statistics is called a mixed model as both fixed-effects (product differences) as random-effects (assessor differences and assessor-by-product interactions) are present in the modeling and analysis approach. Incorporating random consumer effects for the analysis of e.g. consumer preference data or data from conjoint experiments is on one hand necessary to obtain the proper conclusions from such data and on the other hand similarly leads to mixed models. In the simplest of cases a mixed-effects model (mixed model) analysis can be handled by simple averaging combined with the use of the proper error term coming from a simple ANOVA decomposition of the data. Two often occurring examples of this situation are:

1. Complete consumer preference data with just a single product factor, that is, just a collection of different products coded in a single variable (as opposed to a multifactorial setting), calling for a two-way (block) ANOVA, where the error term is simply the residual error.
2. Complete sensory profile data similarly with just a single product factor, calling for either a 2-way or 3-way ANOVA mixed model depending on the presence of a blocking (replication) factor such as session or product batch. And hence calling for using either the panelist-by-product mean square as the error term or a combination of this with blocking-by-product (Næs, Brockhoff, & Tomic, 2010).

These cases are exactly those covered by the open source software package PanelCheck (Mat & Ås, 2008). However, these simple approaches of analysis have their limitations. With missing values or with more complex study designs one would often benefit from a more detailed analysis. The PanelCheck tool can still be a valuable tool in that using missing values imputation and considering all products together it will in most cases be able to provide some relevant ANOVA information for the situation at hand, and by the

* Corresponding author.
E-mail address: alku@dtu.dk (A. Kuznetsova).

http://dx.doi.org/10.1016/j.foodqual.2014.08.004
0950-3293/

way the by-attribute ANOVA is only a small part of what PanelCheck has to offer. The more detailed univariate analysis of variance provided in this paper becomes relevant in the sensory context in the following example list of situations.

- Unbalanced sensory profile data (for example due to missing observations).
- Incomplete consumer preference data.
- 2-(or higher) way product structure in sensory profile data. (Beck, Jensen, Bjoern, & Kidmose, 2014).
- 2-(or higher) way product structure in consumer preference data (Conjoint) (Jaeger, Mielby, Heymann, Jia, & Frost, 2013).
- Extending Conjoint to include consumer background/design variables or factors/covariates.
- Complex blocking, product replication, product batch structures in as well sensory as consumer preference data.
- Extending external preference mapping to include product and consumer background/design factors/covariates.

Even though commercial and open source software exist for the relevant mixed modeling in such situations, it maintains to be a challenge to apply mixed models for a sensory practitioner. The questions arise as to which model to consider, which variables to include and what the interpretations of the results are. We have developed an **R** package named lmerTest (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2013) that will help answer these questions. Moreover, it will do almost all the model selection work for the practitioner and present the results of the model selection process together with post hoc analyses in a nice and user-friendly way.

The paper is organized such that first, in Section 2, we define basic mixed-effects models, then in Section 3 we discuss why mixed-effects models are important for sensory profile and consumer data and introduce aspects of the model building approaches. In Section 4 we present the lmerTest package and the details of automated mixed modeling, and in Section 5 give four examples showing the usefulness of the automated mixed modeling together with the package in the situations mentioned above. The paper ends with discussions and conclusions in Section 6.

Theory: basic mixed models

Let us consider a simple example of a sensory experiment where we have I assessors, J products and R replicates. This type of data can be described by a mixed ANOVA for replicated two-way data, where as effects we have factors A (assessor) and B (product). A reasonable model can then be written as

$$y_{ijr} = \mu + a_i + \beta_j + d_{ij} + \epsilon_{ijr} \quad (1)$$

where a_i and β_j are main effects for factors A and B and d_{ij} is the effect corresponding to interaction between A and B. If we consider the effects of factor A random, then this implies that the effects a_i (assessor) and d_{ij} (interaction between assessor and product) are random:

$$a_i \sim N(0, \sigma^2_{assessor})$$
$$d_{ij} \sim N(0, \sigma^2_{assessor \times product})$$
$$\epsilon_{ijr} \sim N(0, \sigma^2_{error})$$

where all the random-effects are independent of each other.

A commonly used test statistic for fixed effects hypotheses is F-test statistic. For complex mixed models, e.g. with unbalanced data sets, the F-test statistic will generally not be exactly F-distributed. The common approach is to assume that the test statistic approximately follows an F-distribution and calculate an approximate number of denominator degrees of freedom. Two degree of freedom approximations well-known in the statistical literature are Satterthwaite's (Satterthwaite, 1946) and Kenward-Roger's approximations (Kenward & Roger, 1997). Both of these are implemented in the lmerTest package.

Mixed effects model building

When building a mixed effects model a number of questions arise such as which effects to consider as random, which ones as fixed and which effects to include at all.

Considering the assessor effects random is generally regarded by the sensory field as the proper approach (Lawless & Heymann, 2010). The reason to consider them random is based on the interest in the population of assessors rather than to specific assessors. This means that we want to know the variation among assessors rather than estimates of effects of each assessor and to be able to properly account for that. So in model (1) we are interested in estimating $\sigma^2_{assessor}$ and $\sigma^2_{assessor \times product}$. Moreover, Næs and Langsrud (1996) showed that in situations with interactions between assessors and products, considering assessors as fixed-effects may lead to a conclusion that differences between products are larger than they really are. Therefore the assumption of random assessor effects is usually the most appropriate. For consumer tests, following the same arguments, treating the consumer effect as random is also the most natural. The same goes for the replication/session effect if present.

Having decided on which effects to include as random and which as fixed, the question arises as to which approach of model selection to use. Model selection in general, and selection of regression and ANOVA type models in particular, are controversial topics with many highly opinionated papers in the statistical literature (Jiang, Rao, Gu, & Nguyen, 2008; Ibrahim, Zhu, Garcia, & Guo, 2011; Fan & Li, 2012; Scheepers, Tily, Levy, & Barr, 2013; Peng & Lu, 2012). A particular challenge for model selection of mixed-effects models is how to handle the two types of effects; random-effects and fixed-effects. If the random effects are not well chosen, this will affect the estimates and the hypothesis tests of the fixed-effects. Vice versa, variation in the response variable not modeled in terms of fixed-effects can partly end up in the random effects. In this paper we take a rather heuristic, but practical data-driven approach to the problem and consider the backwards selection approach based on step-wise deletion of model terms with high p-values (Diggle, Heagerty, Liang, & Zeger, 2002; Zuur, Ieno, Walker, Saveliev, & Smith, 2009). In this approach the largest possible model is considered at the first place which includes all possible fixed and random effects that are at least in principle supported by the design. Then the simplification of the random structure is performed. Finally the fixed effects are also incrementally eliminated following the principle of marginality, that is the effects that are contained in any other effects are retained in the model when the effects that they are contained in are found to be significant according to the specified Type 1 level. Lower order interactions are contained in the higher order interactions, so when the higher order interactions are found to be significant, the lower order interactions are kept in the model. The marginality principle is used to enhance interpretability of the various fixed effects and tests thereof. The random effects are part of the overall covariance structure and there is no tradition nor reason for applying a similar principle for these effects. The most important random effects should be included to model the variance structures as good as possible. Even it could be quite meaningful to allow for the pooling effect that would be the consequence of eliminating a random main effect while keeping a random