

Implementation and Application of Bayesian Three-Level IRT Random Intercept Latent Regression Model

Hung-Yu Huang¹ and Su-Pin Hung²

¹Department of Applied Psychology, Hsuan Chuang University

²Department of Educational Psychology and Counseling, National Taiwan Normal University

Multidimensional item response theory (MIRT) has received much attention developing into many different models. Traditionally, the standard IRT or MIRT models have two-level structure. Three-level IRT latent regression models were proposed with a MML algorithm, the predictors in IRT latent regression, however, have been assumed to be error-free. Recently, research has explored the application of a Bayesian IRT approach. This study aims to explore the Bayesian three-level IRT random intercept latent regression model (Bayesian 3L-IRT-RILRM) and assess the accuracy of its parameter recovery and efficiency. All of simulations were based on the one-parameter logistic model under the 3L-IRT-RILRM. Three tests with 20 items in each test were analyzed and 40 clusters, each containing 50 examinees, were simulated. The generated data sets were fitted to the four different models: the proposed model, the two-level latent regression model, the conventional MIRT model, and the conventional unidimensional IRT models respectively. The computer program WinBUGS with Metropolis-Hastings sampling was implemented to estimate model parameters. The Bayesian model-data fit checking techniques, such as posterior predictive model checking (PPMC), pseudo Bayes factor (PsBF) and Bayesian DIC, were implemented to choose which model was better. The results of PPMC produce an analytic index which can identify the 3L-IRT-RILRM as the best model. Furthermore the proposed model was considered best to describe the generated data through model comparison. The model parameter estimates were recovered fairly well in the framework of the Bayesian approach if the generated data was fitted to the proposed model. If the random intercept in latent regression was ignored, the parameter estimates would be biased and the precision of estimation, as well as the test reliability would be overestimated. Finally, two empirical data sets from the TEPS and BCTEST were used to illustrate the use of 3L-IRT-RILRM as the analytic model for comparison with other competitive models. 3L-IRT-RILRM is reliable and provides the most complete description of real data. Further studies and recommendations are addressed by the authors for extending more general models.

Keywords: *Bayesian inference, item response theory, latent regression, multilevel model, Rasch model*