The Stata Journal (2006)
6, Number 1, pp. 138–143

Review of Multilevel and Longitudinal Modeling Using Stata by Rabe-Hesketh and Skrondal

Rory Wolfe
Department of Epidemiology and Preventive Medicine
Monash University
Melbourne, Australia
rory.wolfe@med.monash.edu.au

Abstract. This article reviews *Multilevel and Longitudinal Modeling Using Stata*, by Rabe-Hesketh and Skrondal.

Keywords: gn0031, longitudinal, multilevel, gllamm, generalized latent variable model

1 Introduction

Multilevel and Longitudinal Modeling Using Stata (Rabe-Hesketh and Skrondal 2005) addresses many interesting datasets in its focus on the application of methods for multilevel and longitudinal data. The authors state that their "emphasis is on explaining the models and their assumptions, applying the methods to real data, and interpreting results" (xxi). As I am an applied statistician who often deals with these challenges, it sounded like my sort of book.

I suspected that the book might simply be a showcase for the authors' `gllamm` command (Rabe-Hesketh, Skrondal, and Pickles 2004) and the underpinning methodology of generalized latent variable modeling (GLVM) (Skrondal and Rabe-Hesketh 2004). This command would be of interest in itself and to an extent my suspicion is true: in this book `gllamm` is used for illustration and comparison with `xtmixed` in situations where in practice one is likely just to use `xtmixed`. However, the book contains much more. The authors provide insightful analyses and helpful interpretations that are useful regardless of your choice of software.

All datasets used in this book can be downloaded from http://www.gllamm.org/books/. The `gllamm` command can be downloaded from within Stata with `ssc install gllamm` if you have an Internet connection; otherwise, the necessary files can be downloaded from http://www.gllamm.org/. The book provides a full description of the syntax for the suite of `gllamm`-related commands in an appendix, and more detail can be found in the `gllamm` manual (Rabe-Hesketh, Skrondal, and Pickles 2004).

The first three chapters of the book deal with models for continuous outcomes based on the normal distribution; the next three cover binary, ordinal, and count outcome data in turn; and the last two chapters cover multilevel models with nested or cross-random effects. I found chapters 1–3 helpful in getting my thoughts straight when

preparing introductory material on reliability and repeated measures for Master's of Public Health students pursuing an advanced statistics course. Although the material in these chapters helped me prepare, I would not recommend it as appropriate reading for these students.

For whom is this book written, then? The authors do not say, but they do expect familiarity with Stata, and although anyone can mimic the commands and obtain the authors' output, one must have considerable familiarity with statistical methods to understand the model formulas presented and follow the parameter interpretations. The book is aimed primarily at experienced applied statisticians who have already been introduced to the challenges of longitudinal and/or multilevel models.

Exercises are provided at the end of each chapter that challenge the reader to use the methods and commands of the book without the full syntax being provided. The authors provide good references to further reading at the conclusion of each chapter.

2 Overview of the book

Chapter 1 introduces variance-components models, provides a bridge from previous versions of Stata with the `xtreg` and `gllamm` commands to version 9 and the `xtmixed` command, and finally covers the issue of obtaining values for individuals who are modeled by a random effect (proposing either maximum likelihood estimation or empirical Bayes prediction). The authors' concise style is an advantage as they compare different Stata commands to fit the same model in this and the next two chapters. The authors use a well-known dataset on peak expiratory flow rates (Bland and Altman 1986) to introduce two- and three-level variance-components models (three-level models in chapter 7) and to estimate intraclass correlation coefficients. These models provide estimates of the standard deviation (SD) that is required to construct limits of agreement (Bland and Altman 1986), although users must remember the difference between SDs for single determination and SDs for change (Chinn 1991). Multilevel variance-components models provide a framework for generalized modeling of agreement or reliability studies (Dunn 2004). As these models are only a subset of GLVM, we begin to see the breadth of that methodology.

The introduction to random-intercept models in chapter 2 focuses on the important distinction between within- and between-cluster effects and on residual checks at levels 1 and 2 of the model. The data used here are from taxpayer returns (showing little evidence of a reduction in tax liability when using a tax consultant or preparer, rather than doing it oneself) and illustrate the between- and within-cluster effects. Chapter 3 covers random-intercept random-slope models; those familiar with MLwiN (Rasbash, Browne, and Goldstein 2003) will recognize the data on exam results from school children. The authors provide a clear explanation for avoiding, in general, the `xtmixed` default of assuming zero correlation between level 1 and level 2 random effects. Using `gllamm`, they arrive at estimated school rankings for test performance with 95% confidence intervals, thereby touching on an important but controversial area for education and health systems analysis. A second dataset on child growth is used nimbly

to further illustrate these models as growth curves. A helpful section is included to demonstrate these models when specified in a two-stage formulation.

Chapter 4 contains a comprehensive introduction to logistic and probit regression. While introducing the first dataset on toenail infections, the authors raise the important issue of missing data in longitudinal studies. Disappointingly, there are no references to further reading on how to deal with this issue in practice. The authors' approach of ignoring the fact that some subjects have missing data is not generally to be recommended, as it may result in biased estimates; see, for example, Little and Rubin (2002). Full consideration is given to marginal and conditional (population averaged and subject specific) model structures and interpretations. A total of 162 of 294 patients had outcome 0 at each of the seven visits (but 32 of these had missing data for one or more visits) and 14 patients had outcome 1 at all their visits (six with missing data for at least one visit). The rest generally improved from 1 to 0 over time. The authors do not provide a description of these subject-specific response patterns (marginal patterns are displayed), which are the natural starting place for understanding the random-effect modeling, and in particular the large random-effects variance seen in this example and the empirical Bayes predictions. However, they do try to aid the interpretability of the results, including a neat plot of Bayes predictions. This chapter closes with brief but welcome notes on conditional logistic regression and generalized estimating equation approaches to clustered binary outcome data.

Chapter 5 covers ordinal outcomes fitted with ordered logistic and ordered probit models. There is a reasonably extensive introduction to the basic models and parameter estimation before introducing random effects to the model. Quickly, however, the authors reach elaborate models and demonstrate how straightforward they are to fit with `gllamm`. Importantly, a rationale is provided for each complex model for research questions. One current limitation of `gllamm` is that for the scaled models one is restricted to probit regression.

Chapter 6 covers Poisson models for count data and ties the phenomenon of overdispersion into a brief introduction to the general model. The chapter includes an exploration of quasilikelihood, as well as generalized estimating equations and conditional Poisson regression. This chapter highlights the limitation of assuming normality for latent variables in `gllamm`, and the authors demonstrate by using lip cancer rates (also analyzed in Skrondal and Rabe-Hesketh 2004) results for a "nonparametric" random effect in which a discrete distribution is assumed for a, typically small, integer number of values.

The analysis of peak expiratory flow rates continues in chapter 7 with the full dataset being used and a three-level model being fitted. This analysis is an example of a continuous outcome, analyzed with the `xtmixed` command. There is a demonstration of `gllamm` fitting a three-level model for binary outcomes. Finally, chapter 8 covers the issue of crossed random effects and includes two examples, investment data from firms crossed with years and school attainment scores with secondary schools crossed with primary schools. These examples illustrate the possibilities of `xtmixed`, so this chapter is limited to continuous outcomes; the authors acknowledge that `gllamm` is too slow for

this application. The school data can be explored further, as prompted in an exercise at the end of the chapter: a fixed effect for "family ability" explains much of the apparent variability at both secondary and primary school levels. This analysis could be pursued further by using empirical Bayes predictions to provide school rankings, along the lines of chapter 3.

3 Strengths

Rather than providing an exhaustive introduction to multilevel and longitudinal models, the book generally gives an excellent concise treatment, with references to further reading. The authors do provide more detail on certain issues than is usually encountered—in particular, parameter interpretation. These subsections will be especially useful for applied statisticians.

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 confidence 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

The GLVM methodology encompasses a rich class of models, so a computer routine in Stata for GLVM is welcome. However, `gllamm` is not as user-friendly as it could be. The syntax presents a challenge and often takes a long time to complete its estimation routine. Section 4.7.1 in the book provides tips for speeding up estimation, and a more thorough description of the speed issue, with proposed ways to minimize the estimation time, can be found in the `gllamm` manual (Rabe-Hesketh, Skrondal, and Pickles 2004). The challenge of syntax is not restricted to `gllamm`. My most perplexing moment in undertaking this review came when I inadvertently put a space between a top cluster-level variable and the following colon in `xtmixed`. Do try it.

Being specific to the structure of `xtmixed` in Stata 9, the book may have a short shelf life in places. It is to be hoped that some of that command's limitations, and hence the need to resort to `gllamm`, will be removed in later versions of Stata. Similarly, if Stata introduces a built-in command to fit random-effects logistic regression models, chapter 4 of this book will need updating.

In addition to minor limitations mentioned in the previous section, it is surprising that the book does not provide an example of longitudinal modeling for latent trajectories, e.g., by extending the boys and girls growth curve example, that would use `gllamm`'s option for a discrete distribution for random effect.

5 Conclusion

This book provides a nice accompaniment to the authors' other books (Skrondal and Rabe-Hesketh 2004; Rabe-Hesketh and Everitt 2004), which have previously been reviewed in the *Stata Journal* (Winter 2004; Newson 2005), and the `gllamm` manual (Rabe-Hesketh, Skrondal, and Pickles 2004). Although there is overlap between this book and the `gllamm` manual, this book contains a new array of datasets. For any reader not already familiar with Stata, the introductory book (Rabe-Hesketh and Everitt 2004) is a better starting place than this one. If you are intrigued by GLVM's flexibility (noting that this book covers only a portion of potential applications), you should progress to the methodology book (Skrondal and Rabe-Hesketh 2004).

Overall, the authors do an excellent job of motivating their analyses from research questions, providing insightful descriptions of the observed data, and describing parameter estimates in terms that are understandable to the casual reader. Similarly, they provide helpful descriptions of difficult parameterizations and complex estimation problems.

This book is aimed at experienced applied statisticians who have already spent time grappling with longitudinal and/or multilevel models in Stata (or people who wish to spend their time that way). The insightful applications and interpretations of a range of models combined with the ability to easily replicate all of the authors' analyses make this book a trusty companion.

6 References

Bland, J. M., and D. G. Altman. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* i: 307–310.

Chinn, S. 1991. Statistics in respiratory medicine. 2. Repeatability and method comparison. *Thorax* 46: 454–456.

Dunn, G. 2004. *Statistical Evaluation of Measurement Errors: Design and Analysis of Reliability Studies*. London: Arnold.

Little, R. J. A., and D. B. Rubin. 2002. *Statistical Analysis with Missing Data*. 2nd ed. New York: Wiley.

Newson, R. 2005. Review of Generalized Latent Variable Modeling by Skrondal and Rabe-Hesketh. *Stata Journal* 5: 130–133.

Rabe-Hesketh, S., and B. Everitt. 2004. *A Handbook of Statistical Analyses Using Stata*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC.

Rabe-Hesketh, S., and A. Skrondal. 2005. *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles. 2004. GLLAMM Manual. Technical report, University of California Berkeley, Division of Biostatistics. http://www.bepress.com/ucbbiostat/paper160/.

Rasbash, J., W. Browne, and H. Goldstein. 2003. *MLwiN 2.0 Command Manual*. Boca Raton, FL: Chapman & Hall/CRC.

Skrondal, A., and S. Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: Chapman & Hall/CRC.

Winter, N. 2004. Review of A Handbook of Statistical Analyses Using Stata by Rabe-Hesketh and Everitt. *Stata Journal* 4: 350–353.

About the author

Rory Wolfe is a medical statistician at Monash University who has worked on the methodology and application of longitudinal and multilevel models. He has contributed Stata commands to aid the analysis of ordinal outcomes, which can be downloaded from the SSC archive.