Computational Statistics and Data Analysis 103 (2016) 91–110

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Multidimensional and longitudinal item response models for non-ignorable data

Vera Lúcia F. Santos [1], Fernando A.S. Moura [a], Dalton F. Andrade [b,*], Kelly C.M. Gonçalves [a]

[a] Departamento de Estatística, Universidade Federal do Rio de Janeiro (UFRJ), Caixa Postal 68530, CEP: 21945-970, RJ, Brazil

[b] Departamento de Informática e Estatística, Universidade Federal de Santa Catarina (UFSC), CEP: 88040-900, Florianópolis, SC, Brazil

HIGHLIGHTS

- Multidimensional IRT models are proposed to non-ignorable responses in multiple-choice educational data.
- Proficiencies of the examinees and their propensities for omission are jointly modeled.
- A model for longitudinal data with non-ignorable missing item responses is also proposed.
- Application to a Brazilian multidimensional and longitudinal educational evaluation is presented.

ARTICLE INFO

Article history:
Received 8 April 2015
Received in revised form 4 May 2016
Accepted 5 May 2016
Available online 11 May 2016

Keywords:
Bayesian inference
Education evaluation
Non-ignorable missing data
MCMC

ABSTRACT

A multidimensional item response approach is proposed to model non-ignorable responses in multiple-choice educational data. The model considers latent traits related to individual proficiency as well as the propensity to answer items. Thus, in addition to modeling the probability of scoring on an item, the probability of answering it is also modeled. Simulation studies are presented to evaluate the efficiency of the estimation procedure in recovering the true values of the model parameters considering several particular cases of the dimensions of proficiency and propensity. The simulation study also compares the proposed approach with others commonly applied in practice. A further extension to cope with longitudinal data with non-ignorable missing item responses is also proposed, together with an application to a Brazilian longitudinal educational evaluation study.

1. Introduction

In modern society, tests are widely used in schools, industry and government to assess the progress of students, select individuals or verify the efficiency of educational systems. For proper extraction of information from tests, Item Response Theory (IRT) has been under development since the second half of the twentieth century. As noted by Lord (1983), IRT models the probability of an examinee's responses to test items as conditionally independent, given an unobservable ability (proficiency) parameter. The IRT approach, as originally proposed by Lord (1952) and Rasch (1960), strove to determine individual scores and their precision by analyzing each individual's item responses. Since that time, its extensive use and diffusion have led to the emergence of new methodological challenges, and consequently, new solutions have been proposed

* Corresponding author. Tel.: +55 48 37217545; fax: +55 48 37219770.
E-mail addresses: fmoura@im.ufrj.br (F.A.S. Moura), dalton.andrade@ufsc.br (D.F. Andrade), kelly@dme.ufrj.br (K.C.M. Gonçalves).

1 In Memoriam; Vera Lúcia Filgueira dos Santos was a former Ph.D. student from Departamento de Estatística, Universidade Federal do Rio de Janeiro (UFRJ).

http://dx.doi.org/10.1016/j.csda.2016.05.002
0167-9473/

to overcome them. These methodological developments intended to improve the analysis of results include omitted item response analysis (e.g., Mislevy and Wu, 1996), longitudinal data models (e.g., Andrade and Tavares, 2005), multidimensional models (e.g., Reckase, 2009) and models to cope with polytomous responses (e.g., van der Linden and Hambleton, 1997). This work is motivated by the need to evaluate models that consider the possibility that examinees may not answer all items.

The omitted item responses that arise from tests that have alternative forms, targeted tests, tests developed in multiple steps, adaptive tests and tests with a time limit can probably be attributed to the administrators of the test. Therefore, they are due to the design or nature of the application. In these situations, the omissions might be considered ignorable and do not need to be considered in the estimation of the model parameters. In other situations, missing responses could be entirely attributable to an examinee's decision not to answer one or more items despite having time to appraise them. In such cases, the omissions cannot be considered ignorable. According to Lord (1983), the probability of an examinee omitting an item can depend on both ability and temperament. Therefore, it is reasonable to include in the model a latent parameter that is jointly related to the proficiency and governs the probability of answering an item. This parameter is referred to here as the propensity to answer an item.

There is a vast IRT literature concerning how to cope with missing item responses when a test is given once. Lord (1974) proposes to assign the score of an omitted item, in a multiple-choice test, a value equal to the inverse of the number of alternatives, assuming that the omission was intentional. Ludlow and O'Leary (1999) compare the proficiency estimate obtained when the non-observed item response is treated as an incorrect answer (assuming intentional omission) with that obtained when the missing response is assumed to be due to a lack of time. Patz and Junker (1999) propose to treat omitted items as unknown parameters to be estimated in the Markov Chain Monte Carlo (MCMC) algorithm. Moustaki and O'Muircheartaigh (2000) propose a latent trait model to obtain information about attitude for nominal data, treating omission as an additional category.

Holman and Glas (2005) and Pimentel (2005) propose a different method of approaching the missing data issue. They model the joint probability of scoring/non-scoring on and not missing/missing an item as a function of the item features, the examinee's ability and his propensity to omit an item. Holman and Glas (2005) propose an IRT model that includes parameters related to the indicator variables of scoring on and missing an item. They also present a simulation study in which they evaluate the effect of omission on the estimation of these parameters via maximum likelihood. An application with a real dataset is also presented. Pimentel (2005), in turn, extends the approach of Holman and Glas (2005) to polytomous items.

The model proposed in this paper can be viewed as an extension of the approach of Holman and Glas (2005) because it also addresses non-ignorable omission by introducing an indicator of an observation variable; however, differently from those authors, we allow the proficiencies of the examinees and their propensities for omission to be multidimensional. Moreover, all inference is performed using a full Bayesian approach, and the correlation between the proficiencies and propensities is assumed to be unknown. A further extension to cope with longitudinal data with non-ignorable missing item responses is also proposed, together with an application to a Brazilian longitudinal educational evaluation study.

This article is organized as follows. Section 2 describes the proposed multidimensional model for coping with non-ignorable missing item responses and its extension to the longitudinal case. Section 3 describes the prior distributions used and discusses several important computational issues related to estimating and sampling from the posterior distributions of the parameters of interest. Section 4 presents a simulation study addressing the special case in which both the proficiency and propensity vectors are unidimensional. This simulation study was performed to evaluate the efficiency of the Bayesian estimation procedure in recovering the true values of the model parameters under scenarios different from that investigated by Holman and Glas (2005). In Section 5, we report a further simulation study conducted to assess the effect of a missing-data process for the case in which the proficiency vector has two components. In Section 6, we present an additional simulation study of the case in which both the proficiency and propensity vectors are bi-dimensional. Section 7 presents a real application to a Brazilian longitudinal educational dataset, including the results of an analysis comparing the proposed approach with two others that are commonly applied in practice. Section 8 offers conclusions and suggestions for further research. An Appendix contains further details of the MCMC algorithm used.

2. Multidimensional models with missing item responses

Suppose that a multiple-choice test with I items is administered to J examinees. Let $\boldsymbol{Y}_{..}$ be a matrix of dimension $J \times I$ with elements denoted by Y_{ij} such that $Y_{ij} = 1$ if the jth examinee scores on the ith item and $Y_{ij} = 0$ otherwise. Analogously, we define the indicator matrix $\boldsymbol{R}_{..}$, with entries of $R_{ij} = 1$ if Y_{ij} is observed and $R_{ij} = 0$ when Y_{ij} is not observed for $i = 1, \ldots, I$ and $j = 1, \ldots, J$. As in a multidimensional IRT model, the proficiency vector of the jth examinee is denoted by $\boldsymbol{\theta}_{1j.} = (\theta_{1j1}, \ldots, \theta_{1jM})'$, where M represents its dimension. In what follows, we call the latent Q-vector $\boldsymbol{\theta}_{2j.} = (\theta_{2j1}, \ldots, \theta_{2jQ})'$ the "propensity vector" for the jth examinee to answer an item. Thus, the vector of all latent parameters for the jth examinee is denoted by $\boldsymbol{\theta}_{.j.} = (\boldsymbol{\theta}_{1j.}, \boldsymbol{\theta}_{2j.})'$ and is of dimension $M + Q$. Furthermore, let $\boldsymbol{\eta}_i$ and $\boldsymbol{\zeta}_i$ be the model parameters related to the ith item, which are associated with the distributions of $\boldsymbol{Y}_{..}|\boldsymbol{R}_{..}$ and $\boldsymbol{R}_{..}$, respectively. According to Holman and Glas (2005), one possibility is to consider the following model:

$$M_1 = \prod_{i,j} P(Y_{ij}|R_{ij}, \boldsymbol{\theta}_{1j.}, \boldsymbol{\eta}_i) P(R_{ij}|\boldsymbol{\theta}_{2j.}, \boldsymbol{\zeta}_i) f(\boldsymbol{\theta}_{1j.}, \boldsymbol{\theta}_{2j.}|\boldsymbol{\nu}). \tag{1}$$