Listening. Learning. Leading.®

Research Report

ETS RR–11-40

The Sensitivity of Parameter Estimates to the Latent Ability Distribution

Xueli Xu

Yue Jia

The Sensitivity of Parameter Estimates to the Latent Ability Distribution

Xueli Xu and Yue Jia
ETS, Princeton, New Jersey

October 2011

Technical Review Editor: Dan Eignor

Technical Reviewers: Frank Rijmen and Matthias von Davier

Abstract

Estimation of item response model parameters and ability distribution parameters has been, and will remain, an important topic in the educational testing field. Much research has been dedicated to addressing this task. Some studies have focused on item parameter estimation when the latent ability was assumed to follow a normal distribution, whereas others have utilized nonparametric or semiparametric techniques to substitute the normal ability assumption. However, both approaches have their limitations. A normal ability assumption is not flexible enough to reflect possible deviations from symmetry, whereas the nonparametric and semiparametric techniques used to capture possible nonnormal features of the latent ability have difficulty in reaching satisfactory estimates for certain quantities of the ability distribution such as quantiles. Hence a continuous generalized skew normal (GSN) distribution was applied in this study to better capture the possible underlying asymmetric ability distribution. In addition, simultaneous estimation of both the item parameters and the distributional parameters was employed. The performance of the GSN was compared with the normal ability assumption in terms of item parameter and distributional parameter recoveries, based on a series of simulation studies. The results showed that (a) under the Rasch model, both the item parameter estimates and the distributional parameter estimates are robust to the misspecification of the ability distribution, and (b) under the two-parameter logistic model, although the distributional parameter estimates are fairly robust, the item parameter estimates are slightly more sensitive to the misspecification of the ability distribution, especially when the underlying ability distribution is highly skewed.

Key words: generalized skew normal distribution, skewed normal distribution, quantiles, EM algorithm, RAL algorithm

Acknowledgments

The authors would like to thank Frank Rijmen, Andreas Oranje, and Dan Eignor for helpful suggestions and comments.

Introduction

The latent ability distribution is a fundamental concept in item response theory (IRT) and plays an important role in the field of educational testing. For example, in the National Assessment of Educational Progress (NAEP) reporting, the basic characteristics of the latent ability distribution, such as mean, standard deviation, and quantiles, are utilized to monitor changes in students' achievements over time.

Usually, a normal distribution is assumed for the latent ability for several reasons. One reason is that for a normal distribution, the first two moments (mean and standard deviation) are sufficient to describe the entire distribution. Although it is true that we can assume any distributional form for the latent ability because of its latency, a normal distribution is the simplest and most convenient one to use. However, one wonders what will happen when the population is composed of seemingly different subgroups. An available example is the NAEP assessment. The representative student sample that takes the NAEP assessment includes students from different races and ethnicities and socioeconomic statuses (SES). Historically, race–ethnicity and SES have often been regarded as important factors affecting academic performance. Thus the so-called NAEP population might follow a distribution that deviates from a normal distribution. Consequently, this situation calls for other forms of distributions for the latent ability.

Calibration of item response model item parameters and the ability distribution has been, and will remain, an active research area in the educational testing field. It is well known that the parameters in traditional IRT models, such as the one-, two-, or three-parameter logistic model, are not identifiable without certain constraints. Consequently, much research and practice in calibration focus on fixing the latent ability distribution. For example, a standard normal distribution is usually assumed for the latent ability random variable. Such practice can be found in computer programs such as BILOG (Mislevy & Bock, 1991) and PARSCALE (Muraki & Bock, 2003). Conversely, there is research in which simultaneous estimation of both the item parameters and the distributional parameters is emphasized, in which the constraints are imposed on the item parameters. For instance, Sanathanan and Blumenthal (1978) used a form of the expectation–maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977) to obtain estimates for the Rasch model (Rasch, 1960), by which the latent ability is treated as a random sample from a normal distribution with unknown mean and variance. Rigdon and Tsutakawa (1983) proposed using a computationally simpler modification of the EM algorithm to estimate

the parameters of the Rasch model, by which the latent ability was also treated as a normal distribution with unknown mean and variance. For the two-parameter probit model, Bock and Aitkin (1981) proposed approximating the prior distribution of the latent ability by a discrete distribution over a finite number of ability levels and using the EM algorithm to estimate the item parameters. In addition, some nonparametric or semiparametric methods were applied to estimate the item parameters when the latent ability distribution was allowed to be nonnormally distributed (Bouezmarni, Rijman, & De Boeck, 2008; Heinen, 1996; Laird, 1978; Lindsay, Clogg, & Grego, 1991; Molenberghs & Verbeke, 2005; Woods & Thissen, 2004). Although these methods have been successfully implemented in many software packages, some distributional parameters, such as quantiles, may be difficult to estimate using discrete levels of the ability distribution because of the noncontinuity of the approximation (Aitkin & Aitkin, 2004).

In this report, we discuss the parameter estimates for the Rasch model and for the two-parameter logistic (2PL; Lord & Novick, 1968) model, in which the extended form of the EM algorithm of Rigdon and Tsutakawa (1983) is applied. In the models illustrated in this report, the latent ability distribution is treated as either a normal distribution with unknown mean and variance or a generalized skew normal distribution with unknown location, scale, and skewness. The primary goals of this paper are to (a) discuss the applicability of the EM algorithm in the case where a distribution other than the normal distribution was assumed for the latent ability, (b) check the stability of item parameter estimates across different ability distribution assumptions, and (c) check the stability of ability inference, such as the mean, standard deviation, and quantiles, across different ability distribution assumptions.

The report is organized as follows. Section 2 introduces the generalized skew normal distribution. This is followed by the algorithm to estimate the parameters of the mixed Rasch model and mixed 2PL model in section 3. A simulation study is conducted in section 4, and results are summarized in section 5. Section 6 provides a brief conclusion and discussion.

Generalized Skew Normal Distribution

The generalized skew normal (GSN) distribution, a special form of the generalized skew elliptical (GSE) distribution (Ma, Genton, & Tsiatis, 2005), will be used in fitting the ability distribution in the Rasch and 2PL models. The density of a random variable with a GSN

distribution is defined through a normal distribution and a skewing function, as follows:

$$f(\theta) = \frac{2}{\sigma}\phi\Big(\frac{\theta-\xi}{\sigma}\Big)\pi\Big(\frac{\theta-\xi}{\sigma}\Big), \quad \theta \in R, \tag{1}$$

where $\phi(\cdot)$ is the probability density function of a standard normal distribution, the function $\pi : R \to [0,1]$ satisfies $\pi(\theta) + \pi(-\theta) = 1$, and π is a continuous function. We refer to π as the *skewing function.* This distribution can be simplified to the skewed normal (SN) distribution (Azzalini & Dalla Valle, 1996) if the skewing function π takes the form of the cumulative normal distribution:

$$f(\theta) = \frac{2}{\sigma}\phi\Big(\frac{\theta-\xi}{\sigma}\Big)\Phi\Big(\frac{\theta-\xi}{\sigma}\Big).$$

It is noticed that the function π in (1) is a semiparametric function because its form is unknown. In estimation, the function π is approximated by functions that satisfy $\pi(\theta)+\pi(-\theta) = 1$. In fact, the cumulative distribution function (CDF) of any symmetric distribution could be a candidate for this function. Certainly we would select a CDF that is computationally easy to handle. In this report, we approximate the skewing function by

$$\pi_K(\theta) = H\left[P_K\left(\frac{\theta-\xi}{\sigma}\right)\right], \tag{2}$$

where H is the logit link function (i.e., $H(\theta) = 1/[1 + \exp(-\theta)]$) and P_K is an odd polynomial function of order K. One may wonder why the logit link function is chosen instead of the CDF of a normal distribution. This choice is mainly due to the fact that the normal CDF does not have a mathematical form, whereas the logit link function does. A polynomial of order 3 ($K = 3$) was used in this study because research (Genton, 2004) has shown that $K = 3$ is sufficient to capture the features of the most commonly used distributions up to three modes.

Estimation of Mixed Models

The extended EM algorithm used by Rigdon and Tsutakawa (1983) was applied in this study to simultaneously estimate item parameters and distributional ability parameters in the mixed Rasch and 2PL models. Consider J items with item parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J)$ and a random sample of N subjects with real-value ability parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$ selected from a prior distribution indexed by a parameter vector $\boldsymbol{\gamma}$. Given $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, the joint distribution of response vector $\boldsymbol{y}$ and $\boldsymbol{\theta}$ is

$$f(\boldsymbol{y}, \boldsymbol{\theta}|\boldsymbol{\beta}, \gamma) = p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\beta})p(\boldsymbol{\theta}|\boldsymbol{\gamma}), \tag{3}$$

where $p(\boldsymbol{\theta}|\boldsymbol{\gamma})$ is the prior probability density function (pdf) of $\boldsymbol{\theta}$ with unknown parameters $\boldsymbol{\gamma}$. Moreover, the posterior distribution of $\boldsymbol{\theta}$, given $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, is

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}) &\propto p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\beta})p(\boldsymbol{\theta}|\boldsymbol{\gamma}) && (4) \\ &\propto \prod_i p(\theta_i|\gamma) \prod_j p(y_{ij}|\theta_i, \beta_j). && (5) \end{aligned}$$

We consider a set of estimates of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ that maximizes the marginal likelihood function

$$l(\boldsymbol{\beta}, \boldsymbol{\gamma}|\boldsymbol{y}) = \int f(\boldsymbol{y}, \boldsymbol{\theta}|\boldsymbol{\beta}, \boldsymbol{\gamma}) d\theta. \tag{6}$$

These estimates can be obtained from an extended EM algorithm. The extended EM algorithm starts with some provisional estimates $(\boldsymbol{\beta}^0, \boldsymbol{\gamma}^0)$ of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ and finds the value of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ iteratively, which maximizes

$$E[\log f(\boldsymbol{y}, \boldsymbol{\theta}|\boldsymbol{\beta}, \boldsymbol{\gamma})|\boldsymbol{y}, \boldsymbol{\beta}^0, \boldsymbol{\gamma}^0]. \tag{7}$$

Equation (7) can be written as

$$\sum_i \int \log p(\theta_i|\boldsymbol{\gamma}) p(\theta_i|\boldsymbol{y}, \boldsymbol{\beta}^0, \boldsymbol{\gamma}^0) d\theta_i + \sum_j \sum_i \int \log p(y_{ij}|\theta_i, \boldsymbol{\beta}_j) p(\theta_i|\boldsymbol{y}, \boldsymbol{\beta}^0, \boldsymbol{\gamma}^0) d\theta_i, \tag{8}$$

where i is the index for subjects and j is the index for items. One can note that the estimation of $\boldsymbol{\gamma}$ involves only the first component of (7), and the maximization with respect to $\boldsymbol{\beta}$ is concerned with the second series only. One may estimate $\boldsymbol{\beta}$ in component-wise manner using only a single series over i in the double series.

As usual, the parameter estimates for $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ can be obtained by taking the first derivative of the log-likelihood in (7). The $\boldsymbol{\beta}$ consists of item difficulty parameters in the Rasch model or item difficulty and discrimination parameters in the 2PL model. If the latent ability is treated as a normal distribution, the $\boldsymbol{\gamma}$ includes the mean and variance of a normal distribution. Conversely, it includes the location, scale, and skewness parameters if a GSN distribution is instead assumed for the latent ability distribution.

When the latent ability distribution is treated as normally distributed, the estimates of the mean (denoted as μ) and variance (indexed by σ^2) of the normal distribution are updated by the following:

$$\begin{aligned} \mu^{(t+1)} &= \frac{1}{N} \sum_i \int \theta_i p(\theta_i|\boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{2(t)}) d\theta_i && (9) \\ \sigma^{2(t+1)} &= \frac{1}{N} \sum_i \int (\theta_i - \mu^{(t+1)})^2 p(\theta_i|\boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{2(t)}) d\theta_i, && (10) \end{aligned}$$

where $(\boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{2(t)})$ are the estimates from the previous iteration.

When the ability distribution is assumed to follow a GSN distribution in estimation, the estimates of the location (denoted as μ) and scale (indexed by σ) are updated by an extended form of the regular asymptotically linear (RAL) estimator proposed by Ma et al. (2005). In particular, they are the values of (μ, σ) that solve the following two nonlinear equations:

$$\sum_i \int \left[\frac{\theta_i - \mu}{\sigma}\left(2\pi\left(\frac{\theta_i - \mu}{\sigma}, \hat{\alpha}\right) - 1\right) - 2\pi_1\left(\frac{\theta_i - \mu}{\sigma}, \hat{\alpha}\right)\right] p(\theta_i | \boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{(t)}) d\theta_i = 0 \tag{11}$$

$$\sum_i \int [(\theta_i - \mu)^2 - \sigma^2] p(\theta_i | \boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{(t)}) d\theta_i = 0, \tag{12}$$

where $\hat{\alpha}$ is the estimate of the parameters used in the approximation of the skewing function $\pi(\cdot)$. For the derivation of the preceding nonlinear equations, the reader is referred to Xu (2007).

The estimates of the item parameter b_j in the Rasch model,

$$p(y_{ij} | \theta_i) = \frac{\exp(y_{ij}(\theta_i - b_j))}{1 + \exp(\theta_i - b_j)},$$

can be updated by solving the equation

$$\sum_i y_{ij} = \sum_i \int (1 + \exp(b_j - \theta_i))^{-1} p(\theta_i | \boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{(t)}) d\theta_i. \tag{13}$$

The estimates of the item parameters a_j and b_j in the 2PL model,

$$p(y_{ij} | \theta) = \frac{\exp(y_{ij}(a_j\theta_i - b_j))}{1 + \exp(a_j\theta_i - b_j)},$$

can be updated by solving the following normal equations:

$$\sum_i \int \left[y_{ij} - \frac{1}{1 + \exp(b_j - a_j\theta_i)}\right] \theta_i p(\theta_i | \boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{(t)}) d\theta_i = 0 \tag{14}$$

$$\sum_i \int \left[y_{ij} - \frac{1}{1 + \exp(b_j - a_j\theta_i)}\right] p(\theta_i | \boldsymbol{y}_i; \boldsymbol{\beta}^{(t)}, \mu^{(t)}, \sigma^{(t)}) d\theta_i = 0. \tag{15}$$

Owing to the nonidentifiability of the parameters in the Rasch and 2PL models, certain constraints have to be imposed to obtain unique estimates. For the Rasch model, the constraint $\sum_j b_j = 0$ is imposed, whereas the constraints $\sum_j b_j = 0$ and $\prod_j a_j = 1$ are implemented for the 2PL model. The maximization with respect to the item parameters (see [13]–[15]) is still performed component-wise for each a_j and b_j, followed by implementation of the constraints.

Simulation Study

In this section, we discuss the estimates of both item parameters and the distributional ability parameters under different latent ability assumptions via simulation study. These data were generated from the Rasch model and from the 2PL model with sets of item parameters and different ability distributions (i.e., normal distribution and SN distribution). A complete factorial design was used in this study, in which two factors were considered. One was the ability-generating distribution, and the other was the ability-fitting distribution. The data sets generated from a normal distribution were calibrated by employing both normal and GSN distributions. Likewise, the data sets generated from a SN distribution were estimated by both normal and GSN distributions. Sixty replications were conducted for each of the four scenarios. The results are the summary of these 60 replications for each case.

The item difficulty parameters of both the Rasch and 2PL models were generated from a standard normal distribution, and the item discrimination parameters in the 2PL model were generated from a uniform distribution $U(0.5, 2.5)$. The sets of $J = 30$ items were applied to groups of $N = 1,000$ hypothetical subjects whose ability values were chosen randomly from either a normal distribution with fixed mean and variance or a SN distribution with fixed location, scale, and skewness. The ability-generating distributions include $N(-0.5, 1)$, $SN(-0.5, 1, -1)$, and $SN(-0.5, 1, 3)$. These three distributions are shown in Figure 1.

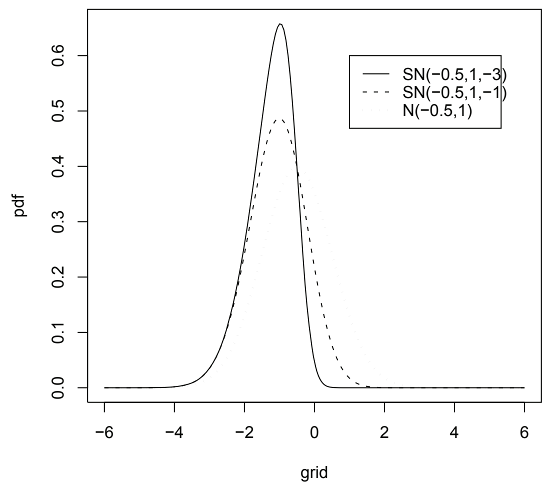

Figure 1. **Probability density function of three ability distributions.**

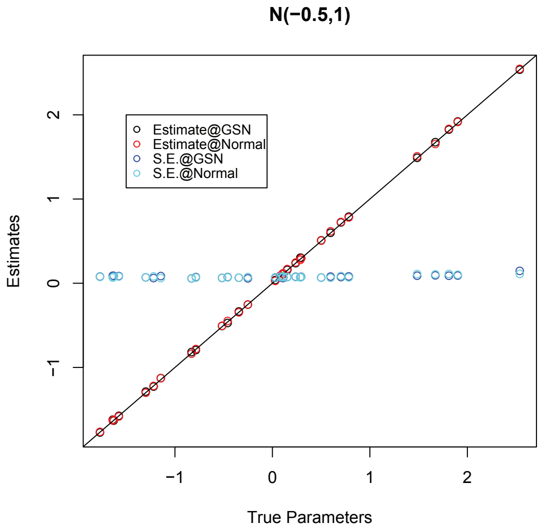

Figure 2. **One-parameter logistic b estimates and SE when data are generated from $N(-0.5, 1)$.**

Results

Rasch Model

Figures 2–4 present the estimates of the Rasch model b parameters as well as the standard error (SE) of these estimates for the data generated from $N(-0.5, 1)$, $SN(-0.5, 1, -1)$, and $SN(-0.5, 1, -3)$, respectively. The x axis in the plot represents the true parameter values, and the y axis represents the estimates. The solid line in each graph represents the symmetric line on which the estimates and the true parameters coincide. The dots that follow the solid line represent the estimates from using either a normal distribution or a GSN distribution as a fitting distribution. The dots that are fairly parallel to a horizontal line around $y = 0$ are the SEs of these estimates. One can note that the estimates and the SEs of the b parameters in the Rasch model are similar across different fitting distributions (normal or GSN). This means that misspecification of the latent ability distribution will have little effect on the parameter estimates in the Rasch model.

Figures 5–7 show the recovery of the quantiles $(0.1, 0.25, 0.5, 0.75, 0.9)$ of the latent ability distribution. In detail, the dashed line in each graph represents the symmetric line on which the estimates and the true values of the quantiles are equal to each other. The red line stands for the

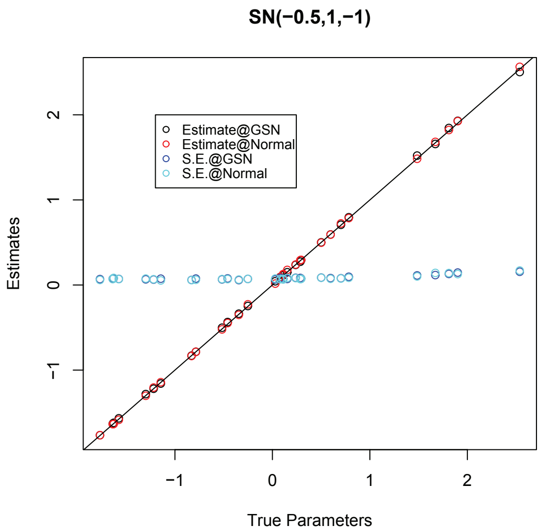

Figure 3. **Same as Figure 2, but for $SN(-0.5, 1, -1)$.**

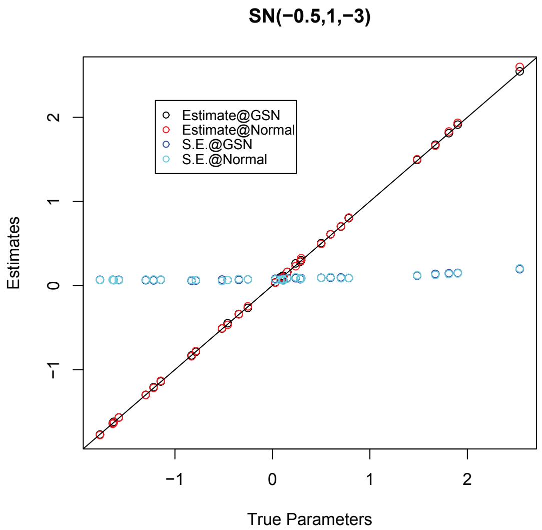

Figure 4. **Same as Figure 2, but for $SN(-0.5, 1, -3)$.**

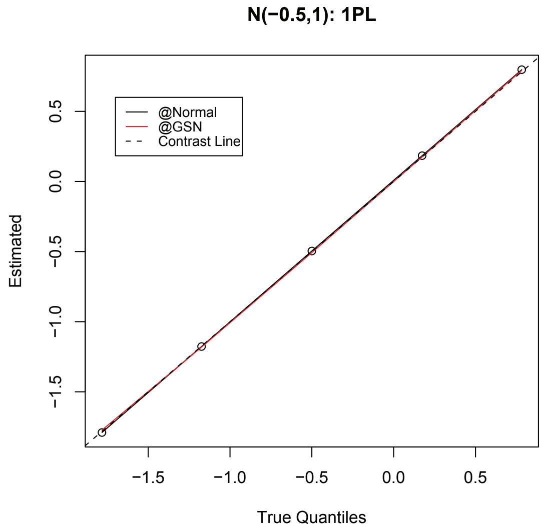

Figure 5. **Quantiles recovery when data are generated from $N(-0.5, 1)$ in the 1PL model.**

estimates from using a GSN to fit the data, whereas the black line represents the estimates from using a normal distribution to fit the data. One can observe that for the data generated from either the $N(-0.5, 1)$ or $SN(-0.5, 1, -1)$ distributions, the use of a normal distribution or a GSN has little effect on recovery of the true values of the quantiles. However, Figure 7 shows that for the data generated from an extreme skewed distribution, such as $SN(-0.5, 1, -3)$, the use of a normal distribution as a fitting distribution does not perform as well as using a GSN distribution in terms of quantile recovery. This can be explained by how well a generating distribution can be approximated by a normal distribution. To be more specific, the $SN(-0.5, 1, -1)$ was approximated well by $N(-1.06, 0.68^2)$ with small error (3.6^{-4}); however, the generating distribution $SN(-0.5, 1, -3)$ was so skewed that it could not be approximated by a normal distribution to a satisfactory level. Conversely, the GSN is flexible enough to recover the quantiles of this extremely skewed distribution.

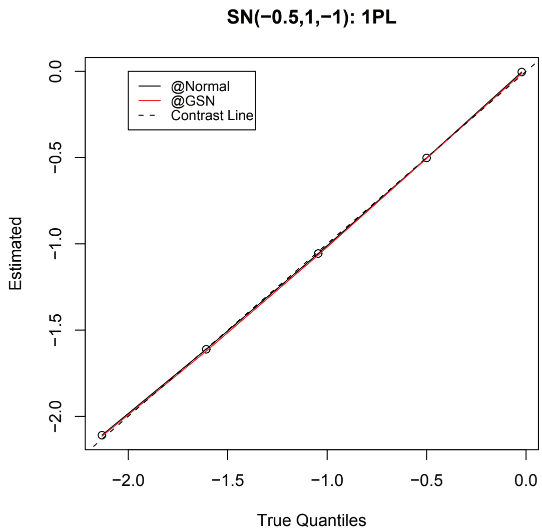

Figure 6. **Same as Figure 5, but for $SN(-0.5, 1, -1)$.**

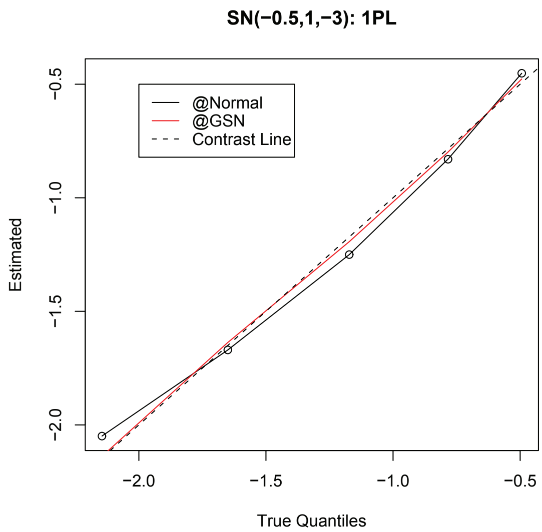

Figure 7. **Same as Figure 5, but for $SN(-0.5, 1, -3)$.**

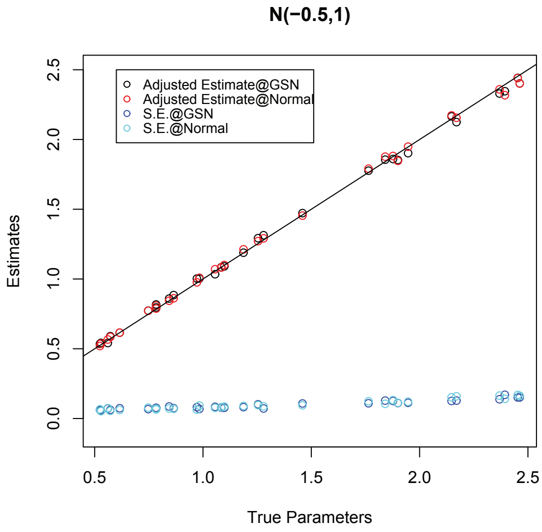

Figure 8. **Two-parameter logistic a estimates and SE when data are generated from $N(-0.5, 1)$.**

2PL Model

Figures 8–11 illustrate the estimates of a and b parameters in the 2PL model as well as the SE associated with these estimates. Again, the solid line in each graph is the symmetric line on which the estimates and true values are the same. The dots close to this line are the estimates from using different fitting distributions (i.e., normal distribution and GSE). The dots that are horizontal represent the SEs. It is noted that the estimates are close to the true values of the parameters, and the SEs are close to zero, no matter what fitting distributions were used to fit the data.

Figures 12 and 13 present the quantile recoveries of the latent ability distribution in the mixed 2PL model. These quantiles are 0.1, 0.25, 0.5, 0.75, and 0.9. Figure 12 shows the quantile recovery when the data are generated from $N(-0.5, 1)$, whereas in Figure 13, the data are generated from $SN(-0.5, 1, -1)$. The dashed line in each figure represents the symmetric line on which the estimates and true values coincide. The red line stands for the recovery from using GSN as the fitting distribution, whereas the black line represents the recovery from using a normal distribution. When using a GSN distribution to fit the data, the quantile recovery at both tails (such as 0.1, 0.25, and 0.9) is satisfactory, whereas the quantiles 0.5 and 0.75 deviate

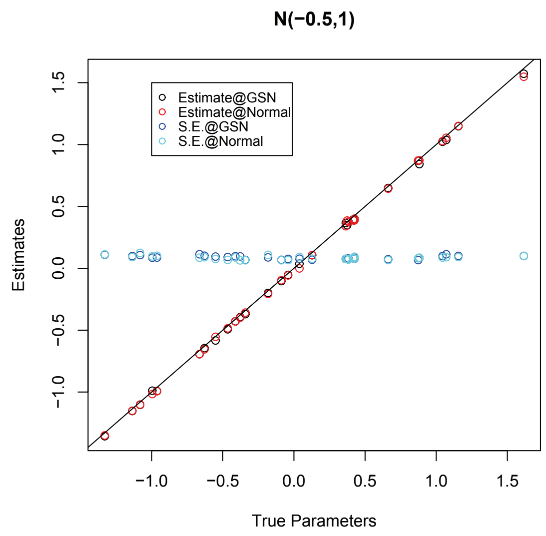

Figure 9. **Same as Figure 8, but for 2PL b estimates.**

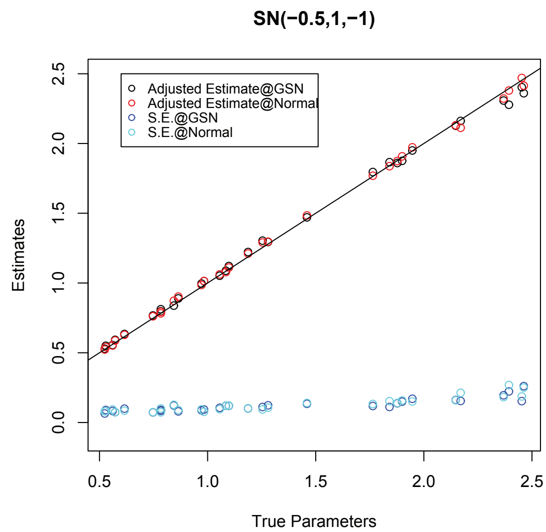

Figure 10. **Two-parameter logistic a estimates and SE when data are generated from $SN(-0.5, 1, -1)$.**

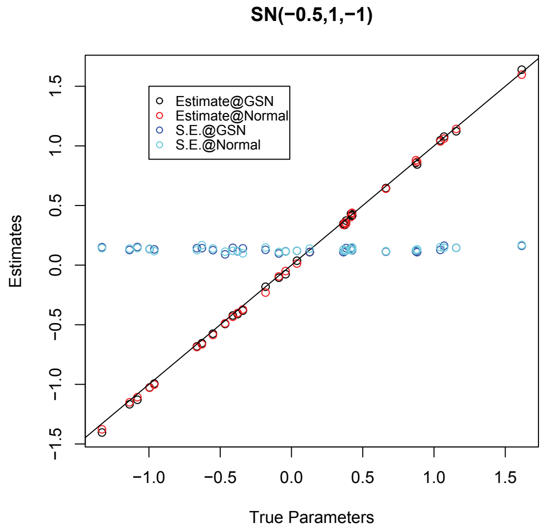

Figure 11. **Same as Figure 10, but for 2PL b estimates.**

slightly from the true values. When using a normal distribution to fit the data, the quantiles are recovered satisfactorily when the data are also generated from a normal distribution. Conversely, if the data are generated from a skewed normal distribution, the recovery of the quantiles at the middle positions deviates slightly from the true values. The results imply that the misspecification of the latent ability distribution has a slight or no effect on the quantile recoveries when the true latent ability is not extremely skewed. We did not report the results for the data generated from the extremely skewed distribution owing to a computational difficulty; however, it is expected that the recovery of the quantiles at the low and high tails of this distribution will be better than the recovery of the quantiles in the middle positions.

Discussion and Conclusion

Since latent ability plays an important role in the IRT modeling framework, researchers and practitioners always have the tendency to justify the latent ability distribution. The primary goal of this report was to discuss the effects of the latent ability distribution on the estimates of the item parameters as well as on the estimates of the ability distributional statistics.

An extended form of the EM algorithm used by Rigdon and Tsutakawa (1983) was applied in fitting the mixed Rasch and 2PL models when both item parameters and distributional parameters were estimated simultaneously. A simulation study was used to illustrate the

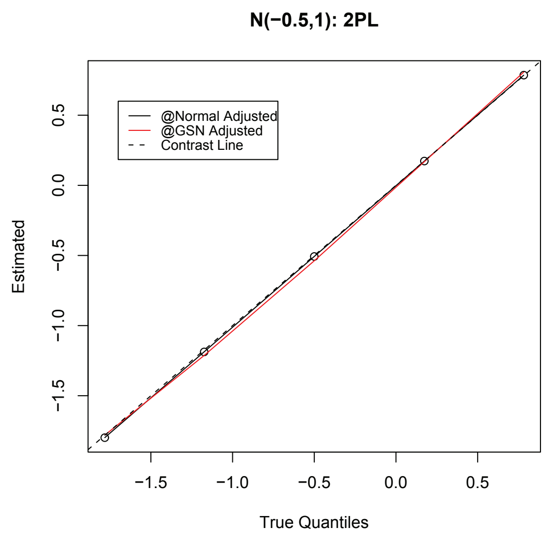

Figure 12. **Quantiles recovery when data are generated from $N(-0.5, 1)$ in the 2PL model.**

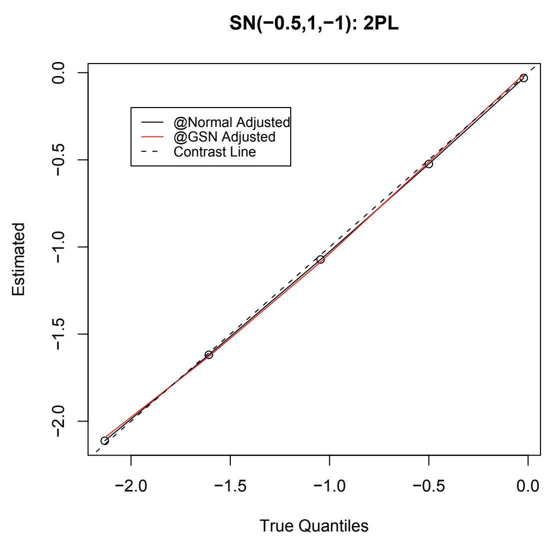

Figure 13. **Same as Figure 12, but for $SN(-0.5, 1, -1)$.**

performance of this algorithm and to check the stability of parameter estimates across different simulation scenarios. The responses of 1,000 simulated examinees on 30 items were generated either from a Rasch model or a 2PL model. The latent ability of each simulated examinee is generated either from a normal distribution or a SN distribution. The fitting distribution of the latent ability was either a normal distribution or a GSN distribution.

For the Rasch model, the item parameter estimates and the ability distributional statistics (including mean, variance, and the quantiles) are similar in all conditions. Even the estimation errors for these statistics are similar in all conditions.

For the 2PL model, the item parameter estimates and the ability distributional estimates are similar to each other for all situations; however, the SEs of the item parameter estimates are slightly larger when the data are generated from a SN distribution than when the data are generated from a normal distribution. This can be seen by comparing Figure 8 with Figure 11 or Figure 9 with Figure 11.

On the basis of the results of this study, we gain confidence in using a normal distribution to fit the mixed Rasch and 2PL models. In fact, there are several advantages to using a normal distribution over other distributions such as a log-cubical distribution (Aitkin & Aitkin, 2004) or a GSN (Ma et al.,2005). First, a normal distribution uses fewer parameters, which will benefit the estimation accuracy of the parameters. Second, a wide range of thin-tail distributions with slight skewness can be well approximated by a normal distribution. For example, in our study, the distribution $SN(-0.5, 1, -1)$ was well approximated by a normal distribution with mean -1.064 and variance 0.682 with error $3.6e-4$. A GSN distribution is more desirable than a normal distribution when the latent ability is extremely skewed such as $SN(-0.5, 1, -3)$. However, it is rare to encounter severe skewness for the latent ability distribution in practice. Moreover, the possible large skewness of the latent ability distribution can be reduced by developing a test to cover a wide range of item difficulties and item discriminations.

All the latent ability distributions discussed in this report are unimodal. The GSN distribution might have an advantages over the normal distribution in the case where the latent ability distribution is a multimodal distribution. Future research can be conducted along these lines.

References

Aitkin, M., & Aitkin, I. (2004). *Percentile estimation for the ability distribution in item response models.* Unpublished technical report.

Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew normal distribution. *Biometrika, 83*, 715–726.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika, 46*, 443–459.

Bouezmarni, T., Rijmen, F. & De Boeck, P., (2008). *Smoothed kernel for generalized liner mixed model.* Retrieved from
$http://www.dms.umontreal.ca/\ bouezma/bouezmarni_Rijmen_deBoeck.pdf$.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

Genton, M. (2004). Skew-symmetric and generalized skew-elliptical distributions. In Genton, M. (Ed.) *Skew-elliptical distributions and their applications - A journal beyond normality.* New York, NY: Chapman & Hall.

Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences.* Thousand Oaks, CA: Sage.

Laird, N. M. (1978). Nonparametric likelihood estimation of a mixing distribution. *Journal of the American Statistical Association, 73*, 805–811.

Lindsay, B. G., Clogg, C. C., & Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models. *Journal of the American Statistical Association, 86*, 96–107.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, UK: Addison-Wesley.

Ma, Y., Genton, M. G., & Tsiatis, A. A. (2005). Locally efficient semiparametric estimators for generalized skew-elliptical distributions. *Journal of the American Statistical Association, 100*, 980–989.

Mislevey, R., & Bock, R. D. (1989). BILOG 3: Item analysis and test scoring using binary logistic models [Computer software]. Mooresville, IN: Scientific Software.

Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data.* New York, NY: Springer.

Muraki, E., & Bock, R. D. (2003). PARSCALE 4: IRT item analysis and test scoring for rating-scale data [Computer software]. Chicago, IL: Scientific Software.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Paedogogiske Institute.

Rigdon, S. E., & Tsutakawa, R. K. (1983). Parameter estimation in latent trait models. *Psychometrika, 48*, 567–574.

Sanathanan, L., & Blumenthal, S. (1978). The logistic model and estimation of latent structure. *Journal of the American Statistical Association, 73*, 794–799.

Woods, C. M., & Thissen, D. M. (2004). *Item response theory with estimation of the latent population distribution using spline based densities.* Paper presented at the annual meeting of the Psychometric Society, Pacific Grove, CA.

Xu, X. (2007). *Estimating the latent ability distribution by using general skew-elliptical distribution.* Paper presented at the annual meeting of the National Council of Measurement in Education, Chicago, IL.