



Assessment of Item Response Model-Data Fit Via Bayesian Limited Information Model Comparison Posterior Predictive Checks

Catherine E. Mintz, Jonathan Templin & Jihong Zhang

To cite this article: Catherine E. Mintz, Jonathan Templin & Jihong Zhang (2020) Assessment of Item Response Model-Data Fit Via Bayesian Limited Information Model Comparison Posterior Predictive Checks, *Multivariate Behavioral Research*, 55:1, 160-160, DOI: [10.1080/00273171.2019.1700772](https://doi.org/10.1080/00273171.2019.1700772)

To link to this article: <https://doi.org/10.1080/00273171.2019.1700772>



Published online: 14 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 176



View related articles [↗](#)



View Crossmark data [↗](#)



Assessment of Item Response Model-Data Fit Via Bayesian Limited Information Model Comparison Posterior Predictive Checks

Catherine E. Mintz , Jonathan Templin , and Jihong Zhang

Department of Psychological and Quantitative Foundations, University of Iowa

Model-data fit is critical to ensure valid interpretations of test scores. Recently, emphasis has been placed on Bayesian psychometric models. Despite a surge in research on Bayesian psychometric models, no research has yet been conducted on explicitly Bayesian methods of model-data fit assessment.

To assess the local fit of Bayesian models, researchers often use *posterior predictive model checking* (PPMC). In PPMC, parameter values are randomly sampled from the parameters' posterior distributions and are used to create a "predicted" data set. Next, a test statistic (e.g., correlation) is computed from the predicted data set. These two steps are repeated a given number of times. Finally, the researcher examines the posterior distribution of the test statistic as compared to the magnitude of the same statistic based on the original data. A purported advantage of Bayesian methodologies is that a range of values, rather than point estimates such as those computed from maximum likelihood estimation, are derived. However, PPMC is still steeped in frequentist ideology, as test statistics are computed and used to evaluate model-data fit.

The current study examines an explicitly Bayesian PPMC method that considers the entire posterior distribution of a target Item Response Theory (IRT) model with a limited-information saturated model. The saturated model for the target IRT model was estimated using the marginals of the item response pattern contingency tables. The limited-information

saturated model-approach was examined across multiple simulated conditions in a factorial design: data generating model type (three-parameter logistic model vs. 2-dimensional); items per factor (6 or 12), latent trait correlations ($\rho = 0$ or $\rho = .3-.8$); sample size (50, 500, or 2000); and estimation model misspecification (under-specified, correctly specified, or over-specified). Kolmogorov-Smirnov (KS) and Kullback-Leibler (KL) statistics were computed to quantify the overlap among the posterior predictive distributions for the saturated, true, and mis-specified models.

Findings suggest that the limited-information saturated model works well if the model is under-specified. The KS statistic appears more effective than traditional PPMC when the model is under-specified or correctly specified, whereas KL appears ineffective as a local fit measure. These results may indicate low sensitivity or convergence issues and require further investigation. The limited-information saturated model appears promising. However, more research is needed, specifically the development of local model-data fit statistic for over-specified models as well as global model-data fit statistics.

ORCID

Catherine E. Mintz  <https://orcid.org/0000-0002-8959-0013>

Jonathan Templin  <http://orcid.org/0000-0001-7616-0973>