Preface

Yanyan Sheng

Department of Counseling, Quantitative Methods,
and Special Education,
Southern Illinois University,
Carbondale, IL 62901, USA
Email: ysheng@siu.edu

Psychometric theories provide a framework to evaluate the psychometric properties of an instrument, such as item characteristics, test development, test-score equating, and differential function analysis. These theories rely on formulating a statistical model to specify the relationship among latent and observed variables while making certain assumptions about them.

The last two decades have seen an explosion in the popularity and use of Bayesian methods with psychometric models, largely as a result of the advances in sampling-based approaches to inference and the availability of enhanced computational technologies. Bayesian statistics, while using the prior belief to help derive the posterior distribution, offers an alternative perspective to probability and inference. It is well-suited to address the increasingly complex phenomena and problems in educational and psychological measurement in that it can effectively tackle more complex and realistic models and problems, specifically as richer sources of data continue to be available. In this sense, the traditional frequentist methods are challenged.

The seminal work by Jim Albert at Bowling Green State University in the early 1990s is generally regarded as the foundation for Bayesian psychometric (or more specifically item response) modelling. Ever since then, much research has been conducted to employ Bayesian methods in developing and estimating modern psychometric models, including item response theory (IRT) and latent class modelling. These studies demonstrated the advantages that Bayesian methods offer in psychometric modelling and call for continued efforts to develop new estimation approaches using Bayesian statistics while improving existing ones, and to carefully implement them in empirical problems that illustrate their practical appeal. *Int. J. Quantitative Research in Education*, Vol. 2, No. 3, 2015

This special issue of *Bayesian statistics in psychometrics* brings together a selection of insightful papers that focus on implementations and applications of Bayesian to psychometric models in general, and item response models more specifically, for dichotomous responses. Specifically, the issue starts with a featured introduction by Albert who provides a summary of the early important literature on Bayesian IRT modelling, an overview of current interesting Bayesian work on, e.g., fitting asymmetric IRT models, multilevel modelling, or evaluations of posterior predictive checking, and an example to showcase recent advances in Bayesian software. The paper is not intended to be a thorough review of the literature on applications of Markov chain Monte Carlo (MCMC) algorithms to IRT models [for a more thorough review on Bayesian psychometric modelling, see Levy (2009)]. Instead, it offers a complete picture about the past and present of Bayesian IRT modelling in theory and practice, and insights on its future directions. Further, the example of using an R package also makes Bayesian fitting of IRT models accessible.

The next two papers are both developing an MCMC algorithm on fitting a complex IRT model. They both involve an evaluation of the performance using Monte Carlo simulations and an illustration of the developed approach with real datasets. Sheng focuses on unidimensional IRT models with the complexity of the model lying on the number of item parameters. The Bayesian fitting algorithm is a straightforward extension of Sahu (2002), and the main contribution of the paper lies in the empirical evidence based on its simulation studies. On the other hand, Azevedo et al. focus on longitudinal IRT models with multiple groups by combining MCMC and a reversible-jump MCMC to simultaneously handle the scaling process and estimate model parameters. The longitudinal model is very complicated involving many parameters and a covariance structure, and the paper offers theoretical evidence on using MCMC for such complex models.

As Albert indicated in his introduction article, much work has been done on developing MCMC algorithms for various IRT models but little is on the use of Bayesian model diagnostic methods such as posterior predictive checking. The paper by Park et al. fills this gap by evaluating posterior predictive checking for cognitive diagnostic models (CDMs), which differ from IRT models in that the latent trait is discrete instead of continuous. This study is important given the prevalence of CDMs in the literature and the lack of empirical evaluation of posterior predictive checking. It is hoped that the paper will generate enough interests in finding suitable discrepancy measures and in evaluating the procedure with IRT models.

Although this special issue focuses on Bayesian psychometric modelling, it does not exclude itself from benefits of traditional frequentist approaches. Instead, as the last two articles have demonstrated, Bayesian fitting with IRT models can be combined with frequentist approaches in evaluating measurement aspects such as person fit and item bias. Sinharay, in this study replicating a part from de la Torre and Deng (2008), proposed a revised procedure that combines posterior predictive checking with a standardised version of the l_z statistic (a popular likelihood-based person-fit test statistic) along with the weighted likelihood estimate of person abilities. On the other hand, Yao and Li, utilising BMIRT (Yao, 2003) – software specifically for implementing Metropolis Hastings to multidimensional IRT models, developed multidimensional extensions of Lord’s chi-square test and Raju’s area measure for detecting adverse DIF (item bias). These two papers, with desired findings, mark an interesting step-forward in the IRT-based person-fit assessment and DIF literatures. They also suggest the possibility and benefit of working with Bayesian and frequentist together on a certain psychometric problem.

In general, these contributions exemplify the current development in:

1. formulating Bayesian fitting algorithms with complex IRT models using MCMC
2. empirically evaluating Bayesian model diagnostic methods with CDMs
3. combining Bayesian with frequentist in person-fit and bias detection with IRT models.

It is hoped that this special issue will serve as a good reference for future research and developments in Bayesian statistics with psychometric models, and will provide researchers and practitioners in educational and psychological measurement with theoretical developments and empirical evidence in fitting such models, which are necessary to further advance the development of Bayesian psychometric modelling.

References

de la Torre, J. and Deng, W. (2008) 'Improving person fit assessment by correcting the ability estimate and its reference distribution', *Journal of Educational Measurement*, Vol. 45, No. 2, pp.159–177.

Levy, R. (2009) 'The rise of Markov chain Monte Carlo estimation for psychometric modeling', *Journal of Probability and Statistics*, No. 537139, pp.1–18.

Sahu, S.K. (2002) 'Bayesian estimation and model choice in item response models', *Journal of Statistical Computation and Simulation*, Vol. 72, No. 3, pp.217–232.

Yao, L. (2003) *BMIRT: Bayesian Multivariate Item Response Theory [Computer Software]*, CTB/McGraw-Hill, Monterey, CA.