This area is a little more complicated than the Gaussian one, because there is no such unifying concept as the multivariate normal distribution. To begin with, there is the important distinction between marginal and random-effects models. The authors succeed in providing a good perspective on both of these, with particular emphasis on generalized estimating equations (GEEs) and generalized linear mixed models (GLMMs). Specific issues surrounding GEEs, such as the choice of a working correlation matrix, are described in a clear fashion. The discussion on the difference between correlations and odds ratios in this context is perhaps a little brief, but the most important thing is that the reader will understand how to use standard GEEs as well as alternating logistic regressions. The authors provide a good perspective on the similarities and differences between marginal and random-effects models. Chapter 13 presents a comprehensive, yet concise view of the relationships that exist between both models, while at the same time not leaving unanswered the practical question as to when to choose from which model family.

In the special topics part, the missing-data chapter takes a prominent place. Even though relatively short, this chapter develops essential terminology and a perspective on what is typically done in practice, including complete-case analysis and last value carried forward (LVCF), which the authors rightly do not recommend. They then discuss the feasibility and ease of likelihood methods under the missing-at-random assumption and indicate how weighting methods work. There is a general feeling that weighting methods are complicated, and some of the finer points of the semiparametric weighting theory, including weighted GEE, happens to be technical. However, the reader will painlessly acquire a good working knowledge of this methodology from Chapter 13, even though its conciseness may imply that some further reading may be necessary.

Another topic treated is design aspects. The authors also provide a clear perspective on multilevel models. There is quite a bit of confusion as to the relationship between this modeling framework, longitudinal data, and mixed-effects models, for example. Although there are strong relationships and even overlaps, there are differences as well. The authors do a good job clarifying these interrelationships. In particular, they intuitively explain how longitudinal data can be seen as a specific instance of clustered or correlated data, and how many models for longitudinal data can be formulated within the general multilevel modeling framework.

The authors successfully use a very pedagogical and didactic writing style. The many case studies and illustrations are tremendously useful to enhance understanding. The inclusion of problems at the end of every chapter is very useful as well. *Applied Longitudinal Analysis* should be on the shelf of everyone interested in acquiring a modeler's or practitioner's perspective on longitudinal data analysis.

Geert MOLENBERGHS
Limburgs Universitair Centrum

REFERENCE

Searle, S. R., Casella, G., and McCulloch, C. E. (1992), *Variance Components*, New York: Wiley.

Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models.

Anders SKRONDAL and Sophia RABE-HESKETH. Boca Raton, FL: Chapman & Hall/CRC, 2004. ISBN: 1-58488-0007. x + 508 pp. $89.95.

This book encompasses the major developments in the theory and practice of generalized latent variable modeling. It has two fundamental features that make it one of the most comprehensive reference books in the field: an up-to-date guide to multilevel and structural latent variable modeling and estimation, plus a multidisciplinary set of illustrative examples from biology, medicine, psychology, education, sociology, political science, economics, marketing, and other areas. These are extremely enlightening for experienced practitioners in the many areas in which latent variable modeling can be used to analyze data. As a complementary toolkit, the authors' Stata package GLLAMM (generalized linear latent and mixed models) and associated datasets, which are available for free download, can be used to replicate the examples in Chapters 9–14.

As the authors state in the Preface, the interplay between methodology and applications is particularly valuable. The book is organized in such a way the intended reader can always find illustrations of methodology by skipping forward to the application chapters, and statistical background for applications is obtainable by skipping backward to the methodology chapters, with technicalities always kept at a reasonable minimum. There are cross-references between the statistical and the illustrative sections merging the two parts of the book, plus a set of complementary readings at the end of each chapter. However, most of the text is self-contained, so the reader rarely needs to make immediate reference to the cited literature.

The first three chapters offer an introduction to the book's modeling approach. As stressed in Chapter 1, latent variables pervade modern statistics and are becoming more and more transversal as they are used to represent different phenomena, such as true variables measured with error, hypothetical constructs, unobserved heterogeneity, missing data, counterfactuals, and latent responses underlying categorical variables. Chapter 2 describes a wide set of possible response processes, ranging from generalized linear models to latent response models. This chapter is the building block of the illustrative examples reported from Chapters 9–14 and may serve as a general guide to the different types of response processes modeled later in the book. Chapter 3 provides a swift overview of classical latent variable models such as, among others, multilevel regression models, structural equation models, and latent class models. Even if the book presupposes a reasonable amount of knowledge of econometrics, statistics, and matrix algebra, the understanding of some topics, such as the statistical matrix notation of two-level and three-level random-coefficient models, is notably simplified by a set of illustrative displays; these are useful devices also for later chapters.

Chapter 4 merges the introductory section with later materials. It unifies in a general framework both the response processes described in Chapter 2 and the models described in Chapter 3. The framework mostly corresponds to the generalized linear latent and mixed models of Rabe-Hesket et al. (2004), where the general model formulation involves the specification of a set of hierarchical conditional relationships unifying multilevel, structural, and latent class models. It accounts for a wide set of possible regression structures between factors and random coefficients at different levels, and for a wide set of responses and flexible specification of latent variable distributions that can be estimated using nonparametric methods. The Stata program GLLAMM can estimate all of the models discussed in the chapter; but, as the authors stress, the multivariate normal distribution is the only continuous latent variable distribution currently available.

Chapters 5 and 6 focus on the identification and estimation of the statistical models analyzed in the introductory chapters. Chapter 6 offers a rather nontechnical overview of frequentist and Bayesian estimation methods and lists some of the available software. Chapter 7 analyzes the empirical Bayes prediction methods and the maximum likelihood methods commonly used for assigning values to continuous latent variables, as well as the empirical Bayes modal method, which is used for discrete latent variables. Finally, Chapter 8 examines the steps needed for correct model specification and inference.

Comparable books are those of Bartholomew and Knott (1999), who focus on latent variable models and factor analysis; Bollen (1989), who focuses mainly on structural equations with latent variables; and Wansbeek and Meijer (2000), who analyze factor and structural equation models. To my knowledge, the present book is the first to provide a truly unifying generalized approach to latent variable modeling.

I did not identify any glaring omissions in the topics covered in the methodology section, although an issue just marginally mentioned by the authors in Section 3.2.2 is witnessing an increasing amount of new research: even when the endogeneity of latent variables with respect to covariates is accommodated by a structural approach, there may still exist a correlation between residuals and covariates at some level of the hierarchical structure. Such correlations reflect different forms of interactions and externalities (Brock and Durlauf 2001). Particularly in the field of social sciences, the importance of such externalities has led researchers to define different concepts of membership and neighborhood effects that rely on notions of spatial and social distance (Anselin 2003). For example, Durlauf (1996) considered the effects of residential neighborhood on education and income inequality. In these models, income inequality across individuals is determined, at least in part, by characteristics of the neighborhood in which each individual grows up; but it also depends on the characteristics of other surrounding neighborhoods. Although the former type of dependence can be easily accommodated by a structural approach, this is not the case for the latter. The presence of such interactions is more than a nuisance, and the biases introduced in parameter estimation may be of concern.

In the application chapters, the authors analyze general Bayesian methods that may accommodate the case where one believes that spatial and social effects may render the distribution of random effects non-Gaussian. Modeling

approaches with spatial dependence using empirical Bayes methods have been applied in Chapter 11 to examine the geographical distribution of diseases. This chapter analyses disease mapping and small-area estimation using count models with a spatial dependence structure for non-Gaussian random effects that are correlated with surrounding random effects. The Bayesian fitting of the model relies on sampling-based approximations to the distribution of interest via Markov chain Monte Carlo (MCMC) methods. The interested reader may find it useful to compare the results of GLLAMM with those in MLwiN, which also uses MCMC methods (Browne 2003) to estimate a conditional autoregressive distribution of the random effects (see Besag, York, and Molliè 1991).

Overall, I find the book to be an exceedingly valuable reference that would be ideal for graduate-level courses on generalized latent variable modeling. It is very straightforward to build from it a comprehensive course where the statistical section is complemented with a multidisciplinary set of easily replicated examples, because both the datasets and the software are available online. In addition, the book's impressive breadth and depth make it an essential reference for any researcher interested in understanding the state-of-the-art methods and potential applications in latent multilevel, longitudinal, and structural equation modeling.

Luisa CORRADO
University of Rome Tor Vergata

REFERENCES

Anselin, L. (2003), "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics," *International Regional Science Review*, 26, 153–166.

Bartholomew, D. J., and Knott, M. (1999), *Latent Variable Models and Factor Analysis* (2nd ed.), London: Arnold.

Besag, J., York, J., and Molliè, A. (1991), "Bayesian Image Restoration With Two Applications in Spatial Statistics," *Annals of the Institute of Statistical Mathematics*, 43, 1–59.

Bollen, K. A. (1989), *Structural Equations With Latent Variables*, New York: Wiley.

Brock, W. A., and Durlauf, S. N. (2001), "Interactions-Based Models," in *Handbook of Econometrics*, Vol. 5, eds. J. J. Heckman and E. Leamer, Amsterdam: Elsevier, pp. 3297–3380.

Browne, W. (2003), "MCMC Estimation in MLwiN," technical report, Centre for Multilevel Modeling Institute of Education, University of London.

Durlauf, S. N. (1996), "Neighborhood Feedbacks, Endogenous Stratification and Income Inequality," in *Dynamic Disequilibrium Modeling*, eds. W. Barnett, G. Gandolfo, and C. Hillinger, Cambridge, U.K.: Cambridge University Press, pp. 505–534.

Rabe-Hesketh, S., Pickles, A., and Skrondal, A. (2004), "GLLAMM Manual," Working Paper 160, University of California Berkeley, Division of Biostatistics, available at *http://www.bepress.com/ucbbiostat/paper160.*

Wansbeek, T., and Meijer, E. (2000), *Measurement Error and Latent Variables in Econometrics*, Amsterdam: North-Holland.

Microarray Quality Control.

Wei ZHANG, Ilya SHMULEVICH, and Jaakko ASTOLA. Hoboken, NJ: Wiley, 2004. ISBN: 0-471-45344-7. xii + 136 pp. $69.95.

This book is 136 pages long with six chapters and an index. The authors provide a concise description and discussion of the sorts of problems that arise when making and using in-house microarrays. These are typically spotted cDNA arrays, but long-oligo arrays are also discussed. The discussion is concise, but is accessible to nonexperts.

The treatment is more in terms of quality assessment, for which the authors provide much invaluable advice. The first chapter considers quality assessment of the biological samples; subsequent chapters consider issues such as array construction, hybridization, image capture, and, finally, some data-analytic issues. I found the writing to be of high quality and the discussion easy to follow.

The figures are good and appropriate. A set of color plates is also provided. Unfortunately, the authors have retained the rather unpleasant red–green color scheme that is so prevalent (but not helpful to a sizeable proportion of the intended audience).

Each chapter concludes with a list of additional references that can be consulted for a deeper understanding. I found the coverage to be quite good and the references to be up-to-date and relevant.

There are some peculiar statements, or at least ones that seem peculiar to a statistician. For example, on page 98 the authors claim that to apply normalization, we need to assume that the genes arrayed are a simple random sample of genes from the organism being studied. This is not true, because for many (one might even claim all) organisms, we do not know the entire genome and hence cannot provide a simple random sample (and the ones we do know about are not simple random samples either). Moreover, we can normalize arrays even when the genes are not a simple random sample. However, further reading reveals the authors' concerns, which are entirely valid and should be considered. So, while reading this volume, it is important to try to understand the authors' points, even when their descriptions seem incorrect.

I recommend this book to any statistician interested in understanding some of the underlying principles and methods used to generate microarray data. This book can be very helpful if you are working in this area and do not have a good understanding of all facets of the technology. The authors alert you to many of the issues and problems that can arise and affect the experimental data. I find it a useful source of ideas and potential issues. It is important that more attention be paid to quality assessment and improvement.

Robert GENTLEMAN
Fred Hutchinson Cancer Research Center

Design and Analysis of DNA Microarray Investigations.

Richard M. SIMON, Edward L. KORN, Lisa M. MCSHANE, Michael D. RADMACHER, George W. WRIGHT, and Yingdong ZHAO. New York: Springer-Verlag, 2004. ISBN: 0-387-00135-2. x + 199 pp. $59.95.

This volume is 199 pages long, with nine chapters, three appendices, and an index. The book has six authors, but manages to maintain a consistent writing style throughout. The authors' intention is to bridge the gap between statisticians and biologists working on the design and analysis of microarray experiments; in my opinion, they have largely succeeded. The book contains some introductory material on biology and provides many nice descriptions of statistical methods and reasoning.

The weakest aspect of the book is the figures. The reproductions in the text are often of poor quality, and there is no indication that there are color plates for some of the figures. Because these plates are grouped in the middle of the volume, readers are left to find this out for themselves. The authors have chosen to continue with the red-green color scheme (which should probably be abandoned) for the presentation of microarray data.

This is a book to get your biological collaborators to read, because it provides good statistical perspectives on the design of experiments and on other methodologies needed to design and analyze microarray experiments. The treatment is straightforward, but not simple.

References are given at the end of the book and are another minor weak point. The choices seem quite odd, but this is a new and rapidly changing and diverse field, and they may reflect the authors' experience better than mine. For example, in their discussion of normalization there is no mention of the variance-stabilizing methods proposed by Durbin et al. (2002) and Huber et al. (2002).

The machine-learning portion is interesting and complete, although some minor issues arise. For example, there is no discussion of feature selection, which is widely practiced. And distance matrices come a bit too late in the discussion; they are of interest earlier (in Chap. 7).

Appendix A provides a concise description of the basic concepts of molecular biology needed to understand microarray data. Appendix B describes the datasets considered in the examples that were used through the text, and Appendix C demonstrates an analysis using software written by some of the authors of the monograph.

This book can be recommended as a good elementary reference for statisticians that want to work on microarray data. In particular, the authors provide wide-ranging coverage of topics. It is also useful for biologists that want to understand the basis for hypothesis testing and basic concepts of designing and analysing microarray experiments.

Robert GENTLEMAN
Fred Hutchinson Cancer Research Center

This text unifies the principles behind latent variable modeling, which includes multilevel, longitudinal, and structural equation models, as well as generalized mixed models, random coefficient models, item response models, factor models, panel models, repeated-measures models, latent-class models, and frailty models. Since latent-variable models are used by researchers from various disciplines with little or no cross-referencing from other disciplines, unifying these models allows readers to separate what the authors refer to as local jargon from the fundamental elements of these models. needed for nonadditive and more complex models. Chapter 2 describes five additional examples that provide the foundation for more detailed explanations in later chapters. The chapter also includes some of properties of sensitivity analysis and expands on the various settings in which sensitivity analysis is applicable. Chapter 3 presents the examples in greater detail by showing some practical problems in which sensitivity analysis has demonstrated usefulness. Â

@article{Lipovetsky2005GeneralizedLV, title={Generalized Latent Variable Modeling: Multilevel,Longitudinal, and Structural Equation Models}, author={Stan Lipovetsky}, journal={Technometrics}, year={2005}, volume={47}, pages={237 - 248} }. Stan Lipovetsky. Published 2005. Computer Science, Mathematics. Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. Chapman & Hall/CRC, Boca Raton, FL. (This is the reference textbook for understanding/working with Stata gllamm.)Â I would say it is necessary reading if you are utilizing latent variables in a longitudinal framework. It is only about the logic of latent variables though, it has nothing about actually estimating models. Do post back with your experiences, I have some of the references given here already, although I would like to expand my library as well. FWIW, Ken Bollen's Structural equations with latent variables was the next on my reading list (although that is only based on my opinion of his scholarly work). Besides that I would say I enjoy the work of Bengt MuthÃ©n as well.