

Preface

This book is written for behavioral scientists who want to consider adding **R** to their existing set of statistical tools, or want to switch to **R** as their main computation tool. We aim primarily to help practioners of behavioral research make the transition to **R**. The focus is to provide practical advice on some of the widely used statistical methods in behavioral research, using a set of notes and annotated examples. We also aim to help beginners learn more about statistics and behavioral research. These are statistical techniques used by psychologists who do research on human subjects, but of course they are also relevant to researchers in others fields that do similar kinds of research.

We assume that the reader has read the relevant parts of **R** manuals on the CRAN website at <http://www.r-project.org>, such as “An Introduction to **R**”, “**R** Data Import/Export”, and “**R** Installation and Administration”. We assume that the reader has gotten to the point of installing **R** and trying a couple of examples. We also assume that the reader has relevant experiences in using other statistical packages to carry out data analytic tasks covered in this book. The source code and data for some of the examples in the book can be downloaded from the book’s website at: <http://idecide.mskcc.org/yl.home/rbook/>. We do not dwell on the statistical theories unless some details are essential in the appropriate use of the statistical methods. When they are called for, theoretical details are accompanied by visual explanations whenever feasible. Mathematical equations are used throughout the book in the hopes that reader will find them helpful in general, and specifically in reaching beyond the scope of this book. For example, matrix notations are used in the chapters covering linear regression and linear mixed-effects modeling because they are the standard notations found in statistics journals. A basic appreciation of mathematical notations may help the readers implement these new techniques before a packaged solution is available. Nevertheless, the main emphasis of this book is on the practical data analytic skills so that they can be quickly incorporated into the reader’s own research.

The statistical techniques in this book represent many of statistical techniques in our own research. The pedagogical plan is to present straightforward solutions and add more sophisticated techniques if they help improve clarity and/or efficiency.

As can be seen in the first example in Chap. 1, the same analysis can be carried out by a straightforward and a more sophisticated method. Chapters 1–4 cover basic topics such as data import/export, statistical methods for comparing means and proportions, and graphics. These topics may be part of an introductory text for students in behavioral sciences. Data analysis can often be adequately addressed with no more than these straightforward methods. Chapter 4 contains plots in published articles in the journal *Judgment and Decision Making* (<http://journal.sjdm.org/>). Chapters 5–7 cover topics with intermediary difficulty, such as repeated-measures ANOVA, ordinary least square regression, logistic regression, and statistical power and sample size considerations. These topics are typically taught at a more advanced undergraduate level or first year graduate level.

Practitioners of behavioral statistics are often asked to estimate the statistical power of a study design. R provides a set of flexible functions for sample size estimation. More complex study designs may involve estimating statistical power by simulations. We find it easier to do simulations with R than with other statistical packages we know. Examples are provided in Chaps. 7 and 11.

The remainder of this book cover more advanced topics. Chapter 8 covers Item Response Theory (IRT), a statistical method used in the development and validation of psychological and educational assessment tools. We begin Chap. 8 with simple examples and end with sophisticated applications that require a Bayesian approach. Such topics can easily take up a full volume. Only practical analytic tasks are covered so that the reader can quickly adapt our examples for his or her own research. The latent regression Rasch model in Sect. 8.4.2 highlights the power and flexibility of R in working with other statistical languages such as WinBUGS/OpenBUGS. Chapter 9 covers missing data imputation. Chapters 10–11 cover hierarchical linear models applied in repeated-measured data and clustered data. These topics are written for researchers already familiar with the theories. Again, these chapters emphasize the practical data analysis skills and not the theories.

R evolves continuously. New techniques and user-contributed packages are constantly evolving. We strive to provide the latest techniques. However, readers should consult other sources for a fuller understanding of relevant topics. The R journal publishes the latest techniques and new packages. Another good source for new techniques is The Journal of Statistical Software (<http://www.jstatsoft.org/>). The R-help mailing list is another indispensable resource. User contributions make R a truly collaborative statistical computation framework. Many great texts and tutorials for beginners and intermediate users are already widely available. Beginner-level tutorials and how-to guides can be found online at the CRAN “Contributed Documentation” page.

This book originated from our online tutorial “Notes on the use of R for psychology experiments and questionnaires.” Many individuals facilitated the transition. We would like to thank them for making this book possible. John Kimmel, former editor for this book at Springer, first encouraged us to write this book and provided continuous guidance and encouragement. Special thanks go to Kathryn Schell and Marc Strauss and other editorial staff at Springer on the preparation of the book. Several anonymous reviewers provided suggestions on how to improve the book.

We are especially indebted to the individuals who helped supply the data used in the examples, including the authors of the **R** packages we use, and those who make the raw data freely accessible online.

New York
Philadelphia

Yuelin Li
Jonathan Baron

Behavioral Research Data Analysis with R

Li, Y.; Baron, J.

2012, XII, 245 p. 32 illus., 11 illus. in color., Softcover

ISBN: 978-1-4614-1237-3