BOOK REVIEW

Multilevel Analysis: Techniques and Applications

Joop Hox
Mahwah, NJ: Erlbaum, 2002

Reviewed by

Jan de Leeuw
UCLA

Target Audience

There are now quite a number of books on multilevel analysis targeted at applied researchers in the behavioral and social sciences. They are not intended primarily for methodologists, psychometricians, or statisticians. The idea is to provide information about this relatively new class of techniques to researchers who want to use them in their research.

The latest entry in the field is the book by Joop Hox, *Multilevel Analysis: Techniques and Applications*, reviewed here. Earlier competitors are the classic work by Bryk and Raudenbush (1992), now Raudenbush and Bryk (2002), and the textbooks by Snijders and Bosker (1999) and Kreft and de Leeuw (1998). The books by Longford (1993) and by Goldstein (1995), no matter how excellent, are written for an essentially different target group, with more statistical sophistication. And there are the edited books by Reise and Duan (2003) and by Leyland and Goldstein (2001). Although these books cover the same material, they cannot really be compared, because they do not speak with a single voice, and their contributions vary wildly in terms of scope and level of difficulty.

Level and Content

So how does Hox's book compare with the others? Let's first look briefly at the technical level: how much statistics and mathematics is required to follow all the details? The order of the existing books is fairly clear. Raudenbush and Bryk (2002) is the most demanding, followed by Snijders and Bosker (1999), then by Hox, and finally by Kreft and De Leeuw (1998). All an applied researcher needs to read Hox from cover to cover is some knowledge of and experience with regression and analysis of variance as it is practiced in the social and behavioral sciences. More importantly, let's now look at the content of the book, chapter by chapter. What does Hox cover? And how does this compare with the other multilevel textbooks?

1. **Introduction to multilevel analysis, 10 pages.** The motivational chapter, which is (surprisingly) very short, emphasizes some of the older sociological literature on ecological fallacies, cross-level inference, and the effect of group characteristics on individual behavior.
2. **The basic two-level regression model: Introduction, 25 pages.** This introduces the hierarchical linear model (HLM) in both single equation and multiple equation form. It discusses estimation (briefly), intra-class correla-

tion, shrinkage and Empirical Bayes, residual plots, software, and notation. This is a solid introduction, though the details can be confusing (cf., below in "Critique").

3. **Estimation and hypothesis testing in multilevel regression, 12 pages.** Maximum likelihood (ML), Bayesian estimation, and the Bootstrap are explained, although not quite correctly in some cases (cf., below in "Critique"). Sometimes the discussion is quite surreal from the statistical point of view, because statistical questions are often decided on the basis of simulation studies. The AIC and BIC are introduced, but not fully discussed.
4. **Some important methodological and statistical issues, 23 pages.** Here, Hox discusses "analysis strategy", which is a stepwise approach of building up an appropriate model. He also discusses centering and interaction. This is a nice chapter, which stays close to the data. Unfortunately, it derails a bit in its discussion of "explained variance." This leads, to what some would consider a wishful juggling of variance ratios.
5. **Analyzing longitudinal data, 30 pages.** Repeated measures analysis fits nicely into the multilevel framework and this chapter explains competently why this is the case. It analyzes some examples in detail.
6. **The logistic model for dichotomous data and proportions, 20 pages.** This chapter has a quick introduction to generalized linear models. It also discusses and compares the MQL, PQL, and numerical integration methods to fit generalized HLMs. It is informative, and has good examples. This is a very active research area in statistics at the moment, and it is impossible for a textbook to be completely up to date. Laplace expansion methods are not mentioned, for instance.
7. **Cross-classified multilevel models, 16 pages.** If there are deviations from a strictly hierarchical data structure (students nested in primary schools and in secondary schools, but primary schools crossed with secondary schools) then design matrices become more complicated. This chapter has good explanations and examples.
8. **The multilevel approach to meta-analysis, 17 pages.** Multilevel analysis, in particular as implemented in the program HLM, is one of the core techniques used in meta-analysis. The chapter describes the way meta-analysis fits into the multilevel framework, and gives examples. Reading about meta-analysis always makes me feel uncomfortable, but in spite of that I think this chapter does a good job.
9. **Multivariate multilevel models, 16 pages.** Multiple outcomes can be incorporated in multilevel analysis by adding an additional level, nested within the other levels. This is a clever trick that is explained and illustrated well in this chapter.
10. **Sample sizes and power analysis in multilevel regression, 24 pages.** Sample size determination and power analysis always have certain ritualistic aspects, with researchers willing to make the wildest assumptions and guesses to satisfy their NIH or APA reviewers. Power analysis for multiparameter models is notoriously problematic. In this chapter, Hox reviews simulation studies and the few analytical results that are available.

11. **Advanced methods of estimation and testing, 28 pages.** In this chapter, true to its daunting title, Hox gives a brief introduction to profile likelihood, Bayesian statistics, MCMC methods, and the bootstrap. And he illustrates this with examples. I am not sure if it is wise to include a chapter like this, because it is unclear if the applied researchers get enough information to work with.
12. **Multilevel factor models, 26 pages.** Here, the book goes off the beaten path. It gives a short introduction to structural equation models (SEM), and then discusses the work of Goldstein, McDonald and Muthén, extending SEM to hierarchical data. This is no longer regression analysis, and new problems arise. Also, the reviewed developments are still incomplete and quite different from the random coefficient approaches to regression analysis. Again, one wonders if it was a good idea to include multilevel SEM in a book of this type.
13. **Multilevel path models, 12 pages.** This chapter includes more work on multilevel SEM. As such, it deals with infinitely complicated multilevel models, with more and more elaborate black-box software to compute parameter estimates, standard errors, and significance tests. It shows that the sins of SEM can be repeated at multiple levels, and that applied researchers need much more greater prior knowledge to venture into these computer programs. As was the case with the previous chapter, one may argue that this chapter takes us too far astray, but one can also argue that it promotes bad habits.
14. **Latent growth curve model, 12 pages.** In this final SEM chapter, specific latent variable models for growth curves (longitudinal data) are considered. More or less the same comments apply as in the previous chapters. Given the fact that Hox has decided to include SEM, this chapter had to be included as well.

We see that the number of topics covered in this book is huge. It easily beats out Raudenbush and Bryk (2002), Snijders and Bosker (1999), and Kreft and De Leeuw (1998). As a necessary consequence, the book is "wide" rather than "deep." This is, of course, not always a problem, because Hox gives many references and those who want to go deeper can find details in the relevant literature. Discussing so many different topics at a rather superficial level, however, may not be useful to some.

Not much of the contents is based on Hox's own research; what we see is mostly a compendium of other people's results. As compendiums go, this one is done competently, although the book often does not try to explain the technical aspects of the techniques. This creates a problem, especially in the later chapters, where more and more complicated models are explained less and less adequately. Fortunately, in these later chapters Hox puts more and more emphasis on the examples.

One of the consequences of the "wide" approach taken by Hox is that some of the key concepts used in the text remain vague, presumably because they are too "technical". For example, it is not explained why maximum likelihood is the ruling method of estimation in statistics, what the precise difference is between full and restricted maximum likelihood, and why we add two times the degrees of freedom to the log-likelihood to get the AIC. This has the unfortunate corollary that some of the concepts and techniques appear as black boxes, forcing the reader's local statistician in the role of magician. Of course, many statisticians very much like to play that role, but the consensus these days is that too much of this role-playing is bad, both for the statistician and for the client.

Critique

My main concern with the book is that there are a number of places the author says things that are imprecise or incorrect. Let me give some examples from just the first three chapters.

1. It is not true that data with a hierarchical structure necessarily come from a multistage sample (pp. 5, 81). There are many sampling designs that lead to hierarchical data. In fact, Hox himself, on the same page, describes cluster sampling and stratified sampling. Multilevel techniques are even useful if we look at all pupils in all schools in the district, and we have no sample at all.
2. Random coefficient (RC) models are more general than multilevel models, and mixed linear models are again more general. Variance components models are less general than mixed linear models. But certainly these terms do not describe the same class of models (p. 11), and in fact there is nothing inherently hierarchical about mixed models or RC models or variance component models.
3. RC models do not get their name from the fact that intercepts and slopes vary (p. 12) over groups. Recent statistical literature discusses varying coefficient models in detail, in which regression coefficients are different but not necessarily random. In RC models, coefficients are modeled as random variables, and that explains their name.
4. Across all classes, regression coefficients have a distribution with some mean and variance (p.12). This is fundamentally wrong. The regression coefficients in each class are random variables and they have a distribution that is generally different for different classes. This basic error, which seems to confuse sample correlation and population correlation, is repeated in other places (p. 31).
5. The intra-class correlation is not the expected correlation between two randomly chosen units in the same group (p. 15). It is the population value of the correlation between any two units in the same group.
6. The likelihood function is not the probability of observing the sample data (p. 16). In the case of continuous data, the probability of observing the sample data is always zero.
7. The deviance does not have a chi square distribution (p. 16), but an approximate chi square distribution if the sample size is large (and in multilevel models it is not immediately obvious what sample size is).
8. With large samples, ML estimates are usually robust against misspecification (p. 37). This is certainly not true in general. ML produces consistent estimates if the residuals are non-normal, but so do OLS estimates, which are even unbiased.

9. It is never the case that the iterative ML algorithms sometimes never stop (p. 39). Because there are a number of iterative ML algorithms, Hox needs to identify the one to which he is referring. For instance, is he referring to an EM, iterative reweighted least squares, or Newton algorithm? In addition, with badly specified models and boundary solutions the convergence speed can become sublinear, but we will still have convergence (theoretically).
10. A generalized least squares algorithm is not the same thing as one step of the maximum likelihood algorithm (p. 39). Again to which algorithm is he referring? Which likelihood, for that matter? And one step from where?
11. The distinction between GEE and ML is not "tricky" (p. 40). In fact, this distinction has been very clearly explained in very many places.

Hox is clearly most at ease when he is actually analyzing and presenting examples. The book is an uneasy hybrid of an encyclopedia of other people's research and a manual for one or more computer programs. It is not clear to me why Hox wanted to write a book discussing so many elaborations of multilevel techniques, some of which are quite esoteric and tentative, especially when his strength is clearly to perform and explain actual data analyses. I can recommend the book because of these recipes and examples, and as a possible gateway to more serious literature, but not as a textbook for explaining the methodology.

A final minor point—but one that annoys me mightily, although perhaps irrationally so—is that the book looks needlessly amateurish because of the ugly typesetting, especially of in-line mathematical material. In many places, subscripts are too close or too high, symbols float above the line, and typeface is sometimes unclear (e.g., see the paragraph in the middle of page 34).

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models. Applications and data analysis methods.* Newbury Park, CA., Sage.

Goldstein, H. (1995). *Multilevel statistical models.* London: Arnold.

Kreft, I, & de Leeuw, J. (1998). *Introducing multilevel modeling.* London: Sage.

Leyland, A. H., & Goldstein, H. (Eds.). (2001). *Multilevel modelling of health statistics.* Chichester, England: Wiley.

Longford, N. T. (1993). *Random coefficient models.* Oxford, England: Clarendon Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reise, S. P., & Duan, N. (Eds.) (2003). *Multilevel modeling. Methodological advances, issues, and application.* Mahwah, NJ: Erlbaum.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling.* London: Sage.