A model of skew item response theory

Jorge L. Bazán, Márcia D. Branco, Heleno Bolfarine
Departament os Statistics
University of São Paulo

Dezember 2003

Summary

In this paper, we introduce a skew item response theory (IRT) for modelling the probability of correctly answearing an item by n subjects in a test with I itens. We extend the ogive normal model (see Albert and Ghosh, 2000) by considering the skew normal distribution function as a link function.
Additionally, considering a Bayesian approach, skew normal distributions are considered as priors for the latent variables to model subject's ability related to human condition which can be far from normality as sujested in Samejima (1997). Model properties are investigated and a Bayesian estimation approach is developed by using the MCMC methodology for simulating from the posterior distribution of item parameters and latent variables. An application with a data set from mathematics test is reported.
Keywords: link skew-probit, item response theory, bayesian estimation, Slice sampling, normal ogive model.

Abstract

The Item response theory is used to model a set of multivariate dichotomous response from n individuals that are submitted to test with I items. This model consider latent variables associate to each individual, that explains to subjects ability and a set of parameters associate to the item under consideration. The IRT model the probability of the sucess p_{ij} as $p_{ij} = F_{ij}(m_{ij}) = \Phi(m_{ij})$ $i = 1, \ldots, n$, $j = 1, \ldots, p$, where F_{ij} is called item curve characteristic and F_{ij}^{-1} is a link function, $\Phi(.)$ is the distribution function of the $N(0,1)$ distribution and $m_{ij} = a_j U_i - b_j$, where a_j and b_j are parameter associate to the items and U_i is the latent variables associate to the individual. To complete the model, the variables U_i are assumed standard normal.

In this paper a new skew link function is considered, the skew normal distribution given by Azzalini (1985). The skew normal distribution is an important class of asymmetric distribution that has as special case the normal distribution.

Samejima (2000) also proposed a new skew link function by considering item characteristic curve that generalize a logit model. Chen, Dey and Shao (1999) considered the a particular skew normal link function in the context of generalized linear models. A random variable has a standard skew normal distribution if its probability density function (pdf) is given by

$$f_Z(z) = 2\phi_1(z)\Phi_1(\lambda z), \tag{1}$$

We consider the notation $Z \sim SN(\lambda)$, where λ is the skewness parameter. Note that, if $\lambda = 0$ the normal pdf is recovered. We can show the cdf of the skew normal is given by

$$\Phi_{SN}(z;\lambda) = \int_{-\infty}^{z} 2\phi_1(u)\Phi_1(\lambda u)du = 2\Phi_2\Big(\Big(\begin{array}{c} z \\ 0 \end{array} \Big); \Big(\begin{array}{c} 0 \\ 0 \end{array} \Big), \Big(\begin{array}{cc} 1 & -\delta \\ -\delta & 1 \end{array} \Big)\Big), \tag{2}$$

when $|\delta| < 1$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ and $\Phi_2(.)$ to the distribution function the bivariate standard normal distribution with correlation $-\delta$, which we denoted for easiness as $\Phi_2((z,0)^T, -\delta)$.

This is result indicates that the distribution skew-normal availed in z point also is obtain by considering a bivariate standard normal distribution with correlation $-\delta$ availed in $(z,0)$ point. This is result is important by considering that several computational algorithms is available for the calculation of the integral in the bivariate normal distribution.

The item characteristic curve (ICC) for the conditional probability of a correct response in item j given the latent variable U_i the subject i, in the SN-IRT model is given by

$$p_{ij}(u) \equiv Prob[Y_i = 1 \mid u] = \Phi[m_{ij};\lambda_j] = \int_{-\infty}^{m_{ij}} \phi(z;\lambda_j)dz, \tag{3}$$

In the above expression, this probability is expressed as a function of the quantity U_i and the parameters $\eta_j = (\alpha_j, \beta_j)^T$ and λ_j, which are parameters associated with item j. We used

$$p_{ij}(u) = 2\Phi_2\Big(\Big(\begin{array}{c} m_{ij} \\ 0 \end{array} \Big); \Big(\begin{array}{c} 0 \\ 0 \end{array} \Big), \Big(\begin{array}{cc} 1 & -\delta_j \\ -\delta_j & 1 \end{array} \Big)\Big) = 2\Phi_2\big[(m_{ij},0)'; -\delta_j\big] \tag{4}$$

where $|\delta_j| < 1$ and $\delta_j = \frac{\lambda_j}{\sqrt{1+\lambda_j^2}}$. $i = 1, \ldots, n$ and $j = 1, \ldots, I$.
Note that for $\lambda_j = 0$, (4) reduces to $p_{ij} = 2\Phi(\alpha_j(u - \beta_j))\Phi(0) = \Phi(\alpha_j(u - \beta_j))$, and the symmetric probit-normal or normal ogive model follows.

In this paper we discuss the importance of the new item parameter λ and its relationship with others parameters.
Let $D_{obs} = \mathbf{y}$ the observed data, the likelihood function for the skew-normal IRT by considering is given by

$$L = \prod_{i=i}^{i=n}\prod_{j=1}^{j=I}\Big[2\Phi_2\big[(m_{ij},0)';-\delta\big]\Big]^{y_{ij}}\Big[1-2\Phi_2\big[(m_{ij},0)';-\delta\big]\Big]^{1-y_{ij}} \tag{5}$$

it is equivalent to ogive normal model likelihood.
The Maxima likelihood estimador can be obtain but but some problems are presenting in the literature (see Patz e Junker, 1999).
Considering the likelihood function given in (5), it is possible to implement a Bayesian estimation procedure by incorporating prior distributions for $\boldsymbol{u}$, $\boldsymbol{\eta}$ and $\boldsymbol{\lambda}$. Chen et al. (1999) discuss conditions for the existence of proper posterior distributions for dichotomous models with asymmetric link functions by considering improper prior distributions. For the SN-IRT model, we consider the following general class of prior distribution:

$$\pi(\boldsymbol{u},\boldsymbol{\eta},\boldsymbol{\lambda}) = \prod_{i=1}^{n} g_{1i}(u_i)\prod_{j=1}^{I} g_{2j}(\eta_j)g_{3j}(\lambda_j) \tag{6}$$

where $g_{2j}(\eta_j) = g_{21j}(a_j)g_{22j}(b_j)$ in which g_{21j} and g_{22j} should be proper to guarantee a proper posterior distribution (see Albert & Gosh, 1999 & Ghosh et al. 2001). Following proposals usually considered Rupp et al. (2004), we take $g_{21} \equiv \phi(\mu_a, s_a^2)$ and $g_{22} \equiv \phi(0, s_b^2)$ so that $g_2 \equiv \phi_2(\mu_\eta, \Sigma_\eta)$ with $\mu_\eta = (\mu_a, 0)'$ and $\Sigma_\eta = \begin{pmatrix} s_a^2 & 1 \\ 1 & s_b^2 \end{pmatrix}$.
Additionally, we consider $g_3 = \phi_{SN}(\lambda;\omega)$ with $\mu_a, s_a^2, s_b^2, , \omega$ known values. In more general situations where w is also unknown, the prior structure needs to be enlarged so that prior information are also considered for those parameters. We consider the SN-IRT model in terms of $\delta_j = \lambda_j/(1+\lambda_j^2)^{1/2}$, $j = 1, \ldots, I$, which takes values in the interval $(-1, 1)$, so that we can consider a uniform priori in $(-1, 1)$. To consider an uniform for δ_j is equivalent to consider a t-sduten for λ_j with location $\mu = 0$, scale $\sigma^2 = 1/2$ and $v = 2$ degree of freedom.

In order to implement the MCMC procediment, we introduce two auxiliary latent variables as given in Albert (1992) and Chen et al. (1999). Then, the skew-normal IRT model for $j = 1, \ldots, I$ items and $i = 1, \ldots, n$ subjects by considering $y_{ij} \sim Ber(p_{ij})$ with probability of the success $p_{ij} = \Phi(m_{ij}; \lambda_j)$ where $m_{ij} = a_j u_i - b_j$, it is equivalent to:

$$y_{ij} = \begin{cases} 1, & Z_{ij}^* > 0; \\ 0, & Z_{ij}^* \leq 0. \end{cases},$$

with $Z_{ij}^* \sim N(-\delta_j X_{ij} + m_{ij}, \sqrt{1-\delta^2})$, $Z^* = Z_{ij}|X_{i,j},$, $X_{ij} \sim HN(0,1)$ e $Z \sim SN(m_{i,j}, 1, -\lambda_j)$. Considering this result, the full conditionals have closed form.

An application with a data set from mathematics test is reported. In this application, 14 item of the Mathematical Test available for download in `http://www.minedu.gob.pe/umc/` were applied to 131 students of high socio-economical status. Item response vectors are available from authors upon request.
Bayesian estimation procedures based in MCMC was implemented in WINBUGS by using Slice Sampling method (Neal, 2003). The new item parameter proposed is interpreted in the SN-IRT model.
Discussion empathized for other models skew to consider in future development by include asymmetry in the latent variable. New asymmetric link are discussed.

References

[1] Albert, J. H. (1992). Bayesian Estimation of Normal Ogive Item Response Curves Using Gibbs Sampling. *Journal of Educational Statistics, 17, 251 - 269.*

[2] Albert, J.H. and Ghosh, M. (2000). Item response modeling. *Generalized Linear Models: A Bayesian Perspective*(D. Dey, S. Ghosh and Mallick, eds.), Marcel-Dekker, New York, 173-193.

[3] Arnold, B.C., Beaver, R. J., Groeneveld, R. A., Meeker, W. Q. (1993). The nontrucated marginal of a truncated bivariate normal distribution. *Psychometrika, 58, 471-478.*

[4] Azzalini. A (1985). A class of distributions which includes the normal ones. *Scand. J. Statistical*, *12*, 171-178.

[5] Chen, M-H,,Dey, D.K., Shao, Q-M. (1999). A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, *vol. 94 No.448*,1172-1186.

[6] Patz, R. J., & Junker, B. W. (1999). A straighforward approach to Markov Chain Monte Carlo methods for item response models. *Journal of Educactional and Behavioral Statistics*,24,146-178.

[7] Sahu, S. K.,Dey, D.E., Branco, M. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadia Journal of Statistics*, Vol. 29,217-232.

[8] Samejima, F. (2000). Logistic positive exponent family of models: virtue of asymmetric item characteristics curves. *Psychometrika*, Vol. 65, No. 3, 319-335.

[9] Neal, Radford. (2003). Slice Sampling. *The Annals of Statistics*, Vol. 31, No. 3, 705-767.

[10] van der Linden, W.J. e Hambleton, R. K. (1997). *Handbook of Modern Item response Theory*. New York: Springer-Verlag.