

IRT models for Non-ignorable Missing Data Processes

Cees A.W. Glas

GW/MD, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

Email: Glas@edte.utwente.nl

Jonald L. Pimentel

GW/MD, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

Email: Pimentel@edte.utwente.nl

Keywords: item response theory, latent traits, missing data, non-ignorable missing data, observed covariates

Abstract

Missing data usually present special problems for statistical analyses, especially when the data are not missing at random, that is, when the ignorability principle defined by Rubin (1976) does not hold. This article presents a model-based procedure that handles non-ignorable missing data using item response theory (IRT). The relevant model for the observed data is estimated concurrently with the IRT model for the missing data process. The estimates are obtained using marginal maximum likelihood. As an example, the generalised partial credit model is used to model the observed data while the Rasch model is used to model the missing data process. Simulation studies for dichotomous and polytomous data are presented that show that the bias in the item parameter estimates obtained ignoring the missing data process can be removed or reduced by using the explicit model for the missing data process. It is shown that the IRT model for missing data can also include observed covariates. Using a simulation study, it is investigated to what extent the bias in the parameters of the observed data model can be removed first using only part of the observed covariates and second if only latent variables are included and the observed covariates are not available for estimation procedure.

References

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.