Bayesian Item Response Modeling: Theory and Applications.

Jean-Paul Fox. New York: Springer, 2010. ISBN 978-1-4419-0741-7. xiv + 313 pp. $69.95 (H).

Item response theory is a general paradigm for the design and analysis of questionnaires measuring abilities and attitudes of individuals. Suppose for example, that a multiple choice mathematics test with k items is administered to a group of n students. Let p_{ij} denote the probability that the ith student gets the jth item correct. A item response model represents this probability as a function of the ability θ_i of the student and parameters reflecting the attributes of the jth item of the test. One popular item response model is the *two-parameter logistic model*

$$p_{ij} = \frac{\exp(a_j\theta_i - b_j)}{1 + \exp(a_j\theta_i - b_j)},$$

where a_j and b_j are respectively parameters that measure the discrimination and difficulty qualities of the jth item.

Historically, one learns about the item parameters $\{(a_j, b_j)\}$ and the student abilities $\{\theta_i\}$ by means of classical inferential methods. One can estimate the item parameters by several maximum likelihood methods and assess the goodness of fit of the model using classical methods. But there are several reasons, mentioned in the book's preface, for considering a Bayesian approach in this class of models. First, Bayesian thinking allows for a clear way of incorporating additional information into the problem. This additional information might include beliefs about locations of parameters or beliefs about parameter relationships such as exchangeability. Second, powerful simulation-based computational algorithms are currently available for fitting a wide-class of item-response models. The purpose of this text is to give an introduction to Bayesian modeling and inference for item response data.

Chapters 1, 2, and 3 provide a general introduction to the class of item response models and the Bayesian viewpoint. Chapter 1 introduces the traditional item response models with either a binary or polytomous response. This chapter gives a brief introduction to Bayesian thinking and illustrates the use of the WinBUGS program to fit a two-parameter item response model with normal priors placed on the item parameters. Chapter 2 provides a brief overview of Bayesian hierarchical models used extensively in the last part of the book. One common exchangeable model assumes that the sets of item parameters $(a_1, b_1), \ldots, (a_k, b_k)$, conditional on hyperparameters (μ_ξ, Σ_ξ), have normal(μ_ξ, Σ_ξ) distributions, and the hyperparameters are given a second-stage distribution. The posterior distributions using these hierarchical priors pose considerable computational challenges since marginal posteriors are expressible using high-dimensional integrals. Chapter 3 introduces the Bayesian MCMC computational algorithms that are fundamental in fitting item response models. The Gibbs sampling and Metropolis–Hastings algorithms, and diagnostic methods for assessing MCMC convergence are described. Bayesian hypothesis testing and model selection are introduced together with methods of computing marginal likelihoods using simulation.

Chapter 4 gives an overview of Bayesian MCMC methods for fitting item response models. For a two-parameter logistic item response model with an exchangeable prior on the item parameters, a Metropolis-within-Gibbs algorithm is described. An alternative computational approach is taken by introducing augmented latent continuous data to the estimation problem. In the case of a probit model, one can construct a MCMC algorithm for the parameters and augmented data based on Gibbs sampling. These algorithms are illustrated for fitting a item response models to TIMMS math achievement data and data from the European Social Survey. Chapter 5 provides a general discussion of Bayesian methods for assessing the fit of item response models. This chapter includes a variety of topics, such as the use of the posterior distributions of Bayesian residuals to detect outliers and the use of prior and posterior predictive measures to measure lack and fit.

The remainder of the book provides analyses of more advanced Bayesian item response models. Chapter 6 describes the situation where the respondents are grouped into larger units, say schools, and a multilevel model is used to represent the ability parameters that differ across schools. In the simplest setting, the ability of ith student at the jth school is represented as $\theta_{ij} = \beta_{0j} + \epsilon_{ij}$, where β_{0j} is a school-specific intercept, and the school intercepts $\{\beta_{0j}\}$ are assigned a normal distribution. This multilevel model is described in the general setting where the abilities for the jth school have a regression structure with parameter β_j and the parameters $\{\beta_j\}$ may also have a regression structure at the second-stage. The computational details of fitting this model by MCMC are detailed

and several applications are described. Chapter 7 considers the use of multilevel models for item parameters. Among the standard IRT assumptions, one assumes measurement invariance which measures that the discrimination and difficulty parameters are equal across populations. A multilevel model allows for differences in these item parameters over clusters of respondents. Chapter 8 considers Bayesian models for jointly modeling the ability and speed in completing the test. Multilevel models are considered where ability and speed parameters of respondents can vary across groups. Chapter 9 illustrates the use of hierarchical structures to represent randomized item response models. This class of models have obvious applications in the case where surveys are admininstered on sensitive topics.

I believe that Bayesian multilevel models are potentially very helpful in understanding how characteristics of item response models vary across groups and how item parameters can vary even within a single exam. This book clearly represents the state of the art of Bayesian thinking for modeling item response data and illustrates the use of MCMC algorithms in fitting these models together with some interesting applications. I believe this book would be useful for researchers who wish to start working in this field. Also, as the author suggests in the preface, the book would also have some use for practitioners who are primarily interested in using these Bayesian models for their own datasets. The author has a website that contains R and WinBUGS code for fitting some of the models and there are exercises included in all of the chapters.

Although this book contains much valuable description of Bayesian hierarchical modeling, the book is written in a concise style and the technical level of the book is relatively high. For example, the important chapter on Bayesian fitting of IRTs is only 40 pages, but it contains a large amount of material including a justification for estimating parameters by marginal likelihoods, a precise statement of a Metropolis-within-Gibbs algorithm for a logistic two-parameter model, the latent data representation of the model together with a detailed description of the Gibbs sampler for the two-parameter model with latent data, a discussion of the model identification issues, a discussion of the performance of the MCMC algorithms, and the generalization of the Gibbs sampler for ordinal response data. It seems the reader would need to work through the chapter exercises to get a good understanding of the MCMC algorithms. I don't believe Chapters 1 through 3 provide sufficient background for the reader who is not familiar with item response models and Bayesian modeling. To get the most out of this book, the reader should have a good understanding of classical item response models such as described in Hambleton and Swaminathan (1984), van der Linden and Hambleton (1996), and de Ayala (2009). Also I believe the reader should be familiar with Bayesian methods as contained in Gelman et al. (2003) and multilevel models (Gelman and Hill 2006). With the above cavets, I believe this book makes an important contribution in summarizing much of the important literature in Bayesian IRT and I think it will lead to future books focusing on the use and interpretation of these models from a practitioner's perspective.

<div align="right">

Jim Albert
Bowling Green State University

</div>

REFERENCES

de Ayala, R. J. (2009), *The Theory and Practice of Item Response Theory*, New York: Guilford Publications. [1220]

Gelman, A., and Hill, J. (2006), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, New York: Cambridge University Press. [1220]

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003), *Bayesian Data Analysis* (2nd ed.), Boca Raton: Chapman & Hall. [1220]

Hambleton, R., and Swaminathan, H. (1984), *Item Response Theory: Principles and Applications*, Hingham: Kluwer. [1220]

van der Linden, and Hambleton, R. (1996), *Handbook of Modern Item Response Theory*, New York: Springer. [1220]

Brownian Motion.

Peter Mörters and Yuval Peres. New York, NY: Cambridge University Press, 2010. ISBN 978-0-521-76018-8. xii + 403 pp. $65.00 (H).

The material in the book is preceded by a panegyric page telling us that this is an eagerly awaited textbook that offers a broad and deep exposition of . . . , it does a beautiful job of covering [the material] in a very user-friendly fashion, it is a modern and attractive account of one of the central topics of probability theory, it is a splendid account of the modern theory of Brownian motion, and so on. So what can this reviewer add?