

Contents

Bibliography 1

- R command and R object - Typewriter Font
 - R Package name - Italics
 - Selected Acronyms and Proper Nouns - Italics

This chapter is broken into two parts. The first part is a review of diagnostics methods for linear models, intended to acquaint the reader with the subject, and also to provide a basis for material covered in the second part. Particular attention is drawn to graphical methods.

The second part of the chapter looks at diagnostics techniques for LME models, firstly covering the theory, then proceeding to a discussion on implementing these using `R` code. While a substantial body of work has been developed in this area, there are still areas worth exploring. In particular the development of graphical techniques pertinent to LME models should be looked at.

- *The previous Section (Section 4) is a literary review of residual diagnostics and influence procedures for Linear Mixed Effects Models, drawing heavily on Schabenberger and Zewotir.*
- *Section 4 begins with an introduction to key topics in residual diagnostics, such as influence, leverage, outliers and Cook's distance. Other concepts such as DF-FITS and DFBETAs will be introduced briefly, mostly to explain why they are not particularly useful for the Method Comparison context, and therefore are not elaborated upon.*
- *In brief, Variable Selection is not applicable to Method Comparison Studies, in the commonly used context. Testing a rather simplistic specified model against one with more random effects terms is tractable, but this research question is of secondary importance.*

Appendix to Section 4

As an appendix to section 4, an appraisal of the current state of development (or lack thereof) for current implemenations for LME models, particularly for `nlme` and `lme4` fitted models.

Crucially, a review of internet resources indicates that almost all of the progress in this regard has been done for `lme4` fitted models, specifically the *Influence.ME* R package. (Nieuwenhuis et 2012)

Conversely there is very little for `nlme` models. To delve into this mor, one would immediately investigate the current development workflow for both packages.

As an aside, Douglas Bates was arguably the most prominent R developer working in the LME area. However Bates has now prioritised the development of LME models in another computing environment , i.e Julia.

The `nlme` package

With regards to `nlme`, the torch has been passed to Galecki Galecki & Burzykowski (UMich. and Hasselt respecitely). Galecki & Burzykowski published *Linear Mixed Effects Models using R*. Also, the accompanying R package, nlmeU package is under current development, with a version being released XXXX.

The `lme4` package

The `lme4` package is also under active development, under the leadership of Ben Bolker (McMaster University). According to CRAN, the LME4 package, fits linear and generalized linear mixed-effects models

The models and their components are represented using S4 classes and methods. The core computational algorithms are implemented using the Eigen C++ library for numerical linear algebra and RcppEigen "glue".
(CRAN)

The key issue is that `nlme` allows for the particular specification of Roy’s Model, specifically direct specification of the VC matrices for within subject and between subject residuals. The `lme4` package does not allow for this. To advance the ideas that emanate from Roys’ paper, one is required to use the `nlme` context. However, to take advantage of the infrastructure already provided for `lme4` models, one may change the research question away from that of Roy’s paper. To this end, an exploration of what `textit`influence.ME can accomplish is merited. As a complement to this, one can also consider how to properly employ the R^2 measure, in the context of Method Comparison Studies, further to the work by Edwards et al, namely “An R^2 statistic for fixed effects in the linear mixed model”.

Abstract for “An R^2 statistic for fixed effects in the linear mixed model” Statisticians most often use the linear mixed model to analyze Gaussian longitudinal data.

The value and familiarity of the R^2 statistic in the linear univariate model naturally creates great interest in extending it to the linear mixed model. We define and describe how to compute a model R^2 statistic for the linear mixed model by using only a single model.

The proposed R^2 statistic measures multivariate association between the repeated outcomes and the fixed effects in the linear mixed model. The R^2 statistic arises as a 11 function of an appropriate F statistic for testing all fixed effects (except typically the intercept) in a full model.

The statistic compares the full model with a null model with all fixed effects deleted (except typically the intercept) while retaining exactly the same covariance structure.

Furthermore, the R^2 statistic leads immediately to a natural definition of a partial R^2 statistic. A mixed model in which ethnicity gives a very small p-value as a longitudinal predictor of blood pressure (BP) compellingly illustrates the value of the statistic.

In sharp contrast to the extreme p-value, a very small R^2 , a measure of statistical and scientific importance, indicates that ethnicity has an almost negligible association with the repeated BP outcomes for the study.

Leave-One-Out Diagnostics with `lmeU`

Galecki et al discuss the matter of LME influence diagnostics in their book, although not into great detail.

The command `lmeU` fits a model with a particular subject removed. The identifier of the subject to be removed is passed as the only argument

A plot of the per-observation diagnostics individual subject log-likelihood contributions can be rendered.

Likelihood Displacement

Missing Data in Method Comparison Studies

The matter of missing data has not been commonly encountered in either Method Comparison Studies or Linear Mixed Effects Modelling. However Roy (2009) deals with the relevant assumptions regarding missing data.

Galecki & Burzykowski (2013) tackles the subject of missing data in LME Modelling.

Furthermore the nlmeU package includes the `patMiss` function, which “allows to compactly present pattern of missing data in a given vector/matrix/data frame or combination of thereof”.