Applied Longitudinal Analysis, Second Edition, by G. M. Fitzmaurice, N. M. Laird, J. H. Ware, Hoboken, NJ: Wiley, 2011, ISBN 978-0-470-38027-7, xxv + 701 pp., $125.

This book provides very broad coverage of modern methods for longitudinal data analysis from an applied perspective. The book is grouped into five parts: Introduction to Longitudinal and Clustered Data; Linear Models for Longitudinal Continuous Data; Generalized Linear Models for Longitudinal Data; Missing Data and Dropout; and Advanced Topics for Longitudinal and Clustered Data. Throughout, real examples are used to illustrate the ideas, and snippets of SAS code provide insight into model fitting.

Besides being recognized leaders in the field, the authors are exceptionally skilled at communicating advanced statistical ideas in a way that leads to conceptual intuition. They accomplish this by avoiding too much technical jargon, and by reinforcing the ideas with illustrative graphs and only the necessary equations. In addition, they make it very easy for readers to try fitting the models themselves, as the online supplemental material includes 27 data sets and a dozen examples of SAS code that correspond to sections in the book.

This is the Second Edition of the *Applied Longitudinal Analysis*. The First Edition received several detailed reviews (Demirtas, 2005; Molenberghs, 2005; Robinson, 2007), all of which were very positive. I will therefore place more emphasis on the new material for the remainder of this review.

Most of the first part of the book is devoted to models for continuous longitudinal data. There are chapters devoted to estimation, modeling the mean, modeling the covariance, and mixed effects models. All of these chapters are well written, and provide a good foundation for the remainder of the book. After the chapter on linear mixed effects model, the new edition now

includes a chapter on fixed effects models. This is a welcome addition, as these models are popular in economics and the social sciences, but tend to receive little attention in the biostatistics literature. The chapter begins with a discussion of the assumptions that each model requires and the trade-offs between the two. They then propose to resolve the dilemma of which model to choose by specifying a random effects model with exogenous time-varying covariates decomposed into cross-sectional and longitudinal components. While this option has been proposed before (e.g., Neuhaus and McCulloch, 2006), my impression is that it has not reached the consciousness of many statisticians. I was surprised that this chapter did not include a discussion of conditional likelihood approaches, although this topic is briefly discussed later in the book.

The next chapter covers residuals and diagnostics. The new edition of the book has added some material on cumulative residuals. They present a case study on the relationship between menarche and changes in body fat accretion. The case study very effectively shows how cumulative residuals can be used to diagnose problems with the structural form of the model.

The next six chapters of the book cover generalized linear models for longitudinal data, including generalized linear mixed models (GLMMs) and generalized estimating equations. A chapter was added on approximation methods for GLMMs (penalized quasi-likelihood and marginal quasi-likelihood) and includes discussions of situations where these methods will likely perform poorly.

Two chapters are devoted to missing data. The first provides an overview, and largely focuses on dropout. The next chapter, which is new in the second edition, focuses on multiple

imputation and inverse probability weighting. As with the other chapters, the ideas are illustrated with data examples and SAS code.

The Advanced Topics section of the book now includes chapters on semiparametric regression and power / sample size calculation. The semiparametric regression chapter focuses on penalized splines. This is an excellent addition to the book, as linear models for time trends are often too restrictive. The explanations here are very intuitive, and demonstrate how these models can be fitted with standard mixed models software such as SAS PROC MIXED. The chapter on power and sample size emphasizes simplifying assumptions that can be made to get approximate estimates in a variety of settings. In general, sample size calculations for longitudinal data are challenging due to the more complicated nature of the data. This chapter should prove useful as at least a starting point for obtaining meaningful estimates.

The final two chapters of the book cover repeated measures designs and multilevel models. I would like to have seen a chapter on causal modeling and time-dependent confounding in the Advanced Topics section, but otherwise the book covers all of the major longitudinal data topics.

I highly recommend this book to statisticians and quantitative researchers who encounter longitudinal and/or clustered data. In addition, I think the book would be an excellent choice as the primary textbook in an applied longitudinal data course. There are better sources when it comes to the technical details of the topics covered in this book, but few sources will provide clearer explanations, illustrated with numerous interesting, real-world examples.

REFERENCES

Demirtas, H. (2005). Book review of Fitzmaurice, G. M., Laird, N. M., and Ware, J. H., 2004, *Applied Longitudinal Analysis*. *Statistical Methods in Medical Research* 14: 321-322.

Molenberghs, G. (2005). Book review of Fitzmaurice, G. M., Laird, N. M., and Ware, J. H., 2004, *Applied Longitudinal Analysis*. *Journal of the American Statistical Association* 100: 709-710.

Neuhaus, J. M. and McCulloch, C. E. (2006). Separating between- and within-cluster covariate effects by using conditional and partitioning methods. *Journal of the Royal Statistical Society: Series B* 68: 859-872.

Robinson, A. (2007). Book review of Fitzmaurice, G. M., Laird, N. M., and Ware, J. H., 2004, *Applied Longitudinal Analysis*. *Statistics in Medicine* 26: 3824-3825

Jason Roy

Center for Clinical Epidemiology and Biostatistics

University of Pennsylvania

Philadelphia, PA, 19104, USA