A note on identification and metric issues for skew IRT models

Caio L. N. Azevedo, IMECC/UNICAMP, Brazil
Heleno Bolfarine, IME/USP, Brazil
Dalton F. Andrade, INE/UFSC, Brazil

Abstract

The skew-normal distribution (SND) is a flexible family of densities which preserves some useful properties of the original normal distribution. Some stochastic representations for the SND have been proposed in the literature. The Henze (H) and Sahu, Branco and Dey (SBD) are the two most used ones. On the other hand, the centered parametrization is useful for inference purposes. The main goals of this article are: establish a link between the standard H and SDB skew-normal distributions and use this result to model the latent traits for IRT models. We proved that standard H and SDB distributions are related to each other through a function of the asymmetry parameter and also that they are exactly the same under centered parametrization (CP). Using these results, we showed that the common density obtained through the CP is useful to model the latent traits for unidimensional IRT models. This approach allows to represent asymmetric latent traits behavior and ensures the model identification as well.

Key words : Skew-normal distribution, centered parametrization, IRT, model identification.

MSC : 62F10.

1 Introduction

The skew-normal distribution is a flexible family of densities which preserves some useful properties of the original distribution, see [Genton(2004)] for a broad discussion. It has been used in many fields to model asymmetric data, including in educational assessments, according to [Bazán et al.(2004)Bazán, Bolfa and [Azevedo et al.(2009a)Azevedo, Bolfarine, and Andrade], for example. In order to facilitate inferential aspects, some stochastic representations were proposed in the literature. Two well known examples are that ones considered by [Henze(1986)] and [Sahu et al.(2003)Sahu, Dey, and Branco], henceforth Henze (H) and Sahu, Dey and Branco (SDB) representations. They have different characteristics and, in fact, produce different kinds of skew-normal distributions. The respective densities are given by :

$$p(y^{(H)}) = \frac{2}{\sigma}\phi\left(\frac{y^{(H)}-\tau}{\sigma}\right)\Phi\left[\lambda\left(\frac{y^{(H)}-\tau}{\sigma}\right)\right] \tag{1}$$

$$p(y^{(S)}) = \frac{2}{\sqrt{\sigma^2+\lambda^2}}\phi\left(\frac{y^{(H)}-\tau}{\sqrt{\sigma^2+\lambda^2}}\right)\Phi\left[\frac{\lambda}{\sigma}\left(\frac{y^{(H)}-\tau}{\sqrt{\sigma^2+\lambda^2}}\right)\right], \tag{2}$$

where $Y^{(H)}$ and $Y^{(S)}$ stands for the H and SDB skew normal distributions, respectively, τ represents the location parameter, σ^2 is the dispersion parameter and λ is the asymmetry parameter. Furthermore, in order to avoid some inference problems, [Azzalani(1985)] proposed the centered reparametrization which consists in letting the mean and variance free of the asymmetry parameter. This reparametrization has been widely used considering the H distribution but not for SDB one. For further details, the reader is refereed to [Genton(2004)], [Azzalani(1985)], [Henze(1986)] and [Sahu et al.(2003)Sahu, Dey, and Branco], for example.

On the other hand, there is a growing interest in using asymmetric distributions to model the latent traits in IRT, see for example [Bazán et al.(2004)Bazán, Bolfarine, and Leandro], [Azevedo et al.(2009a)Azeve and [Azevedo et al.(2009b)Azevedo, Bolfarine, and Andrade]. However, there are some identification issues related to the using of the skew-normal distributions to model the latent traits.

The main goal of this work is to study a link between standard H and SDB densities and show that they are exactly the same under centered parametrization. Furthermore, we proved that this common density is a Henze type skew normal density and, therefore, admits the correspondent stochastic representation. We will also indicate how the using of such distribution may help to provide identifiability with latent trait asymmetric distribution for IRT models.

This article is organized as follows. In the Section 1 we reviewed some aspects os skew normal distribution and we described the problem of interest. In the Section 2 we will sutdy the relationships between the H and SDB distributions. In Section 3 we will apply the latter results in some IRT models and in the Section 4 we will outline some appropriated conclusions and comments.

2 The standard H and SDB skew-normal distributions

As we will show latter, it is suitable to build a latent skew normal centered distribution based on the usual standard parametrization of (1) and (2), that is, by setting ($\tau = 0$ and $\sigma^2 = 1$), and we will developed all the results using such densities. Henceforth, we will consider only the standard versions of H and SDB distributions, unless the contrary is stated.

Let $Y^{(H)} \sim SN(0, 1, \lambda)$ and $Y^{(S)} \sim SN(0, 1, \lambda)$ two independent random variables with standard Henze's and Sahu's skew-normal distributions, respectively. In both cases, λ remain as the asymmetry parameter. For details see [Henze(1986)], [Sahu et al.(2003)Sahu, Dey, and Branco] and [Genton(2004)]. The correspondent densities, expectations and variances are given by:

$$
\begin{aligned}
p(y^{(H)}) &= 2\phi\left(y^{(H)}\right)\Phi\left[\lambda\left(y^{(H)}\right)\right], \\
\mathbb{E}(Y^{(H)}) &= \delta\sqrt{\frac{2}{\pi}}, && (3)\\
\mathbb{V}ar(Y^{(H)}) &= 1 - \frac{2}{\pi}\delta^2, && (4)\\
\delta &= \frac{\lambda}{\sqrt{1+\lambda^2}}, && (5)
\end{aligned}
$$

and

$$
\begin{aligned}
p(y^{(S)}) &= \frac{2}{\sqrt{1+\lambda^2}}\phi\left(\frac{y^{(H)}}{\sqrt{1+\lambda^2}}\right)\Phi\left[\frac{\lambda}{\sqrt{1+\lambda^2}}\left(y^{(H)}\right)\right], \\
\mathbb{E}(Y^{(S)}) &= \lambda\sqrt{\frac{2}{\pi}}, && (6)\\
\mathbb{V}ar(Y^{(S)}) &= 1 + \left(1 - \frac{2}{\pi}\right)\lambda^2. && (7)
\end{aligned}
$$

Proposition 2.1. *Let $Y^{(H)}$ and $Y^{(S)}$ the H and SBD skew normal distributions, then the Pearson's skewness coefficient in both case are the same and given by:*

$$\gamma = r\delta^3\left(2r^2-1\right)\left(1-r\delta^2\right)^{-3/2},$$

where $r = \sqrt{\frac{2}{\pi}}$.

Proof: For $Y^{(H)}$ see [Henze(1986)], for example. For $Y^{(S)}$, by definition, we have that:

$$\gamma_{(Y^{(S)})} = \frac{\mathbb{E}\left[(Y^{(S)} - \mathbb{E}(Y^{(S)})\right]^3}{\left[\mathbb{V}ar(Y^{(S)})\right]^{1/3}}. \tag{8}$$

On the other hand, denoting $\mu_S = \mathbb{E}(Y^{(S)})$, it follows that:

$$\mathbb{E}\left[(Y^{(S)} - \mathbb{E}(Y^{(S)})\right]^3 = \mathbb{E}\left[(Y^{(S)})^3\right] - \mu_S^2\mathbb{E}\left[(Y^{(S)})^2\right] + 2\mu_s^3. \tag{9}$$

By using the results of [Sahu et al.(2003)Sahu, Dey, and Branco], the moment generating function of $Y^{(S)}$ is given by:

$$M_{Y^{(S)}}(t) = 2\Phi\left(\lambda t\right)\exp\left\{\frac{t^2}{2}\left(1+\lambda^2\right)\right\}.$$

After some cumbersome but easy algebra we find:

$$M^{(2)}_{Y^{(S)}}(0) = \mathbb{E}\left[(Y^{(S)})^2\right] = 1 + \lambda^2, \tag{10}$$

$$M^{(3)}_{Y^{(S)}}(0) = \mathbb{E}\left[(Y^{(S)})^3\right] = 2\lambda^3\sqrt{\frac{2}{\pi}} + 3\lambda\sqrt{\frac{2}{\pi}}, \tag{11}$$

where $M^{(k)}_{Y^{(S)}}(0) = \left.\frac{\partial M_{Y^{(S)}}(t)}{\partial t^{(k)}}\right|_{t=0}$.

Therefore, using (6), (10) and (11) in (9), it follows that:

$$\mathbb{E}\left[\left(Y^{(S)} - \mathbb{E}(Y^{(S)})\right)\right]^3 = \lambda^3\sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right) . \tag{12}$$

Finally, using (5), (7) and (12) in (8), and gathering the terms properly, we have:

$$\gamma_{(Y^{(S)})} = \frac{\lambda^3\sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right)}{\left(1 + \left(1 - \frac{2}{\pi}\right)\lambda^2\right)^{3/2}} = r\delta^3\left(2r^2 - 1\right)\left(1 - r\delta^2\right)^{-3/2} = \gamma_{(Y^{(H)})} .$$

■

Proposition 2.2. *Let $Y^{(H)}$ and $Y^{(S)}$ the H*

$$Y^{(H)} = \frac{Y^{(S)}}{\sqrt{1+\lambda^2}} .$$

Proof : From [Henze(1986)] and [Sahu et al.(2003)Sahu, Dey, and Branco] we have that

$$\begin{aligned} Y^{(H)} &= \delta T + \sqrt{1-\delta^2}X , \\ Y^{(S)} &= \lambda T + X , \end{aligned} \tag{13}$$

where $T \sim HN(0,1)$ and $X \sim N(0,1), T \perp X$. On the other hand, from (5) it follows that :

$$\begin{aligned} \lambda &= \frac{\delta}{\sqrt{1-\delta^2}} \\ \frac{1}{\sqrt{1+\lambda^2}} &= \frac{1}{\sqrt{1+\frac{\delta^2}{1-\delta^2}}} = \sqrt{1-\delta^2} . \end{aligned} \tag{14}$$

Therefore, dividing $Y^{(S)}$ by $\sqrt{1+\lambda^2}$ and using (5) and (14) it comes that:

$$\left(\frac{\lambda}{\sqrt{1+\lambda^2}}T + \frac{1}{\sqrt{1+\lambda^2}}X\right) = \delta T + \sqrt{1-\delta^2}X = Y^{(H)}\,.$$

Then, under direct parametrization, the univariate Henze and SBD skew normals are related.

■

On the hand, in order to circumvent some inference problems, [Azzalani(1985)] proposed a centered parametrization which can be described as:

$$Y_c^{(\cdot)} = \sqrt{\psi}\left(\frac{Y^{(\cdot)} - \mathbb{E}(Y^{(\cdot)})}{\sqrt{\mathbb{V}ar(Y^{(\cdot)})}}\right) + \mu\,, \tag{15}$$

where (.) stands for some stochastic representation. We will use the notation $Y_c^{(\cdot)} \sim SN_c(\mu, \psi, \gamma)$ to stands for a centered skew normal distribution (CSN) under some stochastic representation. In this case μ and ψ are the mean and variance, respectively. The CSN is parametrized by γ, which is the asymmetry coefficient, instead of λ. Considering the centered parametrization $\mathbb{E}(Y_c^{(\cdot)}) = \mu$ and $\mathbb{V}ar(Y_c^{(\cdot)}) = \psi$ independently from the stochastic representation used. Notice also, that it is possible to consider a CSN, with any mean and variance, from a standard skew-normal distribution (which has the location and dispersion parameters fixed at 0 and 1, respectively). Furthermore, the skewness coefficient of $Y_c^{(\cdot)}$ and $Y^{(\cdot)}$ are the same, see [Azzalani(1985)]. Another feature is stated in the following theorem:

Theorem 2.1. *Let $Y_c^{(H)}$ and $Y_c^{(S)}$ the centered version of H and SDB skew normal distributions. Then:*

$$p(y_c^{(H)}) = p(y_c^{(S)}) = p(y_c)\,,$$

where Y_c is a Henze type skew normal variable.

Proof:

For both cases, from Proposition (2.1), the Pearson's skewness coefficient is given by:

$$\gamma = r\delta^3\left(2r^2-1\right)\left(1-r\delta^2\right)^{-3/2},$$

which implies, after some algebra, that:

$$\lambda = \frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\left(r^2-1\right)}}, \tag{16}$$

$$\delta = \frac{r^{-1}s\gamma^{1/3}}{\sqrt{1+s^2\gamma^{2/3}}}, \tag{17}$$

where $s = \left(\frac{2}{4-\pi}\right)^{1/3}$. By using the Jacobian method and the equation (15), it follows that:

$$\begin{aligned} p(y_c^{(H)}) &= \frac{1}{\sqrt{\psi}}\left[\sqrt{\mathbb{V}ar(Y^{(H)})}\right]\phi\left[\sqrt{\mathbb{V}ar(Y^{(H)})}\left(\frac{y_c^{(H)}-\mu}{\sqrt{\psi}}\right)+\mathbb{E}(Y^{(H)})\right] \\ &\times \Phi\left\{\lambda\left[\sqrt{\mathbb{V}ar(Y^{(H)})}\left(\frac{y_c^{(H)}-\mu}{\sqrt{\psi}}\right)+\mathbb{E}(Y^{(H)})\right]\right\}. \end{aligned} \tag{18}$$

On the other hand, using (17) in (3) and (4), it comes that:

$$\mathbb{E}(Y^{(H)}) = \frac{s\gamma^{1/3}}{\sqrt{1+s\gamma^{2/3}}}, \tag{19}$$

$$\mathbb{V}ar(Y^{(H)}) = \frac{1}{1+s\gamma^{2/3}}. \tag{20}$$

Using (19) and (20) in (18), and gathering the terms properly, we obtain:

$$\begin{aligned}
p(y_c^{(H)}) &= \frac{1}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\phi\left\{\frac{y_c^{(H)}-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right\} \\
&\times \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\left(r^2-1\right)}}\left[\frac{y_c^{(H)}-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right]\right\} \\
&= \frac{1}{\sqrt{\psi^*}}\phi\left\{\frac{y_c^{(H)}-\mu^*}{\sqrt{\psi^*}}\right\} \\
&\times \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\left(r^2-1\right)}}\left[\frac{y_c^{(H)}-\mu^*}{\sqrt{\psi^*}}\right]\right\}, \qquad (21)
\end{aligned}$$

where $\mu^* = \mu - s\gamma^{1/3}$ and $\psi^* = 1 + s^2\gamma^{2/3}$.

On the other hand, for $Y_c^{(S)}$ we have that:

$$\begin{aligned}
p(y^{(S)}) &= \left[\sqrt{\frac{\mathbb{V}ar(Y^{(H)})}{1+\lambda^2}}\right]\phi\left(\sqrt{\mathbb{V}ar(Y^{(S)})}\left(\frac{y^{(S)}-\mu}{\sqrt{\psi}}\right)+\mathbb{E}(Y^{(S)})\right) \\
&\times \Phi\left[\frac{\lambda}{1+\lambda^2}\left(\sqrt{\mathbb{V}ar(Y^{(S)})}\left(\frac{y^{(S)}-\mu}{\sqrt{\psi}}\right)+\mathbb{E}(Y^{(S)})\right)\right]. \qquad (22)
\end{aligned}$$

Using (16) in (6) and (7), it comes that:

$$
\begin{aligned}
\mathbb{E}(Y^{(H)}) &= \frac{sr\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\,(r^2-1)}}\,, & (23)\\
\mathbb{V}ar(Y^{(H)}) &= \frac{r^2}{r^2+s^2\gamma^{2/3}\,(r^2-1)}\,, & (24)\\
\frac{\mathbb{E}(Y^{(S)})}{\sqrt{\mathbb{V}ar(Y^{(S)})}} &= s\gamma^{1/3}\,, & (25)\\
\sqrt{\frac{\mathbb{V}ar(Y^{(S)})}{1+\lambda^2}} &= \frac{r^2\sqrt{1+s^2\gamma^{2/3}}}{r^2+s^2\gamma^{2/3}\,(r^2-1)}\,, & (26)\\
\delta\sqrt{\mathbb{V}ar(Y^{(S)})} &= \frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\,(r^2-1)}\sqrt{1+s^2\gamma^{2/3}}}\,. & (27)
\end{aligned}
$$

Finally, from (23) to (27) in (22), and gathering the terms properly, we obtain:

$$
\begin{aligned}
p(y_c^{(S)}) &= \frac{1}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\phi\left\{\frac{y_c^{(S)}-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right\}\\
&\times\ \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\,(r^2-1)}}\left[\frac{y_c^{(S)}-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right]\right\}\\
&= \frac{1}{\sqrt{\psi^*}}\phi\left\{\frac{y_c^{(S)}-\mu^*}{\sqrt{\psi^*}}\right\}\\
&\times\ \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\,(r^2-1)}}\left[\frac{y_c^{(S)}-\mu^*}{\sqrt{\psi^*}}\right]\right\}\,. \qquad (28)
\end{aligned}
$$

Therefore, comparing (21) with (28), we conclude that:

$$
p(y_c^{(H)}) = p(y_c^{(S)}) = p(y_c)\,. \qquad (29)
$$

Notice also that the density (29) can be written as:

$$
\begin{aligned}
p(y_c) &= \frac{1}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\phi\left\{\frac{y_c^{(S)}-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right\} \\
&\times \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\left(r^2-1\right)}}\left[\frac{y_c-\left(\mu-s\gamma^{1/3}\right)}{\sqrt{\psi\left(1+s\gamma^{2/3}\right)}}\right]\right\} \\
&= \frac{1}{\sqrt{\psi^*}}\phi\left\{\frac{y_c-\mu^*}{\sqrt{\psi^*}}\right\}\Phi\left\{\lambda^*\left[\frac{y_c^{(S)}-\mu^*}{\sqrt{\psi^*}}\right]\right\}, \qquad (30)
\end{aligned}
$$

where $\lambda^* = \frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}(r^2-1)}}$. From (1) we can notic that (30) is, in fact, a Henze type skew normal density. Therefore, from [Henze(1986)], we are able to write:

$$
Y_c = \left(\delta^* T + \sqrt{1-(\delta^*)^2}X\right)\sqrt{\psi^*} + \mu^*, \qquad (31)
$$

where T and X are as defined in (13). The representation given by (31) is useful for inference purposes.

■

Corollary 2.1. *Taking $\mu = 0$ and $\psi = 1$ in Theorem 2.1 we have:*

$$
\begin{aligned}
p(y_c^{(H)}) &= p(y_c^{(S)}) = \frac{1}{\sqrt{\left(1+s\gamma^{2/3}\right)}}\phi\left\{\frac{y_c^{(H)}+s\gamma^{1/3}}{\sqrt{\left(1+s\gamma^{2/3}\right)}}\right\} \\
&\times \Phi\left\{\frac{s\gamma^{1/3}}{\sqrt{r^2+s^2\gamma^{2/3}\left(r^2-1\right)}}\left[\frac{y_c^{(H)}+s\gamma^{1/3}}{\sqrt{\left(1+s\gamma^{2/3}\right)}}\right]\right\}.
\end{aligned}
$$

Proof: It follows directly from Theorem 2.1.

■

3 Application in IRT : latent skew normal distributions

Item response theory is a set of models which deal with certain non-observable attributes, typically named latent traits, see [Lord(1980)] and [Baker and Kim(2004)], for example. Usually, they are treated as random variables and some convenient distribution is considered to them. In this section we will consider the centered skew normal distribution to model the latent traits. We will focus on the one group and the multiple groups framework, according to [Bock and Zimowski(1997)]. In order to illustrate the concepts, we consider the two parameters model, see [Baker and Kim(2004)]. However, the result can be extended to other unidimensional univariate models.

The general framework for one group situation in IRT consists on a set of n examinees answering a test with I items. Let Y_{ij} the answer of the examinee j to the item i, which is assumed to follow a Bernoulli distribution, with 1 indicating the right answer and 0 otherwise, such that:

$$P(Y_{ij}=1|\theta_j,\boldsymbol{\zeta}_i)=\phi\left(a_i\theta_j-b_i\right)\ , \tag{32}$$

where θ_j is the latent trait and (a_i,b_i) are the item parameters. For interpretation and meaning of the parameters, see [Baker and Kim(2004)] for example.

Let us suppose, additionally, that :

$$\theta_j \overset{i.i.d.}{\sim} Y^{(H)}. \tag{33}$$

Under the assumption (33) we have that $\mathbb{E}(Y^{(H)|\lambda})=\delta\sqrt{\frac{2}{\pi}}$ and $\mathbb{V}ar(Y^{(H)}|\lambda)=1-\frac{2}{\pi}\delta^2$, see equations (3) and (4). In this way, the metric of the latent traits is not defined and, if no restrictions are imposed on the item parameters, the model (32) is not identified. This happens because:

$$\begin{aligned}
P(Y_{ij}=1|\theta_j,\boldsymbol{\zeta}_i) &= \Phi(a_i(\theta_j-b_i))=\Phi\left(\frac{a_i}{\alpha}\left(\alpha\theta_j-\alpha b_i\right)\right)\\
&= \Phi\left(\frac{a_i}{\alpha}\left(\alpha\theta_j+\beta-\alpha b_i-\beta\right)\right)\\
&= \Phi\left(-a_i^*\left(\theta_j^*-b_i^*\right)\right)\ ,
\end{aligned} \tag{34}$$

where $\alpha, \beta \in I\!R$ are constants. Therefore, different sets of parameter values can lead to the same likelihood. This means that is not sufficient to set $\mu = 0$ and $\sigma^2 = 1$ in (1) in order to identify the model (32). The same argument is also valid if we consider that $\theta_j \sim Y^{(S)}$. However, assuming that:

$$\theta_j \sim SN_c(0, 1, \lambda_\theta)\,, \tag{35}$$

we will have that $I\!E(\theta_j|\lambda_\theta) = 0$ and $I\!Var(\theta_j|\lambda_\theta) = 1$. This ensures that not only the model (32) is identified but also that all estimates will lie on the metric (0,1), see [Baker and Kim(2004)], for example. Therefore, the using of the CSN assures the identification of the model (32). Notice that, trasformations as stated in (34) are no longer possible. Actually, we claim that, if no restrictions are imposed to item parameters, it is necessary and sufficient to assume (35), in order to have the IRT model identified with the assumption of skew normal distribution to the latent traits.

In a presence of K multiple groups with n_k, $k = 1, ..., K$ examinees each one, we consider that every group is submitted a test with I_k items. Let Y_{ijk} the answer of the examinee j, from group k to the item i, a Bernoulli random variable, with 1 indicating the right answer and 0 otherwise, such that:

$$P(Y_{ijk} = 1|\theta_{jk}, \boldsymbol{\zeta}_i) = \phi\left(a_i\theta_{jk} - b_i\right)\,.$$

Analogously to the above case and using the usual assumptions for multiple groups model, see [Bock and Zimowski(1997)], the following restriction:

$$\begin{aligned} \theta_{j1} &\sim SN_c(0, 1, \lambda_{\theta_1})\,, \\ \theta_{jk} &\sim SN_c(\mu_{\theta_k}, \psi_{\theta_k}, \lambda_{\theta_k}), k = 2, ..., K\,. \end{aligned}$$

is sufficient in order to identify the model (36). The interpretations remain the same. Analogously to the one group model, notice that is not possible to consider the following kind of transformations:

$$
\begin{aligned}
P(Y_{ij1}=1|\theta_{j1},\boldsymbol{\zeta}_i) &= \Phi(a_i(\theta_{j1}-b_i)) = \Phi\left(\frac{a_i}{\alpha}\left(\alpha\theta_{j1}-\alpha b_i\right)\right) \\
&= \Phi\left(\frac{a_i}{\alpha}\left(\alpha\theta_{j1}+\beta-\alpha b_i-\beta\right)\right) \\
&= \Phi\left(-a_i^*\left(\theta_{j1}^*-b_i^*\right)\right) . \qquad (36)
\end{aligned}
$$

Notice also that, if we consider $(\boldsymbol{\theta}_{..}^{(1)'},\boldsymbol{\zeta}^{(1)'})'$ and $(\boldsymbol{\theta}_{..}^{(2)'},\boldsymbol{\zeta}^{(2)'})'$, where $\boldsymbol{\theta}_{..}=(\theta_{11},...,\theta_{n_kK})'$ and $\boldsymbol{\zeta}=(\boldsymbol{\zeta}_1',...,\boldsymbol{\zeta}_I')'$, two different sets of values for the parameters, then:

$$
L(\boldsymbol{\theta}_{..}^{(1)},\boldsymbol{\zeta}^{(1)}) \neq L(\boldsymbol{\theta}_{..}^{(2)},\boldsymbol{\zeta}^{(2)}) ,
$$

once that is not possible to consider transformations as stated in (36). Therefore, in this case the model is also identified and the latent trait scale is fixed in the metric (0,1) as well.

4 Conclusions and final remarks

The skew-normal distribution provides a flexible and easy interpretable way to model asymmetric data and preserves some interesting properties of the usual normal distribution. Also, some stochastic representations were proposed for it, which are useful for inference purposes. The Henze and Sahu, Branco and Dey are the two widely used ones. On the other hand, the centered parametrization helps in avoiding problems concerning some inferential aspects.

On the other hand, the skew-normal offers a promising possibility of flexible and easy interpretation way to model the latent traits in IRT. However, its using requires a special attention concerning the model identification and, consequently, the stablishment of the latent traits metric.

In this article we established a link between the two aforementioned representations and we proved that, in the univariate case for the centered standard parametrization, they generate the same density. Furthermore, we proposed a sufficient condition for the model identification using the skew normal distribution in IRT. We claim that this condition is also necessary if no additional restrictions are imposed to the item parameters. Such condition is based on the using of the centered parametrization of a standard skew normal distribution. Such approach can be useful to analyze asymmetric IRT data as well to provide some directions for the IRT model identification in more complex cases, for example for the multidimensional models.

5 Acknowledgements

The authors would like to thank FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), for the finnancial support through the process number 2008/50046-9.

References

[Azevedo et al.(2009a)Azevedo, Bolfarine, and Andrade] C. L. N. Azevedo, H. Bolfarine, and D. F. Andrade. Bayesian inference for a skew-normal irt model under the centered parametrization: model identification and estimation. *Manuscript under preparation*, 2009a.

[Azevedo et al.(2009b)Azevedo, Bolfarine, and Andrade] C. L. N. Azevedo, H. Bolfarine, and D. F. Andrade. Bayesian inference for a multiple groups model with skew normal latent distributions. *Manuscript under preparation*, 2009b.

[Azzalani(1985)] A. Azzalani. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12:171–178, 1985.

[Baker and Kim(2004)] F. B. Baker and Seock-Ho Kim. *Item response theory : Parameter Estimation Techniques*. Marcel Dekker, Inc, New York, NY, second edition, 2004.

[Bazán et al.(2004)Bazán, Bolfarine, and Leandro] J. L. Bazán, H. Bolfarine, and R. A. Leandro. Bayesian estimation via mcmc for probit-normal model in item response theory. *Technical report, University of São Paulo*, (15), 2004.

[Bock and Zimowski(1997)] D. R. Bock and M. F. Zimowski. *The Multiple Groups IRT, in Handbook of modern Item Response Theory*. Wim J. van der Linden and Ronald K. Hambleton eds, Springer-Verlag, New York, NY, first edition, 1997.

[Genton(2004)] M. Genton. *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Chapman & Hall, London, first edition, 2004.

[Henze(1986)] N. Henze. A probalistic representation of the skew-normal distribution. *Scandinavian journal of statistics*, 13:271–275, 1986.

[Lord(1980)] F. M. Lord. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale : Lawrence Erlbaum Associates, Inc, first edition, 1980.

[Sahu et al.(2003)Sahu, Dey, and Branco] Sujit K. Sahu, Dipak K. Dey, and Marcia D. Branco. A new class of multivariate skew distributions with applications to bayesian regression models. *The Canadian Journal of Statistics*, 31(2):129–150, 2003.