RT-MAE 2007-09

LATENT TRAIT ESTIMATION IN NOMINAL RESPONSE MODEL: LATENT TRAIT ASYMMETRY AND HIERARCHICAL AND EMPIRICAL FRAMEWORK

Caio Lucidius Naberezny Azevedo
and
Dalton Francisco de Andrade

Palavras-Chave: Latent trait, nominal response model, dichotomous models, bayesian estimation, asymmetry distribution.

Classificação QA277.24.

- Março de 2007 -

Latent Trait Estimation in Nominal Response Model : Latent Trait Asymmetry and Hierarchical and Empirical Framework

Caio Lucidius Naberezny Azevedo* and Dalton Francisco de Andrade†

Abstract

The Bock's Nominal Response Model (Bock, 1972) was proposed to improve the **latent trait (ability)** estimation. In this article we discuss the five of most used latent trait estimation methods: Maximum Likelihood (MV) (Baker and Kim, 2004), three bayesian methods, the Expectation a Posteriori (EAP), the Expectation a Posteriori through Monte Carlo integration (EAPMC) and Mode a Posteriori (MAP), that we will call as Classical Bayesian Procedures (CBP) and Metropolis-Hastings algorithm within Gibbs Sampling, that will name MCMC + Metropolis Hastings algorithm (MCMC + MH). We review the MC and EAP methods, propose the EAP through monte carlo integration, develop a suitable modification in MAP and present three schemes of simulation for the MCMC. To illustrate these estimation methods we conduct appropriate simulations. We assume that the item parameters are known or have been estimated by some appropriate method.

Key words : Latent trait, nominal response model, dichotomous models, bayesian estimation, asymmetry distribution.

Classificação AMS : QA 277.24.

1 Introduction

The Bock's Nominal Response Model (Bock, 1972) was proposed to improve the **latent trait (ability)** estimation. Frequently, the MV and CBP procedures have been used to estimative latent traits in IRT models. However, these methods may not produce good results. In counterpoint, the MCMC + MH approach constitutes an interesting alternative because its flexibility, even though it demands a great computational effort. These methods are applied in the latent trait estimation, considering the item parameters known or not. In this article, we focus on the situation where we know the true value of the item parameters. This a common situation, for example, when it is used items from an item baking.

In next section we present the model and introduce some useful notations. In Section 3 we discuss the estimation methods considered in this work, in Section 4 we conducted appropriate simulations and in Section 5 we give some conclusions. The Appendix brings some algebra and results related to the simulation process.

2 The Model

The Nominal Response Model (NRM), see Bock (1997) for example, has the following form

*Department of Statistics, Institute of Math and Statistics - USP / Brazil, *c_naber@yahoo.com.br*

†Departament of Informatic and Statistics - Federal University of Santa Catarina - Brazil, *dandrade@ine.ufsc.br*

$$P_{ijs} = P\left(Y_{ijs} = 1|\theta_j, \zeta_i\right) = \frac{e^{a_{is}(\theta_j - b_{is})}}{\sum_{h=1}^{m_i} e^{a_{ih}(\theta_j - b_{ih})}}, \tag{1}$$

where Y_{ijs}, is the random variable that assumes value 1, if subject j, $j = 1, 2, \ldots, n$ chooses the alternative s, $s = 1, 2, \ldots, m_i$ of item i, $i = 1, 2, \ldots, I$ and 0, otherwise, a_{is} e b_{is} represent the parameters related to the discrimination and the difficult of the category, respectively. In this model we may have negative values for both type of parameters. For the parameter discrimination, we expect negative, or small values, for the wrong alternatives and positive value for the right alternative. This means that higher value for individual ability is associated with higher probability of this subject chooses the right alternative. The difficult parameter represents, in some way, the ability that a subject must have to chooses the referred alternative.

Notice that, the two parameter logistic model (see Baker and Kim, 2004), is not a particular case of the NRM considering only two categories.

This model has three basically assumptions that are essential to the estimation processes;

1. The responses of different subjects are independent,
2. Given the latent trait, the response of the same subject to different items are independent (conditional independency),
3. The probability that a subject chooses an alternative of a specific item can be modelled by the multivariate Bernoulli model, that is,

$$P(\boldsymbol{Y}_{ij.} = \boldsymbol{y}_{ij.}|\theta, \zeta_i) \equiv P(\boldsymbol{Y}_{ij.} = \boldsymbol{y}_{ij.}) = \prod_{s=1}^{m_i} P_{ijs}^{y_{ijs}},$$

where $\boldsymbol{y}_{ij.} = (y_{ij1}, \ldots, y_{ijm_i})^t$ represents a specific set of responses of a subject to a specific item.

In the next section we will discuss the estimation procedures.

3 Latent trait Estimation

In this section we discuss and present the latent traits estimation procedures. We review the Maximum Likelihood, suggest modifications in the CBP procedures and present the MCMC simulation trough three schemes of drawing from posterior densities.

In the latent trait estimation, using ML and CBP, one always consider the item parameter know and we introduce the estimatives (or the true values) in the original likelihood and then it uses this function to estimative them. This generates a kind of profile likelihood (Van Der Vaart and Murphy, 2000) and (Fraser and Reid, 1989), that not necessarily leads to the same estimatives obtained from the original likelihood (Baker and Kim, 2004). Furthermore we use only the estimatives of item parameters and not their standard errors. Another important point is that, to estimative the item parameters we need to consider a (prior) latent distribution that, in general is set to a standard normal (Baker and Kim, 2004), or an empirical distribution estimated by the so-called Nonparametric Maximum Likelihood Estimation (Mislevy, 1984), for example. This may leads to poor estimatives of item parameters and consequently to poor latent trait estimatives. Or even, if the true latent distribution is different from the distribution used in the estimation processes, the latent estimatives may not be so good.

Even in the CBP procedure, the prior distribution is set to a standard normal and the quantities adopted are commonly the maximum (MAP) or the expectation (EAP) values. In the MAP method, one may use a different prior because it is not so complicated to calculate the needed derivates. However, in the EAP, we need to use quadrature points that in general, are not available to asymmetric distributions. Furthermore, it is not possible to evaluate the posterior distribution of latent traits, only some quantities. The classical bayesian methods discussed above are not appropriate to take into latent asymmetry distribution account.

In counterpoint, the MCMC + Metropolis-Hastings simulation proposed by Patz and Junker (1999) is a straightforward method that uses a bayesian approach in item response theory. Particulary, one may choose to estimative simultaneously the item parameters and latent traits, to consider a wide set of priors and to use different quantities to estimative the interesting parameters. In next subsection we will presente/review the estimation methods.

3.1 Maximum Likelihood Estimation

As described in Baker and Kim (2004) and Azevedo (2003), we have the following log-likelihood to an individual latent trait

$$l(\theta_j, \widehat{\Gamma}) \equiv l(\theta_j) = \sum_{i=1}^{I}\sum_{s=1}^{m_i} y_{ijs} \ln P_{ijs} ,$$

where $\widehat{\Gamma}$ represents an estimative for the item free parameters (Azevedo, 2003). After some algebra, we obtain the score function,

$$S(\theta_j) = \sum_{i=1}^{I} \alpha_i^t T_i \left[y_{ij.} - P_{ij.} \right] , \quad (2)$$

and Hessian Matrix

$$H(\theta_j) = -\sum_{i=1}^{I} \left\{ \alpha_i^t T_i W_{ij} T_i^t \alpha_i \right\} , \quad (3)$$

where α_i, T_i and W_i are convenient matrices (Azevedo, 2003). We may notice, by (3), that the Hessian Matrix is non-stochastic, so the Information Matrix is only $I(\theta_j) = -H(\theta_j)$. Therefore, considering $\widehat{\theta}_j$ an estimative of θ_j in iteration t, we can define the Newton-Raphson / Fisher Scoring method (Baker and Kim, 2004) as

Newton-Raphson

$$\widehat{\theta}_j^{(t+1)} = \widehat{\theta}_j^{(t)} - H\left(\widehat{\theta}_j^{(t)}\right)^{-1} S\left(\widehat{\theta}_j^{(t)}\right) , \quad (4)$$

Fisher Scoring

$$\widehat{\theta}_j^{(t+1)} = \widehat{\theta}_j^{(t)} + I\left(\widehat{\theta}_j^{(t)}\right)^{-1} S\left(\widehat{\theta}_j^{(t)}\right), \tag{5}$$

$t = 1, 2, \ldots$, up to achieve an convergence criteria.

In the next subsection we present the modal and expectation bayes estimation.

3.2 Bayes Modal Estimation

In a general way, in the bayes estimation, we need to use the posterior distribution which follows directly from the **Bayes Theorem** (Bernardo and Smith, 1998),

$$\begin{aligned} g_j^*(\theta_j) \equiv Cg(\theta_j|y_{.j.}, \widehat{\Gamma}, \eta) &= C\, P(Y_{.j.}|\theta_j, \widehat{\Gamma})g(\theta_j|\eta) \\ &\propto P(Y_{.j.}|\theta_j, \widehat{\Gamma})g(\theta_j|\eta), \end{aligned} \tag{6}$$

where $Y_{.j.} = (Y_{1j1}, \ldots, Y_{1jm_i}, \ldots, Y_{Ij1}, \ldots, Y_{Ijm_i})^t$, $P(Y_{.j.}|\theta_j, \widehat{\Gamma})$ is the profile likelihood, $g(\theta_j|\eta)$ is a convenient prior, η are the **hyper parameters** (or the populational parameters) and C is a normalization constant. To evaluate the **bayes modal** we need to maximize equation (6). Then taking its natural logarithm we have

$$\ln g_j^*(\theta_j) \equiv l_j^*(\theta_j) \;=\; l(\theta_j) + \ln g(\theta_j|\eta) + const. \tag{7}$$

Differentiating (7), we have the **bayesian estimating equation**

$$S(\theta_j)_B \;=\; \frac{\partial l_j^*(\theta_j)}{\partial\theta_j} = \frac{\partial l(\theta_j)}{\partial\theta_j} + \frac{\partial \ln g(\theta_j|\eta)}{\partial\theta_j}. \tag{8}$$

Notice that the first term in the right-hand side of (8) is exactly the **score function** defined in (2) and, considering a skew-normal distribution, $\theta_j|\eta_j \sim SN(\mu_{\theta_j}, \sigma^2_{\theta_j}, \lambda_{\theta_j})$ (Genton, 2004) it follows that

$$S(\theta_j)_B \;=\; S(\theta_j) - \frac{\theta_j - \mu_\theta}{\sigma^2_\theta} + \left\{\Phi\left[\lambda_{\theta_j}\left(\frac{\theta_j - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\right\}^{-1} \phi\left[\lambda_{\theta_j}\left(\frac{\theta_j - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\frac{\lambda_{\theta_j}}{\sigma_{\theta_j}},$$

where $\eta_j = (\mu_{\theta_j}, \sigma^2_{\theta_j}, \lambda_{\theta_j})^t$, μ_{θ_j} is the mean, $\sigma^2_{\theta_j}$ is the variance, λ_{θ_j} is the asymmetry parameter, $\phi(.)$ and $\Phi(.)$ are the density and cumulative functions of a standard normal distribution. Consequently, the hessian matrix and the Fisher information are given by

Hessian Matrix

$$H(\theta_j)_B \;=\; H(\theta_j) - \left\{\frac{1}{\sigma^2_\theta} + h(\lambda_{\theta_j}, \theta_j)\right\},$$

where,

$$h(\lambda_{\theta_j},\theta_j) = \left\{\Phi\left[\lambda_{\theta_j}\left(\frac{\theta_j-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\right\}^{-2}\phi\left[\lambda_{\theta_j}\left(\frac{\theta_j-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\left(\frac{\lambda_{\theta_j}}{\sigma_{\theta_j}}\right)^2 \times \left(\Phi\left[\lambda_{\theta_j}\left(\frac{\theta_j-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\frac{\lambda_{\theta_j}}{\sigma_{\theta_j}}(\theta_j-\mu_{\theta_j})+\phi\left[\lambda_{\theta_j}\left(\frac{\theta_j-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\right), \quad (9)$$

Fisher Information

$$I(\theta_j)_B = I(\theta_j) + \frac{1}{\sigma_\theta^2} + h(\lambda_{\theta_j},\theta_j),$$

where $H(\theta_j)$ is given by (3). The Fisher information is only the symmetric value of hessian matrix. The iterative process may be applied as described in (4) and (5). In the next subsection we present the bayes expectation estimator.

3.3 Bayes Expectation

From (6) we have that the bayes expectation is given by

$$\mathbb{E}\left[\theta_j|y_{.j.},\widehat{\Gamma},\eta\right] = \frac{\int_{\mathbb{R}}\theta P(Y_{.j.}|\theta,\widehat{\Gamma})g(\theta|\eta)d\theta}{\int_{\mathbb{R}}P(Y_{.j.}|\theta,\widehat{\Gamma})g(\theta|\eta)d\theta}. \quad (10)$$

Generally, the integrals in (10) do not have an explicit form and then they must be solved by some numerical method of integration (Robert and Casella, 1999). Again, considering a $SN(\mu_\theta,\sigma_\theta^2,\lambda_\theta)$ with $\lambda_{\theta_j}=0$, i.e., a normal distribution, the integrals may be solved by Gauss-Hermite method (Stroud and Secrest, 1966), or generating quadrature points considering any value to λ_θ. So (10), in terms of the quadrature points, becomes

$$\begin{aligned}\mathbb{E}\left[\theta_j|y_{.j.},\widehat{\Gamma},\eta\right] &\approx \mathbb{E}\left[\overline{\theta}_j|y_{.j.},\widehat{\Gamma},\eta\right] = \frac{\sum_{l=1}^q\overline{\theta}_l P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})g(\overline{\theta}_l|\eta)}{\sum_{l=1}^q P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})g(\overline{\theta}_l|\eta)}\\ &= \frac{\sum_{l=1}^q\overline{\theta}_l P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})A_l}{\sum_{l=1}^q P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})A_l},\end{aligned}$$

where $\overline{\theta}_l$ and A_l, are the quadrature points and the quadrature weights, respectively. A measure of precision of the EAP estimation is given by the Variance a Posteriori (VAP) that has the following form

$$Var\left[\overline{\theta}_j|y_{.j.},\widehat{\Gamma},\eta\right] = \frac{\sum_{l=1}^q\left\{\overline{\theta}_l-\mathbb{E}\left[\overline{\theta}_j|y_{.j.},\widehat{\Gamma},\eta\right]\right\}^2P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})A_l}{\sum_{l=1}^q P(Y_{.j.}|\overline{\theta}_l,\widehat{\Gamma})A_l}.$$

We need to point that for both MAP and EAP methods we used a asymmetry sample estimative for λ_θ using the observed scores. To generate quadrature points considering any value for λ_θ, it was written a R-function that generates quite similar values to them used for Bilog (Mislevy and Bock, 1990) when $\lambda_\theta = 0$.

On the other hand, if we consider a general Skew-Normal distribution, we may use the Monte Carlo Integration (Robert and Casella, 2000) to solve the integral (10), using the following scheme:

1. Generate m random Uniform (-1,1) numbers, say $\overline{\psi}_1, \ldots, \overline{\psi}_m$.

2. For $r = 1, \ldots, m$, calculate

$$P(Y_{.j.} | \ln(\overline{\Delta}_r), \widehat{\Gamma}) = \prod_{i=1}^{I} \prod_{s=1}^{m_i} P^{*}_{ijs} \text{ , and} \quad (11)$$

$$g(\ln(\overline{\Delta}_r)|\eta), \quad (12)$$

where P^{*}_{ijs} in (11) is given by (1) replacing θ_j by $\ln(\overline{\Delta}_r)$, (12) is a skew-normal density evaluated in the $\ln(\overline{\Delta}_r)$ and $\overline{\Delta}_r = \frac{2}{\overline{\psi}_r - 1} + 1$.

3. Then, evaluates

$$\overline{I}_1 = \sum_{r=1}^{m} \left\{ \frac{2}{|\overline{\Delta}_r(\overline{\psi}_r - 1)^2|} \right\} \left\{ \ln(\overline{\Delta}_r) P(Y_{.j.} | \ln(\overline{\Delta}_r), \widehat{\Gamma}) g(\ln(\overline{\Delta}_r)|\eta) \right\},$$

$$\overline{I}_2 = \sum_{r=1}^{m} \left\{ \frac{2}{|\overline{\Delta}_r(\overline{\psi}_r - 1)^2|} \right\} \left\{ P(Y_{.j.} | \ln(\overline{\Delta}_r), \widehat{\Gamma}) g(\ln(\overline{\Delta}_r)|\eta) \right\},$$

$$E[\overline{\psi}_j | y_{.j.}, \widehat{\Gamma}, \eta] = E[\overline{\psi}_j] = \frac{\overline{I}_1}{\overline{I}_2}, \quad (13)$$

$$\overline{I}_3 = \sum_{r=1}^{m} \left\{ \ln[\overline{\Delta}_r] - E[\overline{\psi}_j] \right\}^2 \left\{ \frac{2}{|\overline{\Delta}_r(\overline{\psi}_r - 1)^2|} \right\} \left\{ P(Y_{.j.} | \ln(\overline{\Delta}_r), \widehat{\Gamma}) g(\ln(\overline{\Delta}_r)|\eta) \right\},$$

$$Var[\overline{\psi}_j | y_{.j.}, \widehat{\Gamma}, \eta] = \frac{\overline{I}_3}{\overline{I}_2}. \quad (14)$$

So, the EAP and VAP estimators through Monte Carlo are given by (13) and (14).
In the next section, we present the MCMC + MH algorithm.

3.4 MCMC + MH algorithm

The MCMC's methods basically consist of simulate random variables through stationary distributions of markov chains (Geman and Geman, 1984). Particulary, in the IRT context, the MCMC + Metropolis - Hastings approach is straightforward and efficient to simulate posterior densities (Patz and Junker, 1999).

We consider three schemes of drawing the posterior distributions (bayesian inference) : classical, empirical and hierarchical ones (Bernardo and Smith, 1998). The first consists basically of choosing suitable priors and keep constant their hyper parameters. In the second scheme we consider hyper priors with the hyper parameters draw from their hyper posteriors, apart from their respective parameters, and then, we consider these simulate values in the posterior densities. From these **empirical posterior densities** we draw values for the parameters of interest. Finally, in the third one, we draw from a joint posterior density of **item parameters and their hyper parameters** and take only the values of the item parameters. We may note that, the second and the third one scheme are different from the hierarchical procedure proposed in Kim et al (1994).

For this purpose we consider, again, a skew-normal density (Genton, 2004), and μ_{θ_j}, $(\sigma^2_{\theta_j}, \psi_{\theta_j})^t$ and $(\lambda_{\theta_j}, \nu_{\theta_j})^t$ as the mean, variance and asymmetry parameter for the prior and kernel densities.

The schemes are described below.

1. **Classical Scheme**

   To simulate $\theta_j^{(t)} \sim g(\theta_j|\widehat{\Gamma}, y_{...})$ (full conditional), for $j = 1, \ldots, n$, considering the prior $\theta_j|\eta_j \sim SN(\mu_{0j}, \sigma^2_{\theta_j}, \lambda_{\theta_j})$, where, $\eta_j = (\mu_{\theta_j}, \sigma^2_{\theta_j}, \lambda_{\theta_j})^t$, through:

   (a) To simulate $\theta_j^{(*)}|(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j}) \sim SN(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j})$.

   (b) To evaluate the vector of accepting probabilities $\theta_j^{(t)} = \theta_j^{(*)}$, $\pi_{\theta_j} = \pi_j\left(\theta_j^{(t-1)}, \theta_j^{(*)}\right) = min\{R_{\theta_j}, 1\}$, where

$$R_{\theta_j} = \frac{L(\theta_j^{(*)}, \widehat{\Gamma}) \exp\left\{-\frac{\left(\theta_j^{(*)} - \mu_{\theta_j}\right)^2}{2\sigma^2_{\theta_j}}\right\} \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(*)} - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\} \Phi\left\{\nu_{\theta_j}\left(\frac{\theta_j^{(*)} - \theta_j^{(t-1)}}{\psi_{\theta_j}}\right)\right\}}{L(\theta_j^{(t-1)}, \widehat{\Gamma}) \exp\left\{-\frac{\left(\theta_j^{(t-1)} - \mu_{\theta_j}\right)^2}{2\sigma^2_{\theta_j}}\right\} \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(t-1)} - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\} \Phi\left\{\nu_{\theta_j}\left(\frac{\theta_j^{(t-1)} - \theta_j^{(*)}}{\psi_{\theta_j}}\right)\right\}}. \quad (15)$$

   (c) To accept every $\theta_j^{(t)} = \theta_j^{(*)}$ with probability π_{θ_j}, otherwise $\theta_j^{(t)} = \theta_j^{(t-1)}$.

2. **Empirical Scheme**

   To simulate $\left(\theta_j^{(t)}, \nu_{\theta_j}^{(t)}\right)^t \sim g\left((\theta_j, \nu_{\theta_j})^t|\widehat{\Gamma}, y_{...}\right)$ (full conditional), for $j = 1, \ldots, n$, considering the prior $\theta_j|\eta_j \sim SN\left(\mu_{\theta_j}, \sigma^2_{\theta_j}, \lambda_{\theta_j}\right)$ and $\nu_{\theta_j}|(\mu_{\nu_j}, \sigma^2_{\nu_j}, \lambda_{\nu_j}) \sim SN(\mu_{\nu_j}, \sigma^2_{\nu_j}, \lambda_{\nu_j})$ through:

   2.1 Latent Trait Asymmetry hyper parameters
   Different Asymmetry Parameters

   (a) To simulate $\nu_{\theta_j}^{(*)}|(\nu_{\theta_j}^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j}) \sim SN(\nu_{\theta_j}^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j})$.

(b) To evaluate the vector of accepting probabilities $\nu_j^{(t)} = \nu_j^{(*)}$, $\pi_{\nu_{\theta_j}} = \pi_j\left(\nu_{\theta_j}^{(t-1)}, \nu_{\theta_j}^{(*)}\right) = min\left\{R_{\nu_{\theta_j}}, 1\right\}$,

$$R_{\nu_{\theta_j}} = \frac{\exp\left\{\frac{-\left(\nu_{\theta_j}^{(*)}-\mu_{\nu_j}\right)^2}{2\sigma_{\nu_j}^2}\right\}\Phi\left[\lambda_{\nu_j}\left(\frac{\nu_{\theta_j}^{(*)}-\mu_{\nu_j}}{\sigma_{\nu_j}}\right)\right]}{\exp\left\{\frac{-\left(\nu_{\theta_j}^{(t-1)}-\mu_{\nu_j}\right)^2}{2\sigma_{\nu_j}^2}\right\}\Phi\left[\lambda_{\nu_j}\left(\frac{\nu_{\theta_j}^{(t-1)}-\mu_{\nu_j}}{\sigma_{\nu_j}}\right)\right]}$$

$$\times \frac{\left\{\Phi\left[\nu_{\theta_j}^{(*)}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}{\left\{\Phi\left[\nu_{\theta_j}^{(t-1)}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}.$$

(c) To accept every $\nu_{\theta_j}^{(t)} = \nu_{\theta_j}^{(*)}$ with probability $\pi_{\nu_{\theta_j}}$, otherwise $\nu_{\theta_j}^{(t)} = \nu_{\theta_j}^{(t-1)}$.

Same asymmetry parameters

(a) To simulate $\nu_\theta^{(*)}|(\nu_\theta^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j}) \sim SN(\nu_\theta^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j})$.

(b) To evaluate the vector of accepting probabilities $\nu^{(t)} = \nu^{(*)}$, $\pi_{\nu_\theta} = \pi_j\left(\nu_\theta^{(t-1)}, \nu_\theta^{(*)}\right) = min\left\{R_{\nu_\theta}, 1\right\}$,

$$R_{\nu_\theta} = \frac{\exp\left\{\frac{-\left(\nu_\theta^{(*)}-\mu_\nu\right)^2}{2\sigma_\nu^2}\right\}\Phi\left[\lambda_\nu\left(\frac{\nu_\theta^{(*)}-\mu_\nu}{\sigma_{\nu_j}}\right)\right]\left\{\prod_{j=1}^{n}\Phi\left[\nu_\theta^{(*)}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}{\exp\left\{\frac{-\left(\nu_\theta^{(t-1)}-\mu_\nu\right)^2}{2\sigma_\nu^2}\right\}\Phi\left[\lambda_\nu\left(\frac{\nu_\theta^{(t-1)}-\mu_\nu}{\sigma_\nu}\right)\right]\left\{\prod_{j=1}^{n}\Phi\left[\nu_\theta^{(t-1)}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}. \quad (16)$$

(c) To accept every $\nu_\theta^{(t)} = \nu_\theta^{(*)}$ with probability π_{ν_θ}, otherwise $\nu_\theta^{(t)} = \nu_\theta^{(t-1)}$.

2.2 Latent traits

(a) To simulate $\theta_j^{(*)}|(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j}^{(t)}) \sim SN(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j}^{(t)})$.

(b) To evaluate the vector of accepting probabilities $\theta_j^{(t)} = \theta_j^{(*)}$, $\pi_{\theta_j} = \pi_j\left(\theta_j^{(t-1)}, \theta_j^{(*)}\right) = min\left\{R_{\theta_j}, 1\right\}$, where R_{θ_j} is given by (15).

(c) To accept every $\theta_j^{(t)} = \theta_j^{(*)}$ with probability π_{θ_j}, otherwise $\theta_j^{(t)} = \theta_j^{(t-1)}$.

3. **Hierarchical Scheme**

To simulate $\left(\theta_j^{(t)}, \nu_{\theta_j}^{(t)}\right)^t \sim g\left((\theta_j, \nu_{\theta_j})^t|\widehat{\Gamma}, y_{...}\right)$ (full conditional), for $j = 1, \ldots, n$, considering the prior $\theta_j|\eta_j \sim SN\left(\mu_{\theta_j}, \sigma_{\theta_j}^2, \lambda_{\theta_j}\right)$ and $\nu_{\theta_j}|(\mu_{\nu_j}, \sigma_{\nu_j}^2, \lambda_{\nu_j}) \sim SN(\mu_{\nu_j}, \sigma_{\nu_j}^2, \lambda_{\nu_j})$, through:

3.1 **Different asymmetry parameters**

(a) To simulate $\nu_{\theta_j}^{(*)}|(\nu_{\theta_j}^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j}) \sim SN(\nu_{\theta_j}^{(t-1)}, \psi_{\nu_j}, \nu_{\nu_j})$ and $\theta_j^{(*)}|(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j}^{(*)}) \sim SN(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_{\theta_j}^{(*)})$.

(b) To evaluate the vector of accepting probabilities $(\theta_j^{(t)}, \nu_{\theta_j}^{(t)})^t = (\theta_j^{(*)}, \nu_{\theta_j}^{(*)})^t$, $\pi_{(\theta_j,\nu_{\theta_j})} = \pi_j\left[(\theta_j^{(t-1)}, \nu_{\theta_j}^{(t-1)}), (\theta_j^{(*)}, \nu_{\theta_j}^{(*)})\right] = min\left\{R_{(\theta_j,\nu_{\theta_j})}, 1\right\}$, where,

$$R_{(\theta_j,\nu_{\theta_j})} = \frac{L(\theta_j^{(*)}, \widehat{\Gamma}) \exp\left\{-\frac{\left(\theta_j^{(*)}-\mu_{\theta_j}\right)^2}{2\sigma_{\theta_j}^2}\right\} \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(*)}-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\}}{L(\theta_j^{(t-1)}, \widehat{\Gamma}) \exp\left\{-\frac{\left(\theta_j^{(t-1)}-\mu_{\theta_j}\right)^2}{2\sigma_{\theta_j}^2}\right\} \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\}} \times R_{\nu_{\theta_j}} \times \frac{\Phi\left[\nu_{\theta_j}^{(*)}\left(\frac{\theta_j^{(*)}-\theta_j^{(t-1)}}{\sqrt{\psi_{\theta_j}}}\right)\right]}{\Phi\left[\nu_{\theta_j}^{(t-1)}\left(\frac{\theta_j^{(t-1)}-\theta_j^{(*)}}{\sqrt{\psi_{\theta_j}}}\right)\right]},$$

and $R_{\nu_{\theta_j}}$ as in (16).

(c) To accept every $(\theta_j^{(t)}, \nu_{\theta_j}^{(t)})^t = (\theta_j^{(*)}, \nu_{\theta_j}^{(*)})^t$ with probability $\pi_{(\theta_j,\nu_{\theta_j})}$, otherwise $(\theta_j^{(t)}, \nu_{\theta_j}^{(t)})^t = (\theta_j^{(t-1)}, \nu_{\theta_j}^{(t-1)})^t$.

3.2 **Same asymmetry parameters**

(a) To simulate $\nu_\theta^{(*)}|(\nu_\theta^{(t-1)}, \psi_\nu, \nu_\nu) \sim SN(\nu_\theta^{(t-1)}, \psi_\nu, \nu_\nu)$ and $\theta_j^{(*)}|(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_\theta^{(*)}) \sim SN(\theta_j^{(t-1)}, \psi_{\theta_j}, \nu_\theta^{(*)})$.

(b) To evaluate the vector of accepting probabilities $\nu_\theta^{(t)} = \nu_\theta^{(*)}$ and $\theta_j^{(t)} = \theta_j^{(*)}$, $\nu_\theta = min\{R_{\nu_\theta}, 1\}$ and $\pi_{\theta_j} = min\left\{R_{\theta_j|\nu_\theta}, 1\right\}$, with $R_{\theta_j|\nu_\theta}$ is given by (15) and

$$R_{\nu_\theta} = \frac{L(\theta^{(*)}, \widehat{\Gamma}^{(t-1)}) \exp\left\{-\sum_{j=1}^n \frac{\left(\theta_j^{(*)}-\mu_{\theta_j}\right)^2}{2\sigma_{\theta_j}^2}\right\}\left\{\prod_{j=1}^n \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(*)}-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\}\right\}}{L(\theta^{(t-1)}, \widehat{\Gamma}^{(t-1)}) \exp\left\{-\sum_{j=1}^n \frac{\left(\theta_j^{(t-1)}-\mu_{\theta_j}\right)^2}{2\sigma_{\theta_j}^2}\right\}\left\{\prod_{j=1}^n \Phi\left\{\lambda_{\theta_j}\left(\frac{\theta_j^{(t-1)}-\mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right\}\right\}} \times \frac{\exp\left\{\frac{-\left(\nu_\theta^{(*)}-\mu_\nu\right)^2}{2\sigma_\nu^2}\right\}\Phi\left[\lambda_\nu\left(\frac{\nu_\theta^{(*)}-\mu_\nu}{\sigma_\nu}\right)\right]\left\{\prod_{j=1}^n \Phi\left[\nu_\theta^{(*)}\left(\frac{\theta_j^{(*)}-\theta_j^{(t-1)}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}{\exp\left\{\frac{-\left(\nu_\theta^{(t-1)}-\mu_\nu\right)^2}{2\sigma_\nu^2}\right\}\Phi\left[\lambda_\nu\left(\frac{\nu_\theta^{(t-1)}-\mu_\nu}{\sigma_\nu}\right)\right]\left\{\prod_{j=1}^n \Phi\left[\nu_\theta^{(t-1)}\left(\frac{\theta_j^{(t-1)}-\theta_j^{(*)}}{\sqrt{\psi_{\theta_j}}}\right)\right]\right\}}.$$

(c) To accept $\nu^{(t)} = \nu^{(*)}$ with probability π_{ν_θ} otherwise, $\nu^{(t)} = \nu^{(t-1)}$ and then, to accept every $\theta_j^{(t)} = \theta_j^{(*)}$ with probability π_{θ_j} otherwise, $\theta_j^{(t)} = \theta_j^{(t-1)}$.

In the next section we present a simulation study.

4 Simulation Study

In order to assess the behavior of the estimation procedures we conducted a simulation study. We considered a set of 30 items with parameters such that the discrimination ones (a) range from 1.1 to 1.8 (for the right alternative) and the difficult ones (b) range from -2.5 to 2.5 (for the right alternative). Concerning the other alternatives the parameters ranges from -1.0 to 1 (discrimination parameter) and from -6.0 to 1.5 (difficult parameter). We also considered three types of latent skew-normal distribution with asymmetry parameters equal -2, 0 and 2 and two sample sizes for the subjects equals to n = 500 and n = 1000, this produced 6 situations to analyze. For every of these situations we considered 5 estimation methods which are : Maximum Likelihood (ML), Mode a Posteriori (MAP), Expectation a Posteriori (EAP), Expectation a Posteriori using Monte Carlo Integration (EAPMC) and finally Monte Carlo Markov Chain + Metropolis-Hastings (MCMC + MH). Additionally, for MCMC procedure we evaluated the mean and median using a independent and burn-in samples. That is, a spaced and non-spaced samples. In the MCMCM + MH approach we took, for both independent (considering a lag 30) and burn-in (taking all sample from the burn-in value) samples, the observed mean and median. For all these methods we keep the parameters values as the real item data set. Finally, we used a set of R = 50 replications, for each of former six situations.

In the ML and MAP estimation we set 40 as the maximum iteration number. For MAP, EAP and EAPMC methods we set in 0 and 1 as the priors mean and variance, respectively, and estimated the asymmetry parameter using a sample asymmetry coefficient. For the EAP procedure, we use 30 quadrature points, more points than the literature suggests, in general around 20 (Muraki and Bock, 1997), in order to deal with the asymmetry latent distribution. For the EAPMC method we considered 5000 uniform numbers in order to ensure the accuracy of this method. Finally, concerning the MCMC simulation we adopted the priors and kernel distributions referred in Section 3.4 with $\mu_{\theta_j} = 0$, $\sigma^2_{\theta_j} = 1$, $\lambda_{\theta_j} = 0$, and $\psi_{\theta_j} = 1$, for $j = 1, \ldots, n$. As we considered only the **Empirical Scheme**, the asymmetry parameter of the kernel latent distribution was estimated concurrently with the entire estimation process. After several preview runs we adopted, for the asymmetry hyper parameters, $\sigma^2_{\nu_j} = 0.10$ and $\psi_{\nu_j} = 0.05$, for $j = 1, \ldots, n$, to the -2 and 0 asymmetry situations and $\sigma^2_{\nu_j} = 0.20$ and $\psi_{\nu_j} = 0.15$, for $j = 1, \ldots, n$ to the 2 asymmetry parameter situation. In a real problem, we can decide what situation we are dealing with by verifying the value of the asymmetry coefficient of the observed scores. The mean and asymmetry hyper parameters $(\mu_{\nu_j}, \lambda_{\nu_j}, \nu_{\nu_j})^t$, were set equal to 0. For the burn-in we considered B = 1000. We made inference taking a total of 8000 MCMC simulations with a lag of 30, that is, with 7000 values and taking the simulated values with distance of 30 observations.

In order to compare the estimation methods we considered 4 statistics, that are :

- Variance : the variance of the estimatives obtained from the 50 replicas.
- Mean Squared Residual (MSR) : the mean of the sums of the squares, that is, is the squared difference between the true value and the estimatives of the 50 replicas.
- Bias : the sum of the two former statistics.
- Correlation : the Pearson's correlation coefficient between the true latent trait and their estimatives.

Table 1 presents these four statistics to all five estimation procedures. The results show that bayesian methods produce better estimatives, specially, in the presence of asymmetry. The correlations (Figures 1 to 6) are quite similar but, the other statistics suggest that, mainly the EAP and MAP, produce more precise results. Considering positive or negative asymmetry we can see that the MCMC methods behaved far from the expected. Basically, we suppose that are two reasons for it. First, for the inference purposes, generally, it is better to use some stochastic representation in order to produce stable estimation results. We also noted in our work that it is not convenient to use the original form of skew-normal densities because its produces full conditionals from which is not possible to drawn directly (without using an auxiliary algorithm), sometimes it generates monotonic acceptance rates and finally and if $\lambda_\theta = 0$, we do not have the symmetry situation studied by Patz and Junker (1999). The second reason is, to simulate the asymmetry parameter, to choose a non-informative prior, e.g. a uniform, may be produce better results (Bazán et al, 2004).

Figures 7 to 24 indicate that the EAPMC method presents large bias and variance, mainly for extreme values. In the presence of asymmetry, the MCMC methods also presented high values for bias and variance. The reasons for this behavior are those that was pointed out in the former paragraph. In general, the EAP and MAP produced the smallest values for the statistics, as we can see in these figures, but this depends on the value of latent trait.

5 Final comments

We considered modifications in some estimation procedures concerning the laten trait estimation in the NRM. As we could see, the EAP and MAP methods produced the best results. However, some improvements may be obtained in the MCMC approach through suitable modifications.

The contributions of this paper were three, basically. First, the introduction of asymmetry latent trait distribution in NRM, which is an extension of work of Bock (1972). The proposed model is less sensitive to depart from normality or even symmetry assumptions. Second, the modifications in MAP in EAP, concerning the prior distribution, permits to deal with asymmetry from the data even though, in our work, we did not consider the asymmetry parameter estimation in those methods. Finally, the simulation study showed the effects of the presence of latent asymmetry and the superiority of EAP and MAP methods.

6 Acknowledgments

This work was developed using results from Master's Dissertation written by the first author under the guidance of second. The financial support was given by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) from which the authors are thankful.

A Appendix

A.1 MAP estimation expressions

From (8) and considering the skew-normal density we have,

$$\ln g(\theta_j|\eta) = const - \frac{(\theta_j - \mu_{\theta_j})^2}{2\sigma^2_{\theta_j}} + \ln \Phi\left[\lambda_{\theta_j}\left(\frac{\theta_j - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right].$$

So, the needed terms are

$$\frac{\partial \ln g(\theta_j|\eta)}{\partial \theta_j} = \frac{\theta_j - \mu_\theta}{\sigma^2_\theta} + \left\{\Phi\left[\lambda_{\theta_j}\left(\frac{\theta_j - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\right\}^{-1} \phi\left[\lambda_{\theta_j}\left(\frac{\theta_j - \mu_{\theta_j}}{\sigma_{\theta_j}}\right)\right]\frac{\lambda_{\theta_j}}{\sigma_{\theta_j}},$$

and, using the product derivate rule,

$$\frac{\partial^2 \ln g(\theta_j|\eta)}{\partial \theta_j^2} = \left\{ \frac{1}{\sigma_\theta^2} + h(\lambda_{\theta_j}, \theta_j) \right\},$$

where $h(\lambda_{\theta_j}, \lambda_{\theta_j})$ is given by (9).

A.2 EAP-MC estimation expressions

First, let us call

$$I_1 = \int_{\mathbb{R}} \theta P(Y_{.j}|\theta, \widehat{\Gamma}) g(\theta|\eta) d\theta \, , I_2 = \int_{\mathbb{R}} P(Y_{.j}|\theta, \widehat{\Gamma}) g(\theta|\eta) d\theta \, , \tag{17}$$

$$I_3 = \int_{\mathbb{R}} [\theta - \mathbb{E}\left[\theta|y_{.j.}, \widehat{\Gamma}, \eta\right]]^2 P(Y_{.j}|\theta, \widehat{\Gamma}) g(\theta|\eta) d\theta \tag{18}$$

To solve them, define $\psi = 1 + \frac{2}{e^\theta - 1}$, and notice that,

$$\theta \to -\infty \Rightarrow \Psi \to -1$$
$$\theta \to \infty \Rightarrow \Psi \to 1$$
$$\theta = \ln\left\{ \frac{2}{\psi - 1} + 1 \right\} = \ln \Delta \tag{19}$$

So, the Jacobian is

$$\frac{\partial \theta}{\partial \psi} = \frac{1}{\Delta} \left\{ \frac{-2}{(\psi - 1)^2} \right\} \tag{20}$$

Then, using (19) and (20) in the integrals (17) and (18), and using **Monte Carlo integration** the expressions (13) and (14) follow.

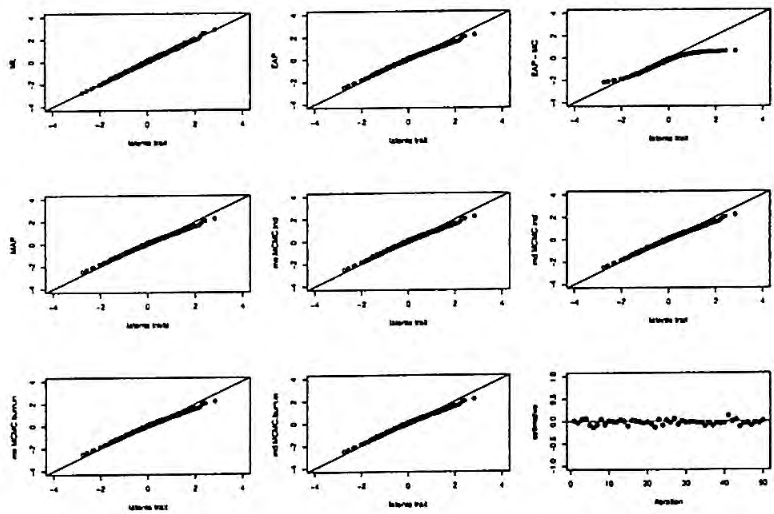

Figure 1: Correlation : true latent trait and estimatives - $\lambda = 0$ and n= 500

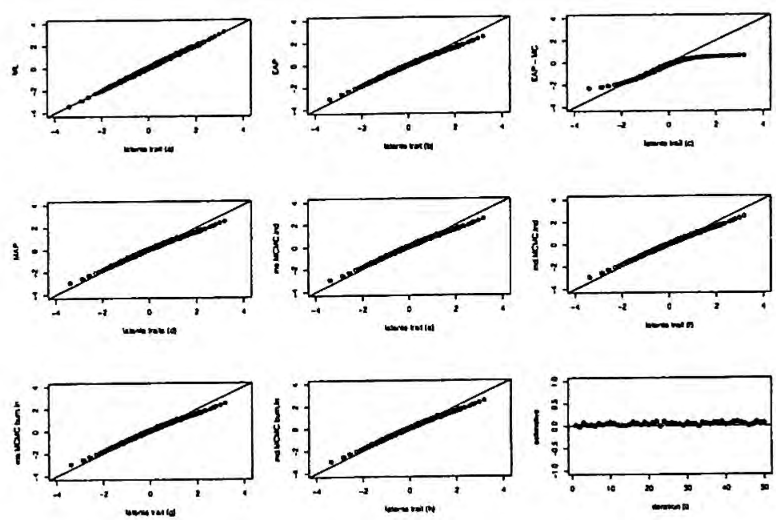

Figure 2: Correlation : true latent trait and estimatives - $\lambda = 0$ and n= 1000

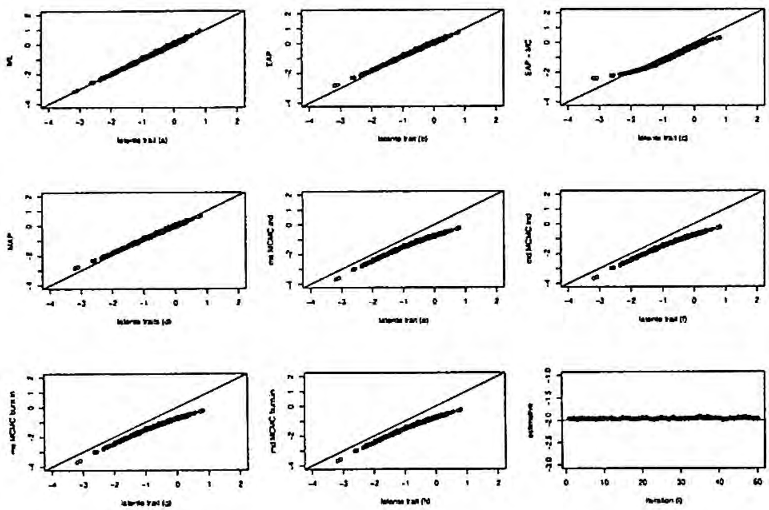

Figure 3: Correlation : true latent trait and estimatives - $\lambda = -2$ and n= 500

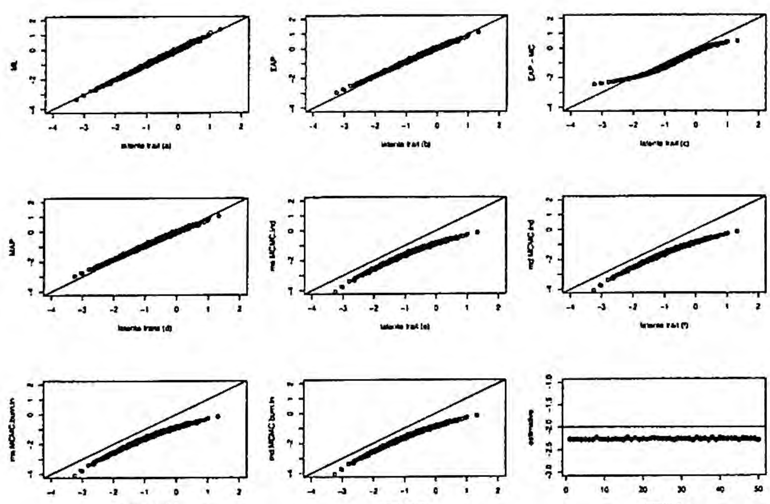

Figure 4: Correlation : true latent trait and estimatives - $\lambda = -2$ and n= 1000

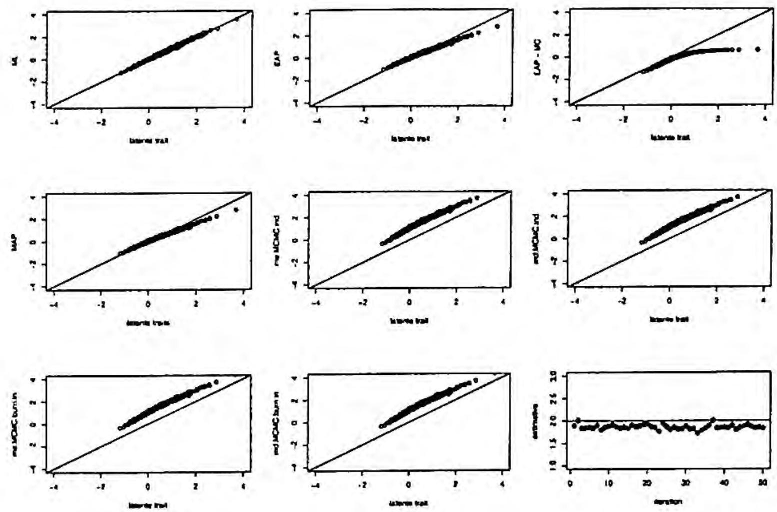

Figure 5: Correlation : true latent trait and estimatives - $\lambda = 2$ and n= 500

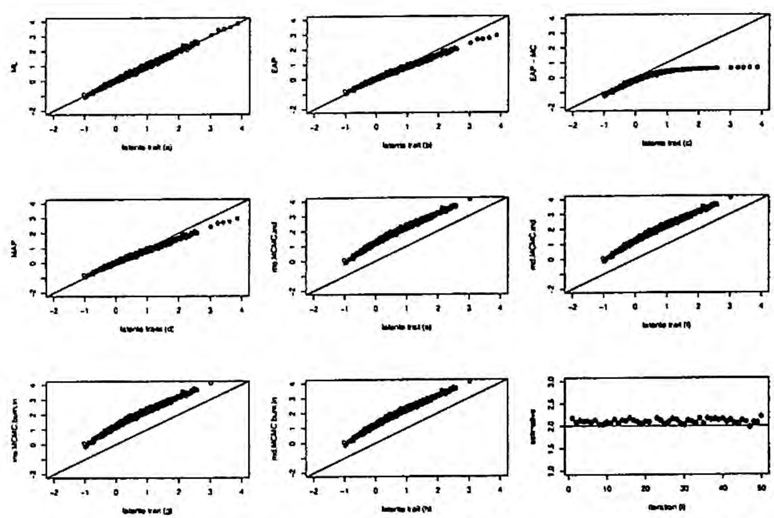

Figure 6: Correlation : true latent trait and estimatives - $\lambda = 2$ and n= 1000

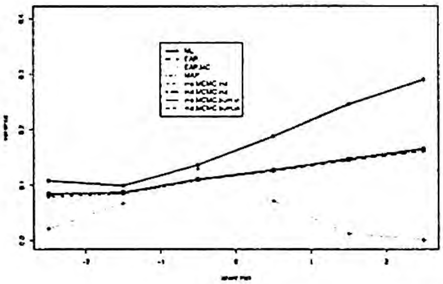

Figure 7: Variance of estimatives : $\lambda = 0$ and n= 500

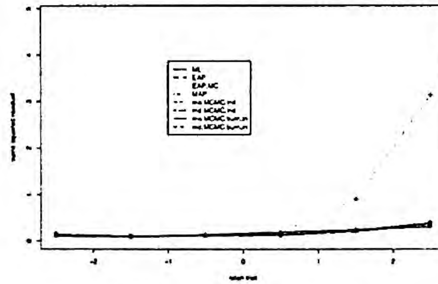

Figure 8: Mean squared residual of estimatives : $\lambda = 0$ and n= 500

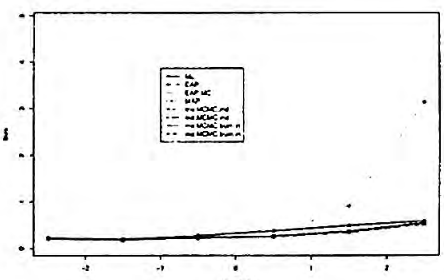

Figure 9: Bias of estimatives : $\lambda = 0$ and n= 500

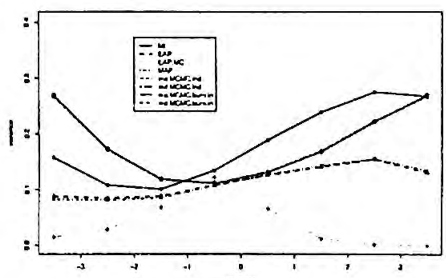

Figure 10: Variance of estimatives : $\lambda = 0$ and n= 1000

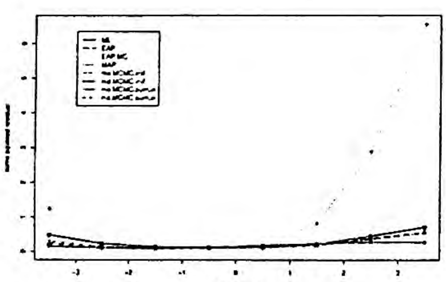

Figure 11: Mean squared residual of estimatives : $\lambda = 0$ and n= 1000

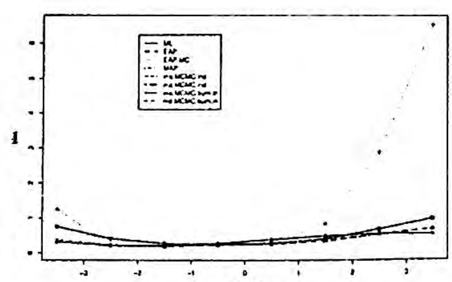

Figure 12: Bias of estimatives : $\lambda = 0$ and n= 1000

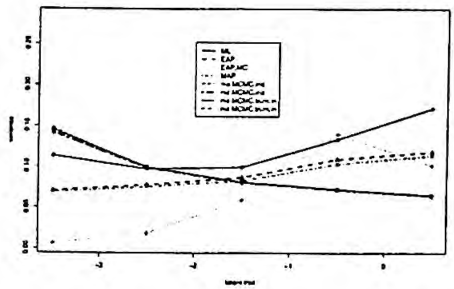

Figure 13: Variance of estimatives : $\lambda = -2$ and n= 500

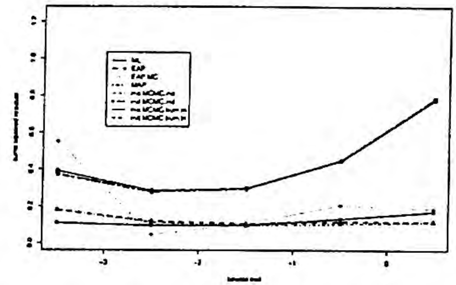

Figure 14: Mean squared residual of estimatives : $\lambda = -2$ and n= 500

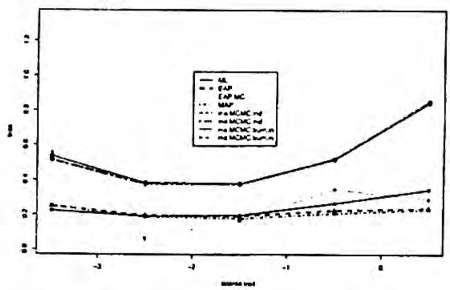

Figure 15: Bias of estimatives : $\lambda = -2$ and n= 500

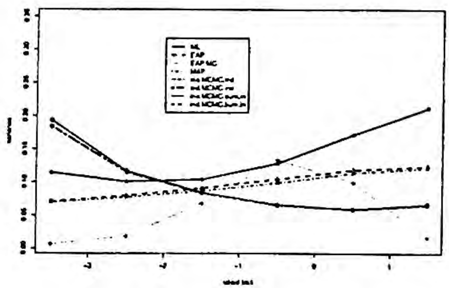

Figure 16: Variance of estimatives : $\lambda = -2$ and n= 1000

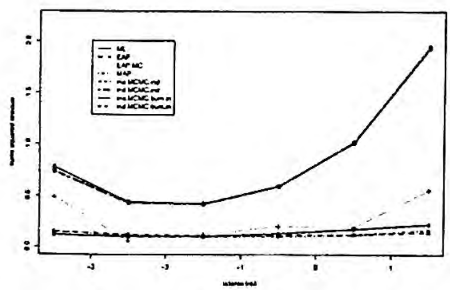

Figure 17: Mean squared residual of estimatives : $\lambda = -2$ and n= 1000

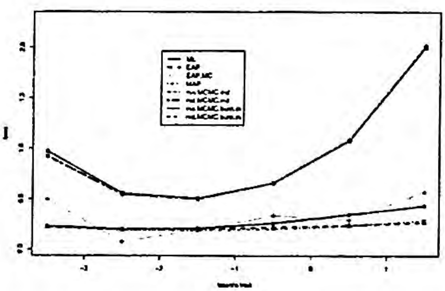

Figure 18: Bias of estimatives : $\lambda = -2$ and n= 1000

Table 1: Statistics of laten trait estimation

λ	n	Estim. Method	Statistics			
			Variance	MSR	Bias	Correlation
-2	500	ML	0.127	0.128	0.255	0.997
		EAP	0.102	0.115	0.218	0.997
		EAP-MC	0.105	0.170	0.276	0.988
		MAP	0.097	0.106	0.203	0.997
		Me.MCMC.ind	0.075	0.445	0.521	0.994
		Md.MCMC.ind	0.075	0.446	0.521	0.994
		Me.MCMC.burn.in	0.075	0.445	0.520	0.994
		Md.MCMC.ind.burn.in	0.075	0.446	0.521	0.994
-2	1000	ML	0.127	0.128	0.255	0.997
		EAP	0.102	0.114	0.216	0.996
		EAP-MC	0.105	0.173	0.277	0.987
		MAP	0.097	0.105	0.201	0.996
		Me.MCMC.ind	0.073	0.602	0.674	0.990
		Md.MCMC.ind	0.072	0.602	0.674	0.990
		Me.MCMC.burn.in	0.073	0.602	0.675	0.990
		Md.MCMC.ind.burn.in	0.072	0.602	0.674	0.990
0	500	ML	0.165	0.166	0.331	0.998
		EAP	0.117	0.135	0.252	0.997
		EAP-MC	0.082	0.288	0.370	0.966
		MAP	0.117	0.136	0.253	0.997
		Me.MCMC.ind	0.118	0.137	0.255	0.997
		Md.MCMC.ind	0.118	0.137	0.255	0.997
		Me.MCMC.burn.in	0.118	0.137	0.255	0.997
		Md.MCMC.ind.burn.in	0.118	0.137	0.255	0.997
0	1000	ML	0.167	0.167	0.334	0.998
		EAP	0.117	0.138	0.255	0.997
		EAP-MC	0.075	0.317	0.391	0.961
		MAP	0.117	0.139	0.255	0.997
		Me.MCMC.ind	0.132	0.156	0.288	0.997
		Md.MCMC.ind	0.132	0.156	0.288	0.997
		Me.MCMC.burn.in	0.131	0.156	0.287	0.997
		Md.MCMC.ind.burn.in	0.131	0.156	0.287	0.997
2	500	ML	0.199	0.200	0.399	0.995
		EAP	0.127	0.157	0.284	0.992
		EAP-MC	0.062	0.474	0.535	0.911
		MAP	0.129	0.160	0.290	0.993
		Me.MCMC.ind	0.203	1.374	1.577	0.990
		Md.MCMC.ind	0.203	1.362	1.565	0.990
		Me.MCMC.burn.in	0.202	1.375	1.577	0.990
		Md.MCMC.ind.burn.in	0.203	1.362	1.565	0.990
2	1000	ML	0.199	0.202	0.400	0.993
		EAP	0.128	0.159	0.286	0.993
		EAP-MC	0.056	0.497	0.553	0.905
		MAP	0.129	0.160	0.290	0.993
		Me.MCMC.ind	0.300	1.880	2.181	0.990
		Md.MCMC.ind	0.299	1.852	2.151	0.990
		Me.MCMC.burn.in	0.300	1.879	2.179	0.990
		Md.MCMC.ind.burn.in	0.298	1.851	2.149	0.990

References

[1] Azevedo, C. L. N. (2003). *Estimation Methods in Item Response Theory.* Unpublisehd Master's Dissertation (In Portuguese). Institute of Mathematics and Statistics. University of Sao Paulo, Brazil (www.teses.usp.br).

[2] Baker, F. B. and Kim, Seock-Ho (2004). *Item Response Theory : Parameter Estimation Techniques.* Statistics, Dekker Series of Textbooks and Monographs, New York, NY.

[3] Bazán, J. L., Bolfarine, H. and Leandro, R. A. (2004). *Bayesian Estimation via MCMC for Probit-Normal model in Item Response Theory.* Techincal Report, IME-USP.

[4] Bernardo, J. M. and Smith, A. F. (2001). *Bayesian Theory.* John Wiley Sons.

[5] Bock, R. D. (1972). Estimating item parameteres and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29 - 51.

[6] Bock, R. D. (1997). The Nominal Categories Model. *In Handbook of Modern Item Response Theory.* Wim J. van der Linden and Ronald K. Hambleton eds. Springer-Verlarg, New York.

[7] Fraser, D. A. S. and Reid, N. (1989). Adjustment to profile likelihood. *Biometrika*, 76, 477-488.

[8] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Transaction on Pattern Analysis and Machine Intelligence*, 6, 721-741.

[9] Genton, M. C. (2004). *Skew-elliptical Distibutions and Their Applications: A Journey Beyond Normality.* Chapman & Hall/CRC.

[10] Kim, Seock-Ho, Cohen, A. S., Baker, F. B., Subkoviak, M. J. and Leonard, T. (1994). An investigation of hierarchical Bayes procedures in item response theory. *Psychometrika*, 59, 405 - 421.

[11] Mislevy, R. J. (1984). Estimating Latent Distribution. *Psychometrika*, 49, 359 - 381.

[12] Mislevy, R. J. and Stocking, M. L. (1989). A Consumer's Guide to LOGISTIC and BILOG. *Applied Psycholigical Measurement*, 13, 57-75.

[13] Muraki, E. and Bock, R. D.(1997). *PARSCALE: IRT Based Test Scoring and Item Analysis for Graded Open-Ended Exercises and Perfomance Tasks. Chicago: Scientific Siftware, Inc.*

[14] Patz, J. R. and Junker, B. W. (1999). A Straightforward Approach to Markov Chain Monte Carlo Methods for Item Response Models. *Journal of Educational and Behavorial Statistics*, 24, 146 - 178.

[15] Robert, C. P. and Casella, G. (2000). *Monte Carlo Statistical Methods.* Springer-Verlag New York, Inc.

[16] Stroud, A. H. and Secrest, D. (1966). *Gaussian Quadrature Formulas.* Englewood Cliffs, New Jersey : Prentice Hall.

[17] Van Der Vaart, A. W. and Murphy, S. A. (2000). On Profile likelihood. *Journal of the American Statistical Association*, 95, 2000.

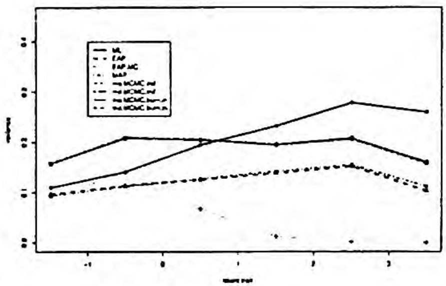

Figure 19: Variance of estimatives : $\lambda = 2$ and n= 500

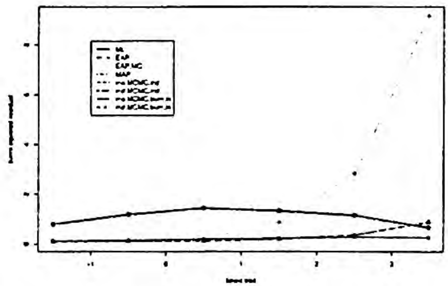

Figure 20: Mean squared residual of estimatives : $\lambda = 2$ and n= 500

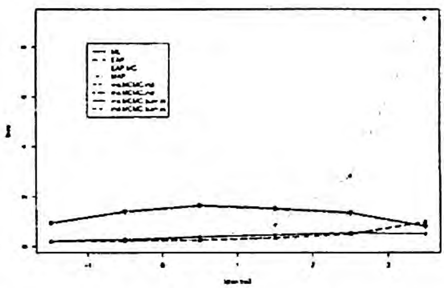

Figure 21: Bias of estimatives : $\lambda = 2$ and n= 500

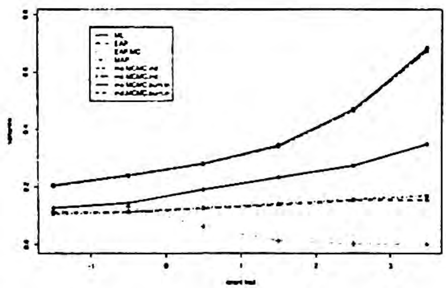

Figure 22: Variance of estimatives : $\lambda = 2$ and n= 1000

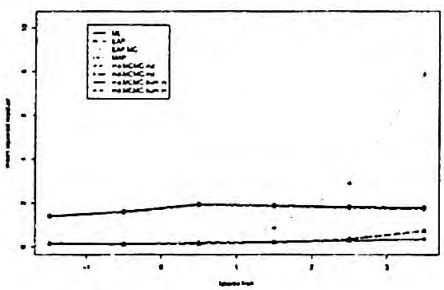

Figure 23: Mean squared residual of estimatives : $\lambda = 2$ and n= 1000

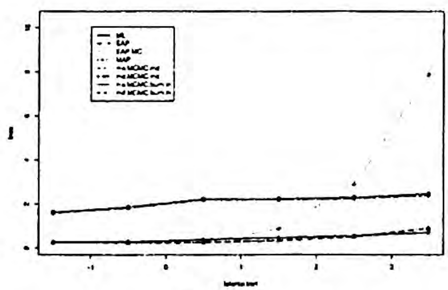

Figure 24: Bias of estimatives : $\lambda = 2$ and n= 1000

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

2007-01 - POLPO, A., PEREIRA, C.A.B. Reliability nonparametric bayesian estimation in parallel systems 2007. 12p. (RT-MAE 2007-01)

2007-02 - GONÇALVES, M., Kolev, N. V., Fabris. A. E. Bounds for quantile-based measures of dependent risks' functions. 2007. 16p. (RT-MAE 2007-02)

2007-03 - MIMBELA, J. A. L, Kolev, N.V. Occupation measure of Markov - modulated risk processes 2007. 12p. (RT-MAE 2007-03)

2007-04 - GONÇALVES, M., Kolev, N. V., Fabris, A. E. Bounds for distorted risk measuress. 2007. 11p. (RT-MAE 2007-04)

2007-05 - BUENO, V.C., A series representation of a parallei systems. 2007. 10p. (RT-MAE 2007-05)

2007-06 - KOLEV, N.V., PAIVA, D. Copula-Based Regression Models. 2007. 27p. (RT-MAE 2007-06)

2007-07 - POLETO, F. Z., SINGER, J.M., PAULINO, C.D., A product-multinomial framework for categorical data analysis with missing responses. 2007. 27p (RT-MAE 2007-07)

2007-08 - KOLEV, N.V., GONCALVES, M. DIMITROV, B. Probabilistic Properties of Sibuya's Dependence Function. 2007. 13p. (RT-MAE-2007-08)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05311-970 - São Paulo, Brasi