Fitzmaurice GM, Laird NM, Ware JH 2004: *Applied longitudinal analysis*. New Jersey: John Wiley & Sons. 506 pp. US$89.95 (HB). ISBN 0 471 21487 6.

This is a systematically written and comprehensive book on fundamentals of longitudinal data analysis, primarily designed towards applied researchers in biostatistics and quantitatively oriented health and medical scientists. A wide range of readers (from statistically challenged medical researchers to technically advanced statisticians) have something to learn from this book. With a wealth of examples drawn from real-life studies, it focuses on modern methods for analysing longitudinal data in an insightful, intuitive and horizon-broadening way. Assumed background is basic knowledge of statistics and comprehension of regression at an intermediate level.

Life was quite simple when the only available method for longitudinal data was repeated measures ANOVA. In the last two decades, we have witnessed a sizeable proliferation of longitudinal data techniques that can operate on a more general class of problems. A textbook that adequately covers modern statistical methods for longitudinal data analysis at a level accessible to a broad audience was missing; the work of the authors seems to fill this gap. Although mainstream books such as Diggle *et al.* (2003),[1] Verbeke and Molenberghs (2000),[2] Davis (2002),[3] Singer and Willett (2003),[4] Lindsey (1999),[5] Hand and Crowder (1996),[6] Brown and Prescott (2003),[7] Davidian and Giltinan (1998),[8] and Vonesh and Chinchilli (1997)[9] also deserve our respect, this book is less technical, but more intuitive and comprehensive than most of the cited books above. This book is refreshingly different from every other longitudinal data analysis book I have seen, although they are not directly comparable given their scopes.

The book starts with descriptions of datasets and general introductory concepts in longitudinal data analysis. Then, the authors focus on methods applicable when the response variable is continuous and assumed to have an approximate multivariate normal distribution: a general linear regression model is introduced, and estimation issues are discussed. In the next few chapters, model-building matters are presented in depth in terms of the mean response and covariance structure, emphasizing the linear mixed model paradigm, along with residual diagnostics for assessing the adequacy of models.

Several subsequent chapters focus on methods for analysing repeated binary responses and repeated count data. Generalized linear models (GLM) are covered as an inferential tool for discrete responses. Extensions of GLM such as semi-parametric marginal models (GEE and its versions) and GLM with random effects are presented. There has been a tremendous interest in these approaches in the biostatistics community in recent years and the authors systematically provide an overview to address these relatively new topics.

The final part of the book covers a number of advanced topics such as missing data in longitudinal studies, design of studies with sample size and power considerations, repeated measures and related designs, and an overview of multilevel models.

A particular strength of the book is that the authors take a balanced view of competing approaches, which they illustrate with numerous biomedical examples that demonstrate the applications of the theory. Most sections include practice problems; however, the answers are not provided. The computing aspects are addressed with SAS implementations. Worked examples are presented in parallel with the methodological development, and adequate detail is given for the reader to be able to reproduce the authors' results using the datasets provided on their website.

In summary, this book concisely presents essential ideas on modern longitudinal data analysis techniques, with many examples, in an intuitively appealing manner, both on analytical and conceptual levels. It has the potential to be a classic text, and an important reference and consulting source.

References

1 Diggle JP, Heagerty P, Liang KY, Zeger SL. *Analysis of longitudinal data*, 2nd edn. Oxford: Oxford University Press, 2003.

2 Verbeke G, Molenberghs G. *Linear mixed models for longitudinal data*. New York: Springer-Verlag, 2000.

3 Davis CS. *Statistical methods for the analysis of repeated measurements*. New York: Springer-Verlag, 2002.

4 Singer JD, Willett JB. *Applied longitudinal data analysis/modeling change and event occurrence*. New York: Oxford University Press, 2003.

5 Lindsey JK. *Models for repeated measurements*, 2nd edn. Oxford: Oxford University Press, 1999.

6 Hand D, Crowder M. *Practical longitudinal data analysis*. London: Chapman and Hall, 1996.

7 Brown H, Prescott R. *Applied mixed models in medicine*. New York: Wiley, 2001.

8 Davidian M, Giltinan DM. *Nonlinear models for repeated measurement data*. Boca Raton, FL: Chapman and Hall/CRC, 1998.
9 Vonesh EF, Chinchilli VM. *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker, 1997.

Reviewed by Hakan Demirtas, Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, USA.

Altman M, Gill J, McDonald MP. 2004: *Numerical issues in statistical computing for the scientist*. New York: John Wiley & Sons, Inc. 266 pp. $89.95. ISBN 0 471 23633 0.

Even with an extensive background in statistical theory, a scientist may have many questions remaining about the implementation of statistical procedures on a given set of data. Those with little interest in the exact implementation of the methodology can soon find themselves having to understand implementation details when 'unexpected' issues manifest themselves occasionally through unclear warnings or errors from standard statistical software. Unfortunately, basic understanding of numerical computing as implemented on modern computational hardware is sometimes lacking in even advanced education. This book provides the researcher with an overview of the issues involved in the implementation and computation of common statistical procedures. This should allow the reader to have more confidence in interpreting and troubleshooting common statistical models. The text also gives the reader the ability to better compare the available statistical software and understand the differences that exist.

The text is divided into two sections. The first section is a general overview of numerical computing issues that includes basic problems inherent to numerical computing, methods of evaluating and comparing statistical software, and computational issues that can interfere with statistical inference. The second section presents six specific areas that pose interesting numerical problems considered by both the authors and leading figures in the field of statistics:

1) Hessian matrices;
2) Markov chain Monte-Carlo estimation;
3) King's ecological inference method;
4) Nonlinear estimation;
5) Spatial Regression Models;
6) Convergence in logistic regression.

The first section of the book requires basic knowledge of computer arithmetic and statistical programming as well as intermediate knowledge of statistics that should be possessed by researchers employing inferential methods. The second section of the book is better understood with a background knowledge of specific statistical methodology and application.

The first section presents issues inherent to computational processing that can affect statistical analyses. This section includes a brief explanation of floating point arithmetic, a commonly misunderstood necessity of computing which leads to many questions in both the S-News and R-Help mailing lists. Also included in this section is an overview of random number generation, which is essential to many modern statistical techniques. In addition, an entire chapter in the first section is devoted to the evaluation of statistical software, providing further information on issues that can influence results of statistical analyses. Specific issues discussed in this section include generation and computation of distribution functions, faults in pseudo-random number generators and accuracy of input/output routines. A description of benchmarking using Standard Reference Datasets (StRD) leads into a comparison of statistical packages and languages. I believe that comparisons between statistical packages should be presented in a more real time media because of the frequency of new releases of software that may include bug fixes and updated algorithms. This is especially important when including open-source products such as R. However, information on benchmarking is important so the reader knows how comparisons should be made.

The second section of the text presents the most interesting reading because of its focus on individual statistical techniques. The chapters are well written and provide information in a variety of topics that researchers use when analysing data.

The first area of focus comprises two key components: the importance of the Hessian in likelihood-based estimation, and possible solutions for practitioners of statistical data analysis when facing a non-invertible Hessian. Anyone who has implemented likelihood-based estimation on even an occasional basis has encountered the warning 'Hessian not invertible' from statistical software. Unfortunately the default reaction of many researchers is to change the question they are trying to answer by adjusting the model to avoid getting the warning or error, with no regard to the fact they are changing the original question. Although in some cases this process may not be harmful, it should not be done without careful