Preface

Linear mixed-effects models (LMMs) are powerful modeling tools that allow for the analysis of datasets with complex, hierarchical structures. Intensive research during the past decade has led to a better understanding of their properties. The growing body of literature, including recent monographs, has considerably increased their popularity among applied researchers. There are several statistical software packages containing routines for LMMs. These include, for instance, SAS, SPSS, STATA, S+, and R. The major advantage of R is that it is a freely available, dynamically developing, open-source environment for statistical computing and graphics.

The goal of our book is to provide a description of tools available for fitting LMMs in R. The description is accompanied by a presentation of the most important theoretical concepts of LMMs. Additionally, examples of applications from various research areas illustrate the main features of both theory and software. The presented material should allow readers to obtain a basic understanding of LMMs and to apply them in practice. In particular, we elected to present several theoretical concepts and their practical implementation in R in the context of simpler, more familiar classes of models such as e.g., the classical linear regression model. Based on these concepts, more advanced classes of models, such as models with heterogenous variance and correlated residual errors, along with related concepts are introduced. In this way, we incrementally set the stage for LMMs, so that the exposition of the theory and R tools for these models becomes simpler and clearer. This structure naturally corresponds to the object-oriented programming concept, according to which R functions/methods for simpler models are also applicable to the more complex ones.

We assume that readers are familiar with intermediate linear algebra, calculus, and the basic theory of statistical inference and linear modeling. Thus, the intended audience for this book is graduate students of statistics and applied researchers in other fields.

Our exposition of the theory of various classes of models presented in the book focuses on concepts, which are implemented in the functions available in R. Readers

interested in a more detailed description of the theory are referred to appropriate theoretical monograph books, which we indicate in the text.

There are a large number of R packages that can be used to fit LMMs. Rather than attempting to describe all of these packages, we focus mainly on two of them, namely, **nlme** and **lme4.0**. In this way, we can provide a more detailed account of the tools offered by the two packages, which include a wide variety of functions for model fitting, diagnostics, inference, etc.

The package **nlme** includes functions, which allow fitting of a wide range of linear models and LMMs. Moreover, it has been available for many years and its code has been stable for some time now. Thus, it is a well-established R tool.

In turn, **lme4.0** is a developmental branch version of the **lme4** package. The latter has been under development for several years. Both packages offer an efficient computational implementation and an enhanced syntax, though at the cost of a more restricted choice of LMMs, as compared to the **nlme** package. At the time of writing of our book, the implementation of LMMs in **lme4** has undergone major changes in terms of internal representation of the objects representing fitted models. Consequently, at the beginning of 2012, a snapshot version of **lme4** has been made available to the R users under the name of **lme4.0**. As we anticipate that **lme4.0** will not undergo any major changes, we decided to present it in more detail in our book. We would like to underscore, however, that the major part of the syntax, presented in the book, will be applicable both to **lme4** and **lme4.0**.

All classes of linear models presented in the book are illustrated using data from a particular dataset. In this way, the differences between the various classes of models, as well as differences in the R software, can be clearly delineated. LMMs, which are the main focus of the book, are also illustrated using three additional datasets, which extend the presentation of various aspects of the models and R functions. We have decided to include the direct output of R commands in the text. In this way, readers who would like to repeat the analyses conducted in the book can directly check their own output. However, in order to avoid the risk of incompatibility with updated versions of the software, the results of the analyses have also been summarized in the form of edited tables.

To further support those readers who are interested in actively using the material presented in the book, we have developed the package **nlmeU**. It contains all the datasets and R code used in the book. The package is downloadable at `http://www-personal.umich.edu/~agalecki/`.

We hope that our book, which aims to provide a state-of-the-art description of the details of implementing of LMMs in R, will support a widespread use of the models by applied researchers in a variety of fields including biostatistics, public health, psychometrics, educational measurement, and sociology.

When working on the text, we received considerable assistance and valuable comments from many people. We would like to acknowledge Geert Molenberghs (Hasselt University and the Catholic University of Leuven), Geert Verbeke (Catholic University of Leuven), José Pinheiro (Novartis AG), Paul Murrell (Auckland University), Przemysław Biecek (Warsaw University), Fabian Scheipl (Ludwig Maximilian University of Munich), Joshua Wiley (University of California, Los

Ann Arbor, MI, USA   Andrzej Gałecki
Diepenbeek, Belgium, and Warszawa, Poland   Tomasz Burzykowski