

A Family of Truncated Skew-Normal Distributions

Hea-Jung Kim¹⁾

Abstract

The paper extends earlier work on the skew-normal distribution, a family of distributions including normal, but with extra parameter to regulate skewness. The present work introduces a singly truncated parametric family that strictly includes a truncated normal distribution, and studies its properties, with special emphasis on the relation with bivariate normal distribution.

Keywords : skew-normal distribution, singly truncated distribution, conditional bivariate normal distribution.

1. Introduction

A random variable W is said to be skew-normal with parameter θ , written $W \sim SN(\theta)$, if its density function is

$$2 \phi(w) \Phi(\theta w), \quad -\infty < w < \infty, \quad (1)$$

where $\phi(w)$ and $\Phi(w)$ denote the standard normal and distribution function, respectively; the parameter θ which regulates the skewness varies in $(-\infty, \infty)$, and $\theta=0$ corresponds to the standard normal density. The random variable W can be expressed in terms of two independent standard normal random variables U and V :

$$W = (V + \theta |U|) / \sqrt{1 + \theta^2}. \quad (2)$$

A systematic treatment of the distribution, developed independently from earlier works, has been given by Azzalini (1985) and Henze (1986). The distribution is suitable for the analysis of data exhibiting a unimodal empirical distribution but with some skewness present, a situation often occurring in practical problems. See Kim (2002) and references therein for the applications of the distribution. As for extensions of the distribution, many classes of distributions are proposed. Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (1999) discussed the multivariate extension of the distribution whose the marginal distributions are scalar skew normal; Branco and Dey (2001) proposed a general class of multivariate skew-elliptical distributions which includes the multivariate extension; Kim (2002) suggested a

1) Professor, Department of Statistics, Dongguk University, Seoul 100-715, KOREA
E-mail : kim3hj@dongguk.edu

class of skew distributions obtained from various scale mixtures of the skew normal distribution.

The major goal of this paper is to introduce a family of truncated skew-normal distributions as a variant of the skew-normal distribution. The interest in the family comes from both theoretical and applied direction. On the theoretical side it enjoys a number of formal properties which resemble those of the skew-normal distribution and produces a class of singly truncated distributions. Further it defines a class of conditional distributions of Y_1 given $Y_i > 0$ (or $Y_i < 0$, for $i=1,2$) when (Y_1, Y_2) has a bivariate normal distribution. This enables us inference for the marginal distribution of Y_1 when the Y_1 and Y_2 variables are truncated in each form of the condition, but only the Y_1 values are observed. Therefore, in the applied viewpoint, the class of distributions provides yet another models that enable us to analyze a truncated (or censored) data set. Immediate examples are personnel selection, clinical study and other screening procedures where data set at hand is frequently based upon an individuals score on one or more screening variables. Arnold et al. (1993) and Cohen (1991) provide examples and a review of the literature in this area, respectively.

The rest of the paper is organized as follows. In section 2 we derive our proposed family of distributions. Several properties and a possible extension are provided for the family. Section 3 provides an application of the family of distributions using a constrained regression problem. This paper is concluded in section 4 with brief discussion.

2. The Family of Distributions

The present section is devoted to the study of the distribution and density functions of the singly truncated skewed-normal random variables. Properties of the distribution are also investigated, together with a graph of the possible shapes of the density.

2.1. Truncated Skew-Normal Distribution

In what follows we write $\Phi(\cdot)$ and $\phi(\cdot)$ to denote the standard normal distribution function and the standard normal density function, respectively. For probabilistic derivation of the truncated skew-normal distribution, following theorem is useful.

Theorem 1. Let U and V are independent standard normal random variables. If $Z=|U|$, the density function of Z given $V < \theta Z$, for any real θ , is

$$c_{\theta}^{-1} \phi(z)\Phi(\theta z), \quad z > 0, \quad (3)$$

where $c_{\theta}^{-1} = 1/4 + 1/(2\pi) \tan^{-1}\theta$.

Proof. $\Pr(Z \leq t | V < \theta Z) = \Pr(Z \leq t, V \leq \theta Z) / \Pr(V \leq \theta Z) = 2 \int_0^t \phi(z) \Phi(\theta z) dz / \Pr(V \leq \theta Z)$

and (3) by differentiation with respect to t . Since $\Pr(V \leq \theta Z) = 2 \int_0^\infty \int_{-\infty}^{\theta z} \phi(v) \Phi(z) \partial v \partial z = 1/2 + 1/\pi \tan^{-1} \theta$.

Definition 1. A random variable Z is a singly truncated skew-normal random variable with the lower truncation point at 0, written $Z \sim TSN^+(\theta)$, $\theta \in (-\infty, \infty)$, if its probability density function is

$$f(z; \theta) = c_\theta \phi(z) \Phi(\theta z), \quad z > 0. \tag{4}$$

The following properties follow immediately from the definition.

Property 1. The $TSN^+(0)$ density is the half standard normal density, the density of $|U|$, where $U \sim N(0, 1)$.

Property 2. As $\theta \rightarrow \infty$, $f(z; \theta)$ tends to the half standard normal density.

Property 3. If Z is a $TSN^+(\theta)$ random variable, then $Z^* = -Z$ is a $TSN^-(\theta)$ random variable having density

$$h(z^*; \theta) = c_\theta \phi(z^*) \Phi(-\theta z^*), \quad Z^* < 0. \tag{5}$$

Property 4. The density (4) is strongly unimodal, i.e. $\log f(z; \theta)$ is a concave function of z .

Theorem 2. Let U and V be independent standard normal random variables. Then

$$\begin{aligned} Z &= \left[\frac{\theta}{\sqrt{1+\theta^2}} |U| + \frac{1}{\sqrt{1+\theta^2}} V \right]^+ \sim TSN^+(\theta), \\ Z^* &= \left[\frac{-\theta}{\sqrt{1+\theta^2}} |U| + \frac{1}{\sqrt{1+\theta^2}} V \right]^- \sim TSN^-(\theta), \end{aligned} \tag{6}$$

where $[X]^+$ and $[X]^-$ denote singly truncated X random variables with the lower truncation point at zero and the upper truncation point at 0, respectively.

Proof. Letting $\alpha = \theta(1 + \theta^2)^{-1/2}$, $\beta = (1 + \theta^2)^{-1/2}$, we have

$$\begin{aligned} \Pr(Z \leq z) &= E [\Pr(Z \leq z | |U|, Z > 0)] \\ &= 2 \int_0^\infty \Pr(V \leq (z - \alpha u) / \beta) \phi(u) du / \Pr(W > 0) \\ &= 2 \int_0^\infty \Phi((z - \alpha u) / \beta) \phi(u) du / \Pr(V < \theta |U|), \end{aligned}$$

and from the relation $\alpha^2 + \beta^2 = 1$ and $\Pr(V < \theta |U|) = 1/2 + 1/\pi \tan^{-1} \theta$, given in the

proof of Theorem 1, it easily follows that

$$\begin{aligned} \frac{d}{dz} \Pr(Z \leq z) &= \frac{2\phi(z) \int_0^\infty (2\pi\beta^2)^{-1/2} \exp\{-(u-\alpha z)^2/(2\beta^2)\} du}{(1/2 + 1/\pi \tan^{-1}\theta)} \\ &= c_\theta \phi(z) \left\{ 1 - \Phi\left(-\frac{\alpha}{\beta} z\right) \right\} = c_\theta \phi(z) \Phi(\theta z). \end{aligned}$$

Similar proof holds for the derivation of the density (5) of the random variable Z^* .

Theorem 2 provides an acceptance-rejection technique which generates a random variable Z with the density (3) that is the following one. Sample U and V from independent $N(0,1)$ distribution. If $\alpha|U| + \beta V > 0$, then put $Z = \alpha|U| + \beta V$, otherwise restart sampling a new pair of variables U and V , until the inequality $\alpha|U| + \beta V > 0$ is satisfied (α, β figuring in the proof of Theorem 2). The same technique applies to generate a random variable Z^* with the density (5).

Corollary 1. Let (Y_1, Y_2) be a bivariate normal random variable with standardized marginals and correlation ρ , and let $\theta(\rho) = \rho/\sqrt{1-\rho^2}$. Then the following conditional distributions are obtained:

$$Y_1 \mid (Y_i > 0, i=1,2) \sim TSN^+(\theta(\rho)) \text{ and } Y_1 \mid (Y_i < 0, i=1,2) \sim TSN^-(\theta(\rho)).$$

Proof. Let $V = (Y_2 - \rho Y_1)/\sqrt{1-\rho^2}$, then V and Y_1 are independent standard normal variables. Further one can express Y_2 as $Y_2 = (\theta(\rho)Y_1 - U)/\sqrt{1+\theta(\rho)^2}$, where $U \sim N(0,1)$ is independent of Y_1 . Using the distributions, we see that Theorem 1 gives

$$Y_1^A \mid U < \theta(\rho)Y_1^A = Y_1^A \mid Y_2 > \rho(Y_1 - Y_1^A) = Y_1 \mid (Y_1 > 0, Y_2 > 0) \sim TSN^+(\theta(\rho)),$$

where $Y_1^A = |Y_1|$. Now, Property 3 gives that $-Y_1^A \mid V < \theta(\rho)Y_1^A \sim TSN^-(\theta(\rho))$.

Since $-Y_1^A \mid V < \theta(\rho)Y_1^A = -Y_1^A \mid Y_2 < \rho(Y_1^A + Y_1) = Y_1 \mid (Y_1 < 0, Y_2 < 0)$, we have the results.

When a class of bivariate distributions parameterized by $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ are considered, using the standardized Y_1 and Y_2 variables we define $X_i = \sigma_i Y_i + \mu_i$, $i=1, 2$. It is clear that the conditional distributions obtained from Corollary 1 depend only on the population mean μ_1 , the standard deviation σ_1 of X_1 and ρ . This implies that Corollary 1 enable us to make inference for untruncated marginal distribution of X_1 when X_1 and X_2 variables are truncated in each form of the condition in Corollary 1, but only

truncated X_1 values are observed.

Let denote by $F(z ; \Theta)$ the distribution function of (4), i.e.

$$\begin{aligned}
 F(z ; \Theta) &= c_{\Theta} \int_0^z \int_{-\infty}^t \phi(t)\phi(u)\partial u\partial t \\
 &= c_{\Theta} \left[\int_0^{\infty} \int_{-\infty}^t \phi(t)\phi(u)\partial u\partial t - \int_z^{\infty} \int_{-\infty}^t \phi(t)\phi(u)\partial u\partial t \right].
 \end{aligned}$$

Using the properties of the function $T(h, a)$, we have

$$F(z ; \Theta) = c_{\Theta} [\Phi(z) - 2T(z, \Theta) - 1/2 + \pi^{-1}\tan^{-1}\Theta]/2, \quad z > 0,$$

where $T(h, a)$ is the function studied by Owen(1956) which gives the integral of the standard normal bivariate density over region bounded by lines $x=h, y=0$, and $y=ax$ in the (x, y) plane. A computer routine which evaluates $T(h, a)$ has been given by Young and Minder (1974). It is known that $T(h, a)$ is a decreasing function of h and $T(-h, a) = T(h, a)$, $T(h, -a) = -T(h, a)$, $T(h, -a) = -T(h, a)$, $2T(h, 1) = \Phi(h)\Phi(-h)$, and $2T(0, a) = \pi^{-1}\tan^{-1}a$. From the Property 3 and the properties of the function $T(h, a)$, we get the following corollaries.

Property 5. If $H(z^*; \Theta)$ is the distribution function of $Z^* \sim TSM(-\Theta)$ having density (5), then $F(z, \Theta) = 1 - H(z^*; \Theta)$, where $z^* = -z$.

Property 6. $F(z, 1) = \frac{4}{3} \Phi(z)^2$.

2.2. Moments

For computing the moment generating function of $Z \sim TSN^+(\Theta)$, we use the next result.

Lemma 1. Let Y_1 and Y_2 be bivariate normal, with zero means, unit variances and correlation coefficient ρ . Then the bivariate normal orthant probability

$$L(a, b ; \rho) = \Pr(Y_1 > a, Y_2 > b) = \Pr(U < (\rho Y_2 - a)/\sqrt{1 - \rho^2}, Y_2 > b), \quad (7)$$

where U is $N(0, 1)$ and independent of Y_2 .

Proof. Let $U = (Y_1 - \rho Y_2)/\sqrt{1 - \rho^2}$. Then U is $N(0, 1)$ and independent of Y_2 for the covariance of U and Y_2 is zero. Thus, by symmetry of the distribution of U ,

$$L(a, b ; \rho) = \Pr(U > (a - \rho Y_2)/\sqrt{1 - \rho^2}, Y_2 > b) = \Pr(U < (\rho Y_2 - a)/\sqrt{1 - \rho^2}, Y_2 > b).$$

Various methods for computation of the bi-variate normal integral $L(a, b; \rho)$ are suggested by Sowden and Ashford (1969) and Joe (1995), among others. An extensive set of tables were published by National Bureau of Standard in 1959 to give $L(a, b; \rho)$ for $a, b = 0(0.1)4.0$ to six decimal places for $\rho = 0(0.05)0.95(0.01)1$ and to seven decimal places for $\rho = 0(0.05)0.95(0.01)1$.

Theorem 3. The moment generating function of $Z \sim TSN^+(\theta)$ is

$$M_Z(t) = c_\theta e^{t^2/2} \{ (\Phi(a) + \Phi(b))/2 - T(a, 1/\theta) \}, \quad (8)$$

where $a = \theta t / \sqrt{1 + \theta^2}$ and $b = t$.

Proof. $E e^{Zt} = c_\theta \int_0^\infty e^{zt} \phi(z) \Phi(\theta z) dz = c_\theta e^{t^2/2} \int_{-t}^\infty \phi(z) \Phi(\theta z + \theta t) dz.$

Putting $\theta z + \theta t = (\rho z - a) / \sqrt{1 - \rho^2}$, we see, from Lemma 1, that

$$E e^{Zt} = c_\theta e^{t^2/2} L\left(-\frac{\theta t}{\sqrt{1 + \theta^2}}, -t; \frac{\theta}{\sqrt{1 + \theta^2}}\right).$$

Using the relation between $L(\cdot)$, $T(\cdot)$, and $\Phi(\cdot)$ functions (see, for example, Sowden and Ashford (1967)) :

$$L(-a, -b; \rho) = L(a, b; \rho) + \Phi(a) + \Phi(b) - 1$$

$$L(a, b; \rho) = 1 - 1/2 [\Phi(a) + \Phi(b)] - T(a, c_1) - T(b, c_2),$$

where $c_1 = (b - \rho a) / (a \sqrt{1 - \rho^2})$ and $c_2 = (a - \rho b) / (b \sqrt{1 - \rho^2})$, we have the result.

Hence, after some algebra, we obtain

$$E[Z] = \frac{c_\theta}{2\sqrt{2\pi}} \left(1 + \frac{\theta}{1 + \theta^2}\right) \quad \text{and} \quad \text{Var}(Z) = \frac{c_\theta}{2} \left(1 + \frac{\theta}{\pi(1 + \theta^2)}\right) - E[Z]^2. \quad (9)$$

2.3. An Extended Class of Densities

The class of truncated skew-normal distribution $TSN_\theta^+ = \{TSN^+(\theta) : \theta \in (-\infty, \infty)\}$ has been introduced and studied its basic properties analytically. One can extend the class TSN_θ^+ by introducing an additional shape parameter ζ .

Theorem 4. Let U and V are independent $N(0, 1)$ random variables. If $Z = |U|$, the density function of Z conditionally on $V < \theta Z + \zeta$, for any real θ and ζ , is

$$c_{\theta, \zeta} \phi(z)\Phi(\theta z + \zeta), \quad z > 0, \tag{10}$$

where $c_{\theta, \zeta}^{-1} = \{ \Phi(\zeta/\sqrt{1+\theta^2}) + 2T(\zeta/\sqrt{1+\theta^2}, \theta) \} / 2$.

Proof. $\Pr(Z \leq t \mid V < \theta Z + \zeta) = 2 \int_0^t \phi(z)\Phi(\theta z + \zeta) dz / \Pr(V \leq \theta Z + \zeta)$

and (10) by differentiation with respect to t . The normalizing constant is obtained from $\Pr(V \leq \theta Z + \zeta) = \Pr\{ (V - \theta |U|) / \sqrt{1 + \theta^2} \leq \zeta / \sqrt{1 + \theta^2} \} = \Pr(W \leq \zeta / \sqrt{1 + \theta^2})$, where $W = (V - \theta |U|) / \sqrt{1 + \theta^2} \sim SN(-\theta)$. The distribution function of W given by Azzalini (1985) yields $\Pr(W \leq \zeta / \sqrt{1 + \theta^2}) = \Phi(\zeta / \sqrt{1 + \theta^2}) + 2T(\zeta / \sqrt{1 + \theta^2}, \theta)$. This gives $c_{\theta, \zeta}^{-1} = \{ \Phi(\zeta / \sqrt{1 + \theta^2}) + 2T(\zeta / \sqrt{1 + \theta^2}, \theta) \} / 2$.

This broader class of distributions defined by the density (10) will for brevity be written $TSN_{\theta, \zeta}^+ = \{ TSN^+(\theta, \zeta) : -\infty < \theta < \infty, -\infty < \zeta < \infty \}$. Note that, for $\zeta = 0$, $TSN^+(\theta, \zeta)$ is equivalent to $TSN^+(\theta)$. Figure 1 shows the shapes of (3) and (10) for various values of θ and ζ . An acceptance-rejection technique which generates a random variable Z with the density (10) that is immediate from Theorem 4: Sample U and V from independent $N(0,1)$ distribution. If $V < \theta |U| + \zeta$, then put $Z = |U|$, otherwise restart sampling a new pair of variables U and V , until the inequality $V < \theta |U| + \zeta$ is satisfied.

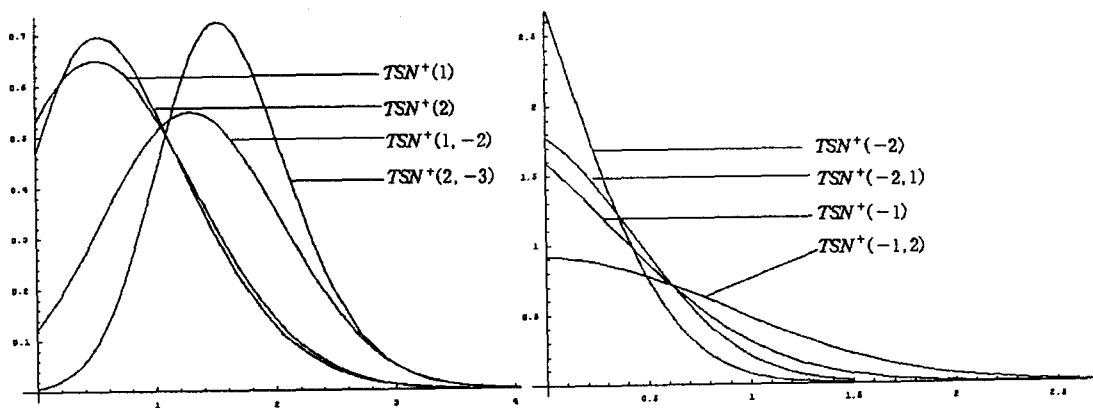


Figure 1. The density functions of $Z \sim TSN^+(\theta)$ and $Z \sim TSN^+(\theta, \zeta)$.

3. Application to A Constrained Regression

In practice, one will often work with the family of distributions generated by the transformation $X = \lambda_1 + \lambda_2 Z$, where $\lambda_2 > 0$ and $Z \sim TSN^+(\Theta, \zeta)$. The density of the random variable X , written $X \sim TSN^+(\Theta, \lambda_1, \lambda_2, \zeta)$ is

$$f(x; \Theta, \lambda_1, \lambda_2, \zeta) = c_{\Theta, \zeta} \lambda_2^{-1} \phi((x - \lambda_1)/\lambda_2) \Phi(\Theta(x - \lambda_1)/\lambda_2 + \zeta), \quad x > \lambda_1. \quad (11)$$

A well-known property of the normal distribution is that, if $Y \sim N(X, \sigma^2)$ where X is a normal random variable, the posterior distribution of X is still normal. An analogous fact is true if a prior X has the probability density function (11). Some simple algebra shows that the posterior density function of X given that $Y = y$ is still of type (11) with $(\Theta, \lambda_1, \lambda_2, \zeta)$ replaced by

$$\Theta(1 + \lambda_2^2/\sigma^2)^{-1/2}, \quad \frac{y/\sigma^2 + \lambda_1/\lambda_2^2}{1/\sigma^2 + 1/\lambda_2^2}, \quad (1/\sigma^2 + 1/\lambda_2^2)^{-1/2}, \quad \zeta + (y - \lambda_1) \frac{\Theta \lambda_2}{\sigma^2 + \lambda_2^2}.$$

Note that the parameter Θ shrinks towards 0, independently of y , and that the updating formulas of the parameters λ_1 and λ_2 are the same as the normal prior case.

Consider now a constrained linear regression setting

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + e_i, \quad i = 1, \dots, n,$$

where $\beta_1 > 0$ and $\beta_2 > 0$ and e_i 's are iid $N(0, \tau^2)$ with known τ^2 . The constraint of positive signs, $\beta_1 > 0$ and $\beta_2 > 0$, in regression coefficients frequently arises in econometric work. For example, an economic theory says that, as explanatory variables, the disposable income and the rate of housing price change have the same positive effect on the dependent variable, the housing price.

With a bivariate normal prior distribution on $\beta = (\beta_1, \beta_2)'$, $\beta \sim N_2(0, \Sigma)$, and the prior knowledge $\beta_1 > 0$ and $\beta_2 > 0$, we obtain constrained prior distributions of β_1/σ_1 and β_2/σ_2 from using Corollary 1:

$$\beta_1/\sigma_1 \mid \beta_1 > 0, \beta_2 > 0 \sim TSN^+(\Theta(\rho)) \text{ and } \beta_2/\sigma_2 \mid \beta_1 > 0, \beta_2 > 0 \sim TSN^+(\Theta(\rho)),$$

where $\Theta(\rho) = \rho/\sqrt{1 - \rho^2}$ and $\Sigma = \{\sigma_{jk}\}$ are known with $\sigma_{jk} = \sigma_j^2$ for $j = k$; $\sigma_{jk} = \sigma_j \sigma_k \rho$ for $j \neq k$. The joint posterior becomes

$$p(\beta \mid \beta_1 > 0, \beta_2 > 0, \text{data}) \propto e^{-[\sum_{i=1}^n (y_i - \beta_1 x_{1i} - \beta_2 x_{2i})^2 / (2\tau^2) + \sum_{j=1}^2 \beta_j^2 / (2\sigma_j^2)]} \prod_{j=1}^2 \Phi(\Theta(\rho) \beta_j / \sigma_j),$$

where $\text{data} = \{y_{[i]}, x_{1i}, x_{2i}; i = 1, \dots, n\}$. Marginal densities of β_1 and β_2 are complicate. Instead simple algebra gives the Gibbs sampler that can be used for inference of the regression model. To apply the Gibbs sampler, we need the conditional posterior distributions which, from above conjugacy property of the $TSN^+(\Theta, \lambda_1, \lambda_2, \zeta)$ density (11) with the

normal distribution, are

$$\beta_1 \mid (\beta_2, \beta_1 > 0, \beta_2 > 0, \text{data}) \sim \text{TSN}^+(\theta^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)}, \zeta^{(1)})$$

$$\beta_2 \mid (\beta_1, \beta_1 > 0, \beta_2 > 0, \text{data}) \sim \text{TSN}^+(\theta^{(2)}, \lambda_1^{(2)}, \lambda_2^{(2)}, \zeta^{(2)}),$$

where

$$\lambda_1^{(1)} = \frac{\sum_{i=1}^n x_{1i}(y_i - \beta_2 x_{2i})/\tau^2}{\sum_{i=1}^n x_{1i}^2/\tau^2 + 1/\sigma_1^2}, \quad \lambda_1^{(2)} = \frac{\sum_{i=1}^n x_{2i}(y_i - \beta_1 x_{1i})/\tau^2}{\sum_{i=1}^n x_{2i}^2/\tau^2 + 1/\sigma_2^2},$$

$$\lambda_2^{(1)} = (\sum_{i=1}^n x_{1i}^2/\tau^2 + 1/\sigma_1^2)^{-1/2}, \quad \lambda_2^{(2)} = (\sum_{i=1}^n x_{2i}^2/\tau^2 + 1/\sigma_2^2)^{-1/2},$$

$$\theta^{(1)} = \theta(\rho)(1 + \sigma_1^2 \sum_{i=1}^n x_{1i}^2/\tau^2), \quad \theta^{(2)} = \theta(\rho)(1 + \sigma_2^2 \sum_{i=1}^n x_{2i}^2/\tau^2),$$

$$\zeta^{(1)} = \frac{\sigma_1 \sum_{i=1}^n x_{1i}(y_i - \beta_2 x_{2i})}{\sigma_1^2 \sum_{i=1}^n x_{1i}^2 + \tau^2}, \quad \zeta^{(2)} = \frac{\sigma_2 \sum_{i=1}^n x_{2i}(y_i - \beta_1 x_{1i})}{\sigma_2^2 \sum_{i=1}^n x_{2i}^2 + \tau^2}.$$

The Gibbs sampler proceeds by alternatively sampling from these two TSN^+ distributions using the acceptance-rejection sampling scheme described in Subsection 2.3.

4. Concluding Remarks

We proposed a singly truncated parametric family of distributions that strictly includes a truncated normal distribution, and studies its properties, with special emphasis on the relation with bivariate normal distribution. In particular, we show that it defines conditional distributions of a truncated bivariate normal distribution. This enables us inference for the marginal distribution of a bivariate normal random variables when they are singly truncated (as the example given in section 4). The likelihood equations for a simple random sample of size n from Z , defined by Definition 1, are readily written down. To compute MLE of Θ , we can use either Newton-Raphson method or the Nelder-Mead algorithm. Because these procedures are quite standard, we omit the details for brevity.

The family of distributions is potentially relevant for practical applications, since there are far fewer distributions available for dealing with truncated data than in untruncated case. Therefore, a further study pertaining to developing applications of the family of distributions in real data analyses is needed, and is left as a future study of interest.

References

- [1] Arnold, B. C., Beaver, R. J., Groeneveld, R. A. and Meeker, W. Q. (1993). The nontruncated marginal of a truncated bivariate normal distribution, *Psychometrika*, 58, 471-488.
- [2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.
- [3] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83, 715-726.
- [4] Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society B*, 61, 579-602.
- [5] Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew elliptical distributions. *Journal of Multivariate Analysis*, 79, 99-113.
- [6] Cohen, A. C. (1991). *Truncated and Censored Samples: Theory and Applications*, Marcel Dekker, New York.
- [7] Henze, N. (1986). A probabilistic representation of the 'skew-normal' distribution. *Scandinavian Journal of Statistics*, 13, 271-275.
- [8] Joe, H. (1995). Approximation to multivariate normal rectangle probabilities based on conditional expectations, *Journal of the American Statistical Association*, 90, 957-966.
- [9] Kim, H. J. (2002). Binary Regression with a class of skewed t link models, *Communications in Statistics*, A 31, 1863-1886.
- [10] Owen, D. B. (1956). Tables for computing bi-variate normal probabilities, *Annals of Mathematical Statistics*, 27, 1075-1090.
- [11] Sowden, R. R. and Ashford, J. R. (1967). Computation of the bi-variate normal integral, *Journal of the Royal Statistical Society*, B 18, 169-180.
- [12] Young, J. C. and Minder, Ch. E. (1974). Algorithm AS76: an integral useful in computing non-central t and bivariate normal probabilities. *Applied Statistics*, 23, 455-457.

[Received February 2004, Accepted March 2004]