Statistical Methods in Medical Research

Book review: Rabe-Hesketh S, Skrondal A 2008: Multilevel and Longitudinal Modeling Using Stata, second edition. College Station, TX: Stata Press. 562pp. USD 59.00. ISBN 978-1-59718-040-5

Daniel Stahl

The online version of this article can be found at:
http://smm.sagepub.com

Additional services and information for *Statistical Methods in Medical Research* can be found at:

Email Alerts: http://smm.sagepub.com/cgi/alerts

Subscriptions: http://smm.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://smm.sagepub.com/cgi/content/refs/18/1/111

Book reviews

Rabe-Hesketh S, Skrondal A 2008: *Multilevel and Longitudinal Modeling Using Stata,* **second edition. College Station, TX: Stata Press. 562pp. USD 59.00. ISBN 978-1-59718-040-5.**

The analysis of longitudinal and other clustered data, such as patients within hospitals, is a common feature in medical research. Sophia Rabe-Hesketh and Anders Skrondal's book provides a comprehensive and thorough introduction to the analysis of such data using the general purpose statistical package Stata. The focus of the book is on multilevel modelling (or generalised linear mixed modelling) approaches. Alternative methods, such as general estimation equations, are only discussed briefly.

The text is written for applied statisticians, graduate students and researchers with experience in statistical modelling (including the regression approach to ANOVA), who have come across repeated measurements or other kinds of clustered data.

The second edition of this book, with almost twice as many pages includes three new chapters and an update to Stata 10 commands, including the new Stata commands **xtmelogit** and **xtmepoisson**. The authors are the developers of the Stata add-on program **gllamm**. It is, therefore, hardly surprising that the authors describe in detail the use of **gllamm**, which expands Stata's own commands considerably, especially for the analysis of categorical or discrete data. **gllamm** is a powerful program to analyse latent variable models, of which generalised linear mixed models are a special case.[1]

The book consists of 13 chapters. The first is a new review chapter and introduces the reader to linear regression and the general linear model. It explains in detail the use and interpretation of dummy coded categorical independent variables and interactions. An understanding of these concepts is a prerequisite for the following chapters. Chapter 2 introduces the concept of random effects or multilevel modelling by means of the variance component model, a simple random intercept model

without any covariates. Chapter 3 expands the random-intercept model by including covariates and in chapter 4 the model is extended further by including interactions between random factors and covariates (random intercept and slope or random-coefficient models). The new chapter 5 provides more details of the analysis of longitudinal data, including discussions of time varying covariates, autoregressive models and growth curves.

The following four chapters apply these concepts to models with non-normal responses: binary responses (Chapter 6, e.g. logit and probit models), ordinal responses (Chapter 7, e.g. ordered logit and ordered probit models), a new chapter on discrete survival times, including survival models with truncated and censored observations (Chapter 8) and count responses (Chapter 9, e.g. Poisson and negative binomial models). Finally, the last two chapters extend the random effects models to higher level models with nested (Chapter 10) and crossed (nonnested models: Chapter 11) random effects. An appendix containing an overview of the syntax of the **gllamm** commands completes the book.

The authors state in their preface that the 'emphasis is on explaining the models and their assumptions, applying the methods to real data, and interpreting results'. They succeed in their goal and provide a comprehensive and lucid account of multilevel modelling. Even complex models are clearly and sufficiently explained with surprisingly little mathematical content. Often neglected topics, such as the empirical Bayes' estimator, are covered in sufficient details. The book discusses many topics which are ignored in other standard textbooks, such as endogeneity, the difference between within and between subject effects in panel models or lagged response and autoregressive models.

The authors present Stata commands and output for all steps of an analyses, starting from initial examination of the data, fitting the model, visualising the results to model diagnostics. The book is, therefore, an excellent combination of

a thorough introduction to multilevel modelling and a tutorial on Stata commands for analysing such models. This is a major advantage to other textbooks that may show how to run the main analysis in software but leave the reader alone with assessing the model fit or visualising the results. The excellent integration of Stata analyses within the text, however, restricts the audience of this book to Stata users.

Examples in the text and exercises at the end of each chapter are stimulating and cover a wide range of real life applications from medical, behavioural and social science research. All data sets and Stata commands used in the text are available from an accompanying webpage or can be downloaded directly from within Stata. Answers to the exercises are only available to instructors adopting the book for a course.

There are only a few topics I missed: One is that problems with missing data are only briefly discussed. The authors explain that multilevel models are less restrictive in the assumption of missingness (missing at random) than complete case or GEE analysis (missing completely at random). However, they do not provide guidance to assess the type of randomness or provide recommendations how to analyse data if missingness is not at random. I also missed a paragraph on model selection, e.g. the use (and its potential problems) of information criteria. My only minor criticism of this book is the index. It is not very extensive, which makes it difficult for the reader to find minor topics, like power or explained variance.

In summary, the book is a very well-written and provides comprehensive guide to multilevel modelling, both from a practical and a theoretical point of view. The second edition of this book will continue to be an invaluable help for every Stata user involved in the analysis of clustered data and it will replace the tattered first edition in my bookshelf.

Reference

1 Skrondal A, Rabe-Hesketh S. *Generalized latent variable modeling: multilevel, longitudinal and structural equation models*. Chapman & Hall/CRC, Boca Raton, FL, 2004.

Reviewed by Daniel Stahl, Department of Biostatistics, Institute of Psychiatry, King's College London, UK.

This book is written for clinical pharmacologists, biopharmaceutical scientists, reviewers from regulatory affairs and biometricians. It also provides a reference for scientists from academia conducting research in the area of biopharmaceutics and clinical pharmacokinetics. The book covers setting up, conduct, analysis and reporting of bioequivalence studies according to regulatory requirements. Example data sets and SAS® code have been used to illustrate methods and to help the reader rework the analysis. Screenshots are also available from nQuery Advisor® for power and sample size determination. The conclusions at the end of each chapter provide an excellent summary. The book provides a good reference text to bioequivalence but will probably be more useful to professionals and research workers with some background to pharmacokinetics rather than those who have no prior knowledge.

The book contains 10 chapters, each with a list of references. The introduction gives definitions of bioavailability, bioequivalence and therapeutic equivalence taken from current guidelines of the US and European regulatory bodies. It also explains when bioequivalence studies are performed, as well as the design and running of such studies. Chapter 2 examines ways to characterise the concentration–time profile in single and multiple dose bioequivalence studies, where the main interest is assessment of 'rate and extent of absorption'. Chapter 3 covers the basic statistical methods necessary for bioequivalence trials, for example the relationship between AUC, dose and clearance, the fundamental properties of the normal and log-normal distributions and hypothesis testing with respect to bioequivalence designs, including a discussion on Type I and Type II errors (consumer and producer risk). The statistical methodology described here is introductory and will be accessible to the non-statistician. It covers the analysis of the standard two-period, two-sequence, crossover design and the consideration of period, carryover and formulation effects.

The focus of chapter 4 is to demonstrate the underlying methodology for the statistical assessment of average bioequivalence. The author introduces the multiplicative model, following on from the additive model introduced

This text is a Stata-specific treatment of generalized linear mixed models, also known as multilevel or hierarchical models. These models are "mixed" in the sense that they allow fixed and random effects and are "generalized" in the sense that they are appropriate not only for continuous Gaussian responses but also for binary, count, and other types of limited dependent variables. Sophia Rabe-Hesketh & Anders Skrondal, 2012. "Multilevel and Longitudinal Modeling Using Stata, 3rd Edition," Stata Press books, StataCorp LP, edition 3, number mimus2, April. Handle: RePEc:tsj:spbook:mimus2. …[Rabe-Hesketh and Skrondal] have produced a comprehensive, nontechnical introduction to multilevel/mixed effect models that would be an excellent choice of text for application-oriented courses on the analysis of longitudinal, panel, or clustered data.… …Sophia Rabe-Hesketh and Anders Skrondal move seamlessly between the methods and language of biostatistics and econometrics… Excerpt from book review: "Multilevel and Longitudinal Modeling Using Stata (Second ed.)" by D. B. Hall. 2010. American Statistician 64: 275. Excerpt from book review: "Multilevel and Longitudinal Modeling Using Stata, Second Edition" by T. M. Loughin. 2008. Biometrics 64: 1310. 2008. Stata Journal 8: 579–582. The book deftly applies multilevel and longitudinal models of various complexity and provides insightful interpretations for model parameters. The authors step through each analysis with complete commands for implementation in Stata. Following the authors through these analyses gives the reader the condence to explore these models further with a combination of the xtmixed and gllamm commands. The exercises provided at the end of each chapter are an excellent place to start. 4 Limitations. Rabe-Hesketh, S., and A. Skrondal. 2005. Multilevel and Longitudinal Modeling Using Stata. College Station, TX: Stata Press. R. Wolfe. 143. Rabe-Hesketh, S., A. Skrondal, and A. Pickles. 2004. GLLAMM Manual. The first edition of Rabe-Hesketh and Skrondal's "Multilevel and Longitudinal Modeling Using Stata" was published in 2005. The second edition was released in 2008, and now this third edition in 2012. With each edition the scope of the model's discussed in the text has increased. This release is in fact a 2-volume work, with the first volume devoted to panel models having a continuous response (or dependent variable), and the second to discrete response panel models. Nearly every panel model in the literature is addressed in the text. Discussion begins with a review of basic Abstract. This article reviews Multilevel and Longitudinal Modeling Using Stata, by Rabe-Hesketh and Skrondal. Discover the world's research. 19+ million members. An introduction to measurement models measures of proximity, agreement and disagreement reliability, consistency and stability designs for reliability studies covariance-components models components of variance statistical methods for categorical measurements. Appendices: expected values linear combinations of measurements the delta technique computer software. View. Show abstract. A Handbook of Statistical Analyses Using Stata. Article. May 2000. AM STAT. Peter A. Lachenbruch. S. Rabe-Hesketh. B. Everitt. View. A Handbook of Statistical Analyses Using Stata. Article. Jan 2004.