Journal of the Korean Statistical Society (2006), **35**: *4, pp 419–439*

BAYESIAN ROBUST ANALYSIS FOR NON-NORMAL DATA BASED ON A PERTURBED-t MODEL[†]

HEA-JUNG KIM[1]

ABSTRACT

The article develops a new class of distributions by introducing a non-negative perturbing function to t_ν distribution having location and scale parameters. The class is obtained by using transformations and conditioning. The class strictly includes t_ν and skew-t_ν distributions. It provides yet other models useful for selection modeling and robustness analysis. Analytic forms of the densities are obtained and distributional properties are studied. These developments are followed by an easy method for estimating the distribution by using Markov chain Monte Carlo. It is shown that the method is straightforward to specify distributionally and to implement computationally, with output readily adopted for constructing required criterion. The method is illustrated by using a simulation study.

AMS 2000 subject classifications. Primary 62H10; Secondary 62F15.
Keywords. Perturbed t-distribution, non-normal data, Bayesian robust analysis.

1. INTRODUCTION

Suppose that the model where a random variable Z is distributed with density $g(z|\delta)$ and that it is desired to make inferences about δ, where δ is a q-dimensional vector of unknown parameters. The usual statistical analysis assumes that a random sample $Z_1, \ldots, Z_n$ from $g(z|\delta)$ can be observed. There are many situations, however, in which such a random sample might not be available, for instance, if it is too difficult or too costly to obtain. Then statistical models have to be developed to incorporate the non-randomness or bias in the observations. Weighted distributions (Rao, 1985) arise when the density of the potential observation z gets distorted so that it is multiplied by some non-negative weight function

Received February 2006; accepted October 2006.
[†]This research was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) KRF-2005-041-C00089.
[1]Department of Statistics, Dongguk University, Seoul 100-715, Korea (e-mail: kim3hj@ dongguk.edu)

$h(z|\delta,\gamma)$, which involves some unknown parameters denoted by a vector γ. Thus, the observed data is a random sample from the following weighted version of $g(z|\delta)$.

$$f(z|\delta,\gamma) = g(z|\delta)\frac{h(z|\delta,\gamma)}{E_\delta[h(Z|\delta,\gamma)]}, \tag{1.1}$$

where the expectation is taken with respect to the distribution $g(z|\delta)$.

A particular class of the weighted distributions (1.1) is the "selection model" in which observations are obtained only from a selected portion of the population of interest. For example, the observation Z of a characteristic of a certain population is measured only for individuals who manifest a certain disease due to cost or ethical reasons (Bayarri and DeGroot, 1992). For such problems, the goal is to find estimator of δ in the presence of the nuisance weight function h.

A different point of view is given by a robustness argument. If $g(z|\delta)$ is the central model of interest, then the weight function h in (1.1) can be seen as a perturbing (or perturbation of $g(z|\delta)$) function. For instance, if g is an elliptical probability density function, then h generates asymmetric outliers in the observed sample from f. The goal is then to derive robust estimator of δ, in the presence of a certain class of the nuisance weight function h. Under the robust argument, the class of weighted distributions includes a vast set of skew distributions, such as skew-normal, skew-t, skew-elliptical and generalized skew-elliptical distributions. Systematic treatments of these distributions have been given by Azzalini (1985), Branco and Dey (2001) and Ma *et al.* (2005), among others. They call the weight function h as "skew function" in a sense that it is useful to modeling random phenomena which have heavier tails than the normal as well as having some skewness.

The major goal of current paper is to introduce yet other subclass of weighted distributions, *i.e.*, a class of perturbed-t_ν distributions, that is useful for selection modeling and robustness analysis. It also suggests a Markov chain Monte Carlo (MCMC) method for estimating the location and scale parameters, δ, of the central part of the model as well as the parameters, γ, in the perturbed function. The interest in studying the class of perturbed-t_ν distributions, strictly including skew-t_ν and t_ν distributions, comes from both theoretical and applied directions. On the theoretical side, it provides a class of distributions that enjoys a number of formal properties which resemble those of skew-t_ν distribution introduced by Branco and Dey (2001) and Kim (2002). From the applied view point, the class is a set of unimodal empirical distributions with presence of skewness and possibly heavy (or light) tail. This implies that the class is useful to modeling random phenomena which have heavier (or lighter) tails than the normal as well as having some skewness. Moreover, the class provides yet other selection models that enable us to analyze a screened data in terms of the sum of truncated and untruncated observations.

The article organized as follows. Section 2 develops a class of perturbed elliptical distributions that is useful for selection modeling and robustness analysis. Section 3 considers the particular case of t_ν distribution. In particular probabilistic and conditional representations of the perturbed-t_ν distribution and its properties are given. In Section 4, some moments of the distribution are derived. In Section 5, we develop an easy implementation technique, MCMC method, for estimating the distribution. In Section 6 a numerical example validating the MCMC method is given. We give few summary remarks in Section 7.

2. The Class of Perturbed Elliptical Distributions

A distribution of $k \times 1$ random vector $\mathbf{X}$, written $\mathbf{X} \sim EC_k(\theta, \Sigma, g^{(k)})$, is said to have k-variate elliptically symmetric (or simply elliptical) distribution with location vector $\theta \in \mathbb{R}^k$ and a $k \times k$ (positive definite) dispersion matrix Σ and the density generator function $g^{(k)}$. The density of $\mathbf{X}$ distribution is given by

$$f(\mathbf{x}|\theta, \Sigma) = |\Sigma|^{-1/2} g^{(k)}((\mathbf{x}-\theta)'\Sigma^{-1}(\mathbf{x}-\theta)), \tag{2.1}$$

for some density generator function $g^{(k)}(u)$, $u \geq 0$, such that

$$\int_0^\infty u^{(k/2)-1} g^{(k)}(u)du = \Gamma(k/2)/\pi^{k/2}. \tag{2.2}$$

This implies that $g^{(k)}$ is a spherical k-dimensional density. When the density function of the elliptical distribution does not exist, we use the characteristic function ϕ and replace the density generator function $g^{(k)}$ in the notation and use $\mathbf{X} \sim EC_k(\theta, \Sigma, \phi)$. By varying the function $g^{(k)}$, distributions with longer or shorter tails than the multivariate normal can be obtained. A comprehensive review of the properties and characterizations of multivariate elliptical distributions can be found in Fang *et al.* (1990) and Fang and Zhang (1990). If we set $k = 2$ and $\Psi = \{\psi_{ij}\}$ with $\psi_{11} = \psi_{22} = 1$ and $\psi_{12} = \psi_{21} = \rho$, the properties of $EC_2(\mathbf{0}, \Psi, g^{(2)})$ distribution yield following theorems. From now on, we will use Ψ to denote this standard form of 2×2 dispersion matrix.

Lemma 2.1. *Let* $\mathbf{W} \sim EC_2(\mathbf{0}, \Psi, g^{(2)})$, $\mathbf{W} = (W_1, W_2)'$. *Then the conditional distribution of* W_2 *given that* $W_1 = w$ *is* $EC_1(\alpha, \beta, g_{q(w)})$ *distribution.*

Here $\alpha = \rho w$, $\beta = 1 - \rho^2$, $g_{q(w)} = g^{(2)}(u + q(w))/g^{(1)}(q(w))$, *where* $g^{(1)}(u) = 2\int_0^\infty g^{(2)}(r^2 + u)dr$ *and* $q(w) = w^2$.

PROOF. Straightforward application of the result of Branco and Dey (2001) yields the result. □

LEMMA 2.2. *Let* $\mathbf{W} \sim EC_2(\mathbf{0}, \Psi, g^{(2)})$, $\mathbf{W} = (w_1, w_2)'$. *If* z *is set to equal to* W_1 *conditionally on* $a < W_2 < b$. *Then the pdf of* z *is*

$$f_Z(z) = \frac{f_{g^{(1)}}(z)\left\{F_{g_{q(z)}}(\lambda_1 b - \lambda z) - F_{g_{q(z)}}(\lambda_1 a - \lambda z)\right\}}{F_{g^{(1)}}(b) - F_{g^{(1)}}(a)} \quad \text{for } z \in \mathbb{R}, \tag{2.3}$$

where $\lambda = \rho/\sqrt{1-\rho^2}$, $\lambda_1 = (1+\lambda^2)^{1/2} = 1/\sqrt{1-\rho^2}$, *and* $f_{g^{(1)}}(\cdot)$ *and* $F_{g^{(1)}}(\cdot)$ *are the pdf and the cdf of* $EC_1(0, 1, g^{(1)})$, *respectively.* $F_{g_{q(z)}}$ *is the cdf of* $EC_1(0, 1, g_{q(z)})$ *with* $q(z) = z^2$.

PROOF. The *pdf* of Z can be expressed as

$$f_Z(z) = P(a < W_2 < b|z) f_{W_1}(z)/P(a < W_2 < b).$$

Using the property of $EC_2(\mathbf{0}, \Psi, g^{(2)})$ (Fang *et al.*, 1990, p. 43) we obtain the marginal distribution, $W_1 \sim EC_1(0, 1, g^{(1)})$. Further Lemma 2.1 gives $W_2 \,|W_1 = z \sim EC_1(\alpha, \beta, g_{q(z)})$, and hence

$$\begin{aligned} P(a < W_2 < b\,|z) &= P\Big((a-\alpha)/\sqrt{\beta} < (W_2 - \alpha)/\sqrt{\beta} < (b-\alpha)/\sqrt{\beta}\,|z\Big) \\ &= F_{g_{q(z)}}(\lambda_1 b - \lambda z) - F_{g_{q(z)}}(\lambda_1 a - \lambda z). \end{aligned}$$

Noticing that $f_{W_1}(z) = f_{g^{(1)}}(z)$ and $P(a < W_2 < b) = F_{g^{(1)}}(b) - F_{g^{(1)}}(a)$, we have the result. □

Noticing, from the proof of Lemma 2.2, that

$$E\left[F_{g_{q(Z)}}(\lambda_1 b - \lambda Z) - F_{g_{q(Z)}}(\lambda_1 a - \lambda Z)\right] = F_{g^{(1)}}(b) - F_{g^{(1)}}(a) \;\text{ for }\; z \in \mathbb{R},$$

the density (2.3) is a proper type of the weighted distribution in (1.1). A generalization of Lemma 2.2 gives following theorem.

THEOREM 2.1. *Let* $\mathbf{X} = (X, Y)'$, *and let* $\mathbf{X} \sim EC_2(\theta, \Sigma, g^{(2)})$ *with* $\theta = (\theta_1, \theta_2)'$, $\Sigma = \{\sigma_{ij}\}$, $\sigma_{ii} = \sigma_i^2$ *and* $\sigma_{12} = \rho\sigma_1\sigma_2$. *Then, the pdf of* $Z \equiv [X \,|a < Y < b]$ *variable is*

$$f_Z^*(z) = \frac{f_{g^{(1)}}(u_1(z))\Big\{F_{g_{q*(z)}}(\lambda_1 u(b) - \lambda u_1(z)) - F_{g_{q*(z)}}(\lambda_1 u(a) - \lambda u_1(z))\Big\}}{\sigma_1\Big\{F_{g_{(1)}}(u(b)) - F_{g_{(1)}}(u(a))\Big\}} \tag{2.4}$$

for $z \in \mathbb{R}$, *where* $u_1(z) = (z - \theta_1)/\sigma_1$, $u(a) = (a - \theta_2)/\sigma_2$, $u(b) = (b - \theta_2)/\sigma_2$, $F_{g_{q*(z)}}$ *is the cdf of* $EC_1(0, 1, g_{q_*(z)})$ *with* $q_*(z) = u_1(z)^2$.

PROOF. Given the density (2.3), the density is obtained through the transformation relations $X = \sigma_1\, W_1 + \theta_1$ and $Y = \sigma_2\, W_2 + \theta_2$. The density of a random variable Z with a perturbed symmetric distribution is defined through an elliptical function and a perturbing function as follows. □

DEFINITION 2.1. *If a random variable* Z *has density function (2.4), then we say that* Z *is a perturbed symmetric random variable with parameters* θ, Σ *and perturbing function*

$$\frac{\Big\{F_{g_{q*(z)}}(\lambda_1 u(b) - \lambda u_1(z)) - F_{g_{q*(z)}}(\lambda_1 u(a) - \lambda u_1(z))\Big\}}{\Big\{F_{g_{(1)}}(u(b)) - F_{g_{(1)}}(u(a))\Big\}}.$$

For brevity we shall also say that Z *is* $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ *random variable.*

It is easily seen that the class of $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ distributions strictly includes skew-normal, skew-elliptical, generalized skew-elliptical, and elliptically symmetric distributions as special members:

(i) When $a = \theta_2$, $b = \infty$, and $g^{(2)}$ is bivariate normal density generator function, $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ reduces to the skew-normal distribution introduced by Azzalini (1985).

(ii) When $a = \theta_2$ and $b = \infty$, $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ distribution is the skew-elliptical distribution by Branco and Dey (2001).

(iii) When $a = \theta_2$, $b = \infty$ and $\rho/\sqrt{1-\rho^2} = 1$, $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ distribution becomes the skew-elliptical distribution by Branco and Dey (2001).

(iv) When $\rho = 0$, $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ distribution is the elliptically symmetric distribution. In other words, the suggested class of perturbed symmetric distributions contains many families of weighted distributions as well as standard elliptically symmetric families. Therefore, the proposed class of distributions is expected to extend earlier results on selection modeling and robustness analysis for the central model of interest.

In this paper, we restrict our attention to the situation where $g^{(2)}$ is bivariate t_ν density generator function, in order to accommodate the application of mixed normal distribution theory in the robustness analysis and the heavy tailed random phenomena with some skewness.

3. PERTURBED-t DISTRIBUTIONS

From Lemma 2.1 and Theorem 2.1, we can get an alternative and convenient expression for the *pdf* of $PS_{(a,b)}(\theta, \Sigma, g^{(2)})$ distribution as

$$f_Z^*(z) = \frac{\int_{u(a)}^{u(b)} g^{(2)}\left(\frac{(r-\rho u_1(z))^2}{1-\rho^2} + u_1(z)^2\right) dr}{\sigma_1\sqrt{1-\rho^2}\left\{F_{g_{(1)}}(u(b)) - F_{g_{(1)}}(u(a))\right\}}, \qquad z \in \mathbb{R}. \tag{3.1}$$

The generator function for a bivariate t_ν-distribution with the degrees of freedom ν is

$$g_\nu^{(2)}(u) = C_1[\nu + u]^{-(\nu+2)/2}, \tag{3.2}$$

where $C_1 = \nu^{\nu/2}\Gamma((\nu+2)/2)/\{\pi\ \Gamma(\nu/2)\}$. It follows from (3.1) that the *pdf* $f_Z^*(z)$ obtained from using the generator function (3.2) is

$$\begin{aligned}
&\frac{C_1\{F_\nu(u(b)) - F_\nu(u(a))\}^{-1}}{\sigma_1\sqrt{1-\rho^2}} \int_{u(a)}^{u(b)} \left[\nu + \frac{\{r - \rho u_1(z)\}^2}{(1-\rho^2)} + u_1(z)^2\right]^{-(\nu+2)/2} dr \\
&= \frac{C_2\{F_\nu(u(b)) - F_\nu(u(a))\}^{-1}}{\sigma_1\{\nu + u_1(z)^2\}^{(\nu+1)/2}} \int_{v_1(z)}^{v_2(z)} C_3\left(\nu + 1 + t^2\right)^{-(\nu+2)/2} dt \\
&= \frac{f_\nu(u_1(z))\{F_{\nu+1}(v_2(z)) - F_{\nu+1}(v_1(z))\}}{\sigma_1\{F_\nu(u(b)) - F_\nu(u(a))\}}, \qquad z \in \mathbb{R},
\end{aligned} \tag{3.3}$$

where $f_\nu(\cdot)$ and $F_\nu(\cdot)$ denote the *pdf* and the *cdf* of a univariate standard t_ν distribution, while $F_{\nu+1}(\cdot)$ is the *cdf* of a univariate standard $t_{\nu+1}$ distribution. Here $C_2 = \nu^{\nu/2}\Gamma((\nu+1)/2)/\{\sqrt{\pi}\ \Gamma(\nu/2)\}$, $C_3 = (\nu+1)^{(\nu+1)/2}\Gamma((\nu+2)/2)/\{\sqrt{\pi}\ \Gamma((\nu+1)/2)\})$, $v_1(z) = (\lambda_1 u(a) - \lambda u_1(z))\sqrt{\nu+1}/\sqrt{\nu + u_1(z)^2}$ and $v_2(z) = (\lambda_1 u(b) - \lambda u_1(z))\sqrt{\nu+1}/\sqrt{\nu + u_1(z)^2}$. We see that the density (3.3) of a random variable with a perturbed-t distribution is defined through a generalized t-density and a perturbing function as follows.

DEFINITION 3.1. *A perturbed-t_ν distribution, written by $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$, is a distribution whose probability density function is of the form*

$$f_Z^*(z) = \frac{f_\nu(u_1(z))}{\sigma_1} \frac{[F_{\nu+1}(v_2(z)) - F_{\nu+1}(v_1(z))]}{[F_\nu(u(b)) - F_\nu(u(a))]}, \quad z \in \mathbb{R}, \tag{3.4}$$

where notations in (3.4) are the same as those in (3.3) and we refer to $[F_{\nu+1}(v_2(z)) - F_{\nu+1}(v_1(z))]/[F_\nu(u(b)) - F_\nu(u(a))]$ as the perturbing function.

Now we will state some interesting properties for the $\mathcal{PT}_{(u(a),u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution as well as the associate examples. Theorem 2.1 immediately gives the following property.

PROPERTY 3.1. *Let $X = (X, Y)'$ be a bivariate t_ν variable with a location vector $\theta = (\theta_1, \theta_2)'$ and a scale matrix $\Sigma = \{\sigma_{ij}\}$, that is $X \sim t_\nu(\theta, \Sigma)$. Then $Z \equiv [X \mid a < Y < b] \sim \mathcal{PT}_{(u(a),u(b))}(\theta_1, \sigma_1, \rho, \nu)$.*

PROPERTY 3.2. *For $T \sim t_\nu(\theta_1, \sigma_1)$, a generalized univariate t_ν with the mean θ_1 and the scale parameter σ_1,*

$$E[F_{\nu+1}(v_2(T)) - F_{\nu+1}(v_1(T))] = F_\nu(u(b)) - F_\nu(u(a)).$$

PROPERTY 3.3. *The distribution function of $Z \sim \mathcal{PT}_{(u(a),u(b))}(\theta_1, \sigma_1, \rho, \nu)$ is*

$$F_Z^*(z) = \frac{H_\nu(u_1(z), u(b)) - H_\nu(u_1(z), u(a))}{\{F_\nu(u(b)) - F_\nu(u(a))\}}, \quad z \in \mathbb{R}, \tag{3.5}$$

where $H_\nu(z_1, z_2)$ denotes the standard bivariate t distribution function whose formula is given by Dunnett and Sobel (1954).

PROPERTY 3.4. *If $Z \sim \mathcal{PT}_{(u(a),u(b))}(\theta_1, \sigma_1, \rho, \nu)$, then*

$$-Z \sim \mathcal{PT}_{(u(a),u(b))}(-\theta_1, \sigma_1, -\rho, \nu).$$

PROPERTY 3.5. *Let $T = (Z - \theta_1)/\sigma_1$, where $Z \sim \mathcal{PT}_{(u(a),u(b))}(\theta_1, \sigma_1, \rho, \nu)$. Then*

$$T \sim \mathcal{PT}_{(u(a),u(b))}(0, 1, \rho, \nu). \tag{3.6}$$

PROPERTY 3.6. *If* $T \sim \mathcal{PT}_{(u(a),u(b))}(0,1,\rho,\nu)$, *then*

$$-T \sim \mathcal{PT}_{(u(a),\,u(b))}(0,1,-\rho,\nu) \equiv \mathcal{PT}_{(-u(b),\,-u(a))}(0,1,\rho,\nu).$$

Following theorem gives a probabilistic representation of the $\mathcal{PT}_{(u(a),u(b))}(0,1,\rho,\nu)$ distribution in terms of mixed normal and mixed truncated normal laws.

THEOREM 3.1. *Conditional on* $\eta \sim Gamma(\nu/2, 2/\nu)$, *let* $U \sim N(0,\eta^{-1})$ *and* $V \sim N(0,\eta^{-1})$ *be independent normal variables. Then the scale mixed distribution of* T *is*

$$T \equiv \frac{\lambda}{\sqrt{1+\lambda^2}} U_{(u(a),\,u(b))} + \frac{1}{\sqrt{1+\lambda^2}} V \sim \mathcal{PT}_{(u(a),u(b))}(0,1,\rho,\nu), \tag{3.7}$$

where $U_{(u(a),\,u(b))}$ *denotes the truncated* $N(0,\eta^{-1})$ *with upper and lower truncation points are* $u(b)$ *and* $u(a)$, *respectively.*

PROOF. Let $\alpha = u(a)$, $\beta = u(b)$, $a = \lambda/\sqrt{1+\lambda^2}$, $b = 1/\sqrt{1+\lambda^2}$, and $f(\eta)$ be the density of η. Then we have

$$\begin{aligned} P(T \le t) &= E_\eta E[P(T \le t\,|U_{(\alpha,\beta)},\eta)] \\ &= \int_0^\infty \int_\alpha^\beta \eta^{1/2} P(V \le (t-au)/b)\phi(\eta^{1/2}u) f(\eta)\partial u \partial\eta / \{F_\nu(\beta) - F_\nu(\alpha)\} \\ &= \int_0^\infty \int_\alpha^\beta \eta^{1/2} \Phi(\eta^{1/2}(t-au)/b)\phi(\eta^{1/2}u) f(\eta)\partial u \partial\eta / \{F_\nu(\beta) - F_\nu(\alpha)\} \end{aligned}$$

and from the relation $a^2 + b^2 = 1$ it easily follows that $f_T^*(t) = dP(T \le t)/dt$ is

$$\begin{aligned} &\int_0^\infty \int_\alpha^\beta \frac{\eta^{1/2}\phi(\eta^{1/2}t)}{\{F_\nu(\beta)-F_\nu(\alpha)\}} \left(\frac{\eta}{2\pi b^2}\right)^{1/2} \exp\left\{-\frac{\eta(u-at)^2}{2b^2}\right\} f(\eta)\partial u\partial\eta \\ &= \int_0^\infty \frac{\eta^{1/2}\phi(\eta^{1/2}t)}{[F_\nu(\beta)-F_\nu(\alpha)]} \left\{\Phi\left(\frac{\eta^{1/2}(\beta-at)}{b}\right) - \Phi\left(\frac{\eta^{1/2}(\alpha-at)}{b}\right)\right\} f(\eta)d\eta \\ &= \int_0^\infty \int_\alpha^\beta \frac{(\nu/2)^{\nu/2}\eta^{\nu/2}}{2b\pi\Gamma(\nu/2)\{F_\nu(\beta)-F_\nu(\alpha)\}} \exp\left\{-\frac{[\nu+t^2+\{(u-at)/b\}^2]\eta}{2}\right\}\partial\eta\partial u \\ &= \frac{f_\nu(t)}{\{F_\nu(\beta)-F_\nu(\alpha)\}b} \left(\frac{\nu+1}{\nu+t^2}\right)^{1/2} \int_\alpha^\beta f_{\nu+1}\left(\left(\frac{\nu+1}{\nu+t^2}\right)^{1/2}\frac{(u-at)}{b}\right) du, \end{aligned}$$

where f_ν and $f_{\nu+1}$ is standard t_ν and $t_{\nu+1}$ densities, respectively. A variable transformation $u' = \{(\nu+1)/(\nu+t^2)\}^{1/2}(u-at)/b$ and letting $\alpha = u(a)$ and $\beta = u(b)$ to the last equation gives the *pdf* of (3.5) obtained from (3.4). □

It is easily seen that the distribution (3.4) leads to a parametric class of distributions that have strict inclusion of t_ν distribution (for the case $\theta_1 = 0$, $\sigma_1 = 1$ and $\rho = 0$) and a perturbed Cauchy distribution (for $\nu = 1$). Also note that (3.7) reduces to a perturbed normal distribution, written $\mathcal{PN}_{(u(a),\, u(b))}(0, 1, \rho)$, when the distribution of η is degenerate with $\eta = 1$.

Thus the representation in (3.7) reveals the structure of the class of $\mathcal{PT}_{(u(a),u(b))}(0, 1, \rho, \nu)$ distributions and indicates the kind of departure from the standard t_ν distribution. Furthermore, the representation provides one-for-one method of generating a random variable Z with density (3.4). The method is the following one. Sample η from $Gamma(\nu/2, 2/\nu)$. Then sample $U_{(u(a),\, u(b))}$ and V from respective $TN_{(u(a),\, u(b))}(0, \eta^{-1})$ and $N(0, \eta^{-1})$ distributions to generate Z from the equation:

$$Z = \theta_1 + \sigma_1 \left\{ \frac{\lambda}{\sqrt{1+\lambda^2}} U_{(u(a),\, u(b))} + \frac{1}{\sqrt{1+\lambda^2}} V \right\}.$$

Here $TN_{(u(a),\, u(b))}(0, \eta^{-1})$ denotes a truncated normal distribution with upper and lower truncation points $u(b)$ and $u(a)$. A sample of $U_{(u(a),\, u(b))}$ is easily extracted from the truncated distribution by using the one-for-one method documented by Devroye (1986). Figure 3.1 depicts various density shapes of $\mathcal{PT}_{(u(a),u(b))}(0, 1, \rho, \nu)$ distribution. Figure 3.1 notes that the distribution induces not only skewness but also high kurtosis to t_ν distribution. Also note from Property 3.6 that the density of $-Z$ can be obtained from changing the sign of ρ.

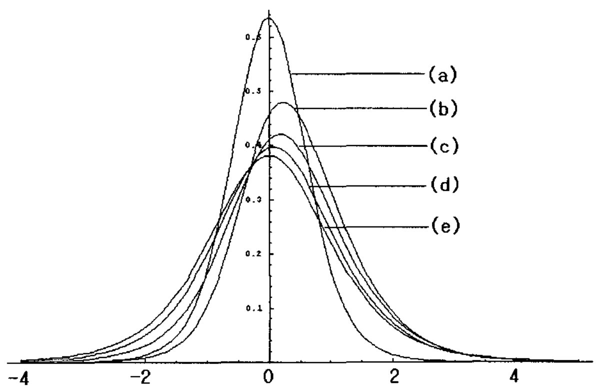

FIGURE 3.1 *Various Shapes of the* $\mathcal{PT}_{(u(a),u(b))}(0, 1, \rho, \nu)$ *densities with* $\nu = 5$*: (a)* $u(a) = -.5$, $u(b) = .5$, *and* $\rho = .8$*; (b)* $u(a) = -.5$, $u(b) = 3$, *and* $\rho = .8$*; (c)* $u(a) = -.5$, $u(b) = 3$, *and* $\rho = .5$*; (d)* $u(a) = -.5$, $u(b) = 3$, *and* $\rho = .2$*; (e) standard student* t_5 *density.*

4. MOMENTS

4.1. Moment Generating Function

In this section we derive the moment generating function (MGF) for the perturbed-t distribution. To compute the moments of $Z \sim \mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$, it suffices to compute the moments of $T \sim \mathcal{PT}_{(\alpha,\, \beta))}(0, 1, \rho, \nu)$ by Property 6.5, where $\alpha = u(a)$ and $\beta = u(b)$. From the proof of Theorem 3.1, we see that T has the density which can be expressed as

$$\int_0^\infty \frac{\eta^{1/2}\phi(\eta^{1/2}t)}{[F_\nu(\beta) - F_\nu(\alpha)]} \left[\Phi\left(\frac{\eta^{1/2}(\beta - at)}{b}\right) - \Phi\left(\frac{\eta^{1/2}(\alpha - at)}{b}\right)\right] f(\eta)d\eta \quad (4.1)$$

for $t \in \mathbb{R}$, where $a = \lambda/\sqrt{1+\lambda^2}$, $b = 1/\sqrt{1+\lambda^2}$, and $f(\eta)$ be the density of $\eta \sim Gamma(\nu/2, 2/\nu)$.

THEOREM 4.1. *Let $T \sim \mathcal{PT}_{(\alpha,\, \beta))}(0, 1, \rho, \nu)$ then its moment generating function is*

$$M_T(\ell) = \frac{E_\eta\left[\{\Phi(\beta_\ell) - \Phi(\alpha_\ell)\}e^{\ell^{\,2}/(2\eta)}\}\right]}{[F_\nu(\beta) - F_\nu(\alpha)]} \quad (4.2)$$

for $\ell \in \mathbb{R}$, where E_η denotes that the expectation is taken with respect to the distribution of $\eta \sim Gamma(\nu/2, 2/\nu)$, $\alpha_\ell = \alpha\eta^{1/2} - a\eta^{-1/2}\ell$ and $\beta_\ell = \beta\eta^{1/2} - a\eta^{-1/2}\ell$.

PROOF. Considering the *pdf* (4.1) and $\ell \in \mathbb{R}$, we have the mgf written as

$$\int_0^\infty \int_{-\infty}^\infty \frac{\eta^{1/2}e^{\ell^{\,2}/(2\eta)}e^{-\eta(t-\ell/\eta)^2/2}}{\sqrt{2\pi}[F_\nu(\beta) - F_\nu(\alpha)]} \times\left\{\Phi\left(\frac{\eta^{1/2}(\beta - at)}{b}\right) - \Phi\left(\frac{\eta^{1/2}(\alpha - at)}{b}\right)\right\} f(\eta)dtd\eta.$$

Using the transformation $t = \eta^{-1/2}w + \eta^{-1}\ell$, one finds

$$M_T(\ell) = \int_0^\infty \frac{e^{\ell^{\,2}/(2\eta)}E_W[\Phi(b(W)) - \Phi(a(W))]}{\{F_\nu(\beta) - F_\nu(\alpha)\}} f(\eta)d\eta,$$

where E_W denotes that the expectation is taken with respect to $W \sim N(0,1)$ variable. Here $a(W) = \eta^{1/2}\alpha/b - a(W + \eta^{-1/2}\ell)/b$ and $b(W) = \eta^{1/2}\beta/b - a(W + \eta^{-1/2}\ell)/b$. Thus we can derive the mgf (4.2) by applying the well-known fact that $E_W[\Phi(hW + k)] = \Phi(k/\sqrt{1+h^2})$ (Zacks, 1981, pp. 53–54). □

4.2. Moments of the Perturbed Distribution

Naturally, the moments of T can be obtained by using the moment generating function differentiation. For example,

$$E[T] = M_T'(\ell)|_{\ell=0} = -\frac{\rho}{\delta_1} E_\eta \left[\eta^{-1/2}\{\phi(\beta^*) - \phi(\alpha^*)\}\right], \tag{4.3}$$

where $\delta_1 = F_\nu(\beta) - F_\nu(\alpha)$, $\alpha^* = \eta^{1/2}\alpha$ and $\beta^* = \eta^{1/2}\beta$. Unfortunately, for higher moments this rapidly becomes tedious.

An alternative procedure makes use of the fact that

$$\frac{d}{dx}[x^{k+1}\phi(x)] = (k+1)x^k\phi(x) - x^{k+2}\phi(x) \tag{4.4}$$

for $k = -1, 0, 1, 2, 3, \ldots$, yields the following result.

Assume η is fixed and let $W = \eta^{1/2}T$. Under the distribution (4.1), the relation (4.4) and integrating by parts gives the following conditional moment.

$$\begin{aligned} E[(k+1)W^k - W^{k+2} \mid \eta] &= \int_{-\infty}^{\infty} \{(k+1)w^k - w^{k+2}\} g(w) dw \\ &= \frac{1}{\Phi(\beta^*) - \Phi(\alpha^*)} \int_{-\infty}^{\infty} \{(k+1)w^k - w^{k+2}\}\phi(w) \\ &\quad \times\{\Phi(\beta^*/b - aw/b) - \Phi(\alpha^*/b - aw/b)\} dw \\ &= \frac{\rho b^{k+1}}{\Phi(\beta^*) - \Phi(\alpha^*)} \Big\{\phi(\beta^*) E[V + a\beta^*/b]^{k+1} \\ &\quad - \phi(\alpha^*) E[V + a\alpha^*/b]^{k+1}\Big\} \end{aligned}$$

for $k \geq -1$. By introducing the distribution η, we have the unconditional moment of $E[(k+1)W^k - W^{k+2}]$, that is

$$\frac{\rho b^{k+1}}{\delta_1} \int_0^\infty \left\{\phi(\beta^*) E[V + a\beta^*/b]^{k+1} - \phi(\alpha^*) E[V + a\alpha^*/a]^{k+1}\right\} f(\eta) d\eta \tag{4.5}$$

for $k = -1, 0, 1, \ldots$, where V is a $N(0, 1)$ variate.

By setting $k = -1, 0, 1$, we obtain three expressions, which may be solved to yield the first three moments of T. Higher moments could be found similarly. One obtains,

$$\begin{aligned} E[T] &= -\frac{\rho}{\delta_1} E_\eta \left[\eta^{-1/2}\{\phi(\beta^*) - \phi(\alpha^*)\}\right] \\ &= \frac{\rho\nu^{\nu/2}}{2\delta_1\sqrt{\pi}} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \left\{A^{-(\nu-1)/2} - B^{-(\nu-1)/2}\right\} \quad \text{for } \nu > 1, \end{aligned}$$

$$\begin{aligned}E[T^2] &= E_\eta[\eta^{-1}] - \frac{\rho^2}{\delta_1} E_\eta\Big[\eta^{-1}\{\beta^*\phi(\beta^*) - \alpha^*\phi(\alpha)^*\}\Big] \\ &= \frac{\nu}{\nu-2} + \frac{\rho^2\nu^{\nu/2}}{2\delta_1\sqrt{\pi}}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\Big\{\alpha A^{-(\nu-1)/2} - \beta B^{-(\nu-1)/2}\Big\} \quad \text{for } \nu > 2, \\ E[T^3] &= -\frac{\rho}{\delta_1} E_\eta\Big[\eta^{-3/2}\Big\{(3-\rho^2+\beta^*\rho^2)\phi(\beta^*) - (3-\rho^2+\alpha^*\rho^2)\phi(\alpha^*)\Big\}\Big] \\ &= \frac{\rho\nu^{\nu/2}}{4\delta_1\sqrt{\pi}\Gamma(\nu/2)}\Big[(3-\rho^2)\Gamma((\nu-3)/2)\Big\{A^{-(\nu-3)/2} - B^{-(\nu-3)/2}\Big\} \\ &\quad + \sqrt{2}\rho^2\Gamma((\nu-2)/2)\Big\{\alpha A^{-(\nu-2)/2} - \beta B^{-(\nu-2)/2}\Big\}\Big] \quad \text{for } \nu > 3,\end{aligned}$$

where $A = \nu + \alpha^2$ and $B = \nu + \beta^2$. By using the Binomial expansion, one can see, from Property 3.3, that the general formula for the moments of $Z \sim \mathcal{PT}_{(\alpha,\,\beta))}(\theta_1, \sigma_1, \rho, \nu)$ is

$$E[Z^k] = \sum_{j=0}^{k} \binom{k}{j} \theta_1^{k-j}\sigma_1^j E[T^j]. \tag{4.6}$$

When $\theta_1 = 0$, $E[Z^k] = \sigma_1^k E[T^k]$. It is also noted that existence of the moments depends on the mixing distribution $\eta \sim Gamma(\nu/2, 2/\nu)$. Given the variance T, $Var(T) = E[T^2] - (E[T])^2$, the skewness of $Z \sim PT_{(a,b)}(\theta, \Sigma, \nu)$ variable is given by the following theorem.

THEOREM 4.2. *Let $\mu^{(3)}_{U_{(u(a),\,u(b))}}$ be the skewness (the standardized third central moment) of $U_{(u(a),\,u(b))} \sim Tt_\nu(u(a),\ u(b))$. Then, for $\nu > 3$, the skewness $\mu_Z^{(3)}$ of the $Z \sim \mathcal{PT}_{(u(a),\,u(b))}(\theta_1, \sigma_1, \rho, \nu)$ variable is*

$$\mu_Z^{(3)} = \frac{\lambda\Big\{\lambda^2\sigma^3_{U_{(u(a),\,u(b))}}\mu^{(3)}_{U_{(u(a),\,u(b))}} + K(\nu)\Big\}}{\sigma^3}, \tag{4.7}$$

where $Tt_\nu(\alpha, \beta)$ denotes a truncated t_ν-distribution with respective lower and upper truncation points α and β, $\sigma^2 = (1+\lambda^2)Var(T)$, $\sigma^2_{U_{(u(a),\,u(b))}} = \{(1+\lambda^2)Var(T) - \nu/(\nu-2)\}/\lambda^2$, and

$$\begin{aligned}K(\nu) = \frac{3\Gamma((\nu-3)/2)\nu^{\nu/2}}{4\delta_1\sqrt{\pi}\,\Gamma(\nu/2)}\Big[&\Big\{(\nu+a(u)^2)^{-(\nu-3)/2} - (\nu+b(u)^2)^{-(\nu-3)/2}\Big\} \\ &- \frac{\nu(\nu-3)}{\nu-2}\Big\{(\nu+a(u)^2)^{-(\nu-1)/2} - (\nu+b(u)^2)^{-(\nu-1)/2}\Big\}\Big],\end{aligned}$$

where $K(\nu) > 0$ for $u(b)^2 > u(a)^2$, $K(\nu) = 0$ for $u(b)^2 = u(a)^2$, $K(\nu) < 0$ for $u(b)^2 < u(a)^2$, $\lambda = \rho/\sqrt{1-\rho^2}$ and $\delta_1 = F_\nu(u(b)) - F_\nu(u(a))$.

PROOF. Let $u(a) = (a - \theta_1)/\sigma_1 = \alpha$ and $u(b) = (b - \theta_1)/\sigma_1 = \beta$, and let $T^* = (1 + \lambda^2)^{1/2}T$, where $T \sim \mathcal{PT}_{(\alpha,\beta)}(0, 1, \rho, \nu)$. Then the standardized third central moment of $Z = \sigma_1 T + \theta_1$ variable is equivalent to that of T^*, *i.e.*, $\mu_Z^{(3)} = \mu_T^{(3)} = \mu_{T^*}^{(3)}$. From Theorem 3.1, we see that $E[T^* - E(T^*)]^3 = E_\eta E[(\lambda U_{(\alpha,\beta)} - \lambda EU_{(\alpha,\beta)} + V)^3 \mid \eta] = \lambda^3 E[U_{(\alpha,\beta)} - EU_{(\alpha,\beta)}]^3 + 3\lambda E_\eta E[(U_{(\alpha,\beta)} - EU_{(\alpha,\beta)})V^2 \mid \eta] = \lambda^3 \sigma^3_{U_{(\alpha,\beta)}} \mu^{(3)}_{U_{(\alpha,\beta)}} + \lambda K(\nu)$, where $K(\nu) = 3E_\eta E[(U_{(\alpha,\beta)} - EU_{(\alpha,\beta)})V^2 \mid \eta]$ and $\eta \sim Gamma(\nu/2, 2/\nu)$. Thus, $\mu_Z^{(3)} = \mu_{T^*}^{(3)} = \lambda \left\{\lambda^2 \sigma^3_{U_{(\alpha,\beta)}} \mu^{(3)}_{U_{(\alpha,\beta)}} + K(\nu)\right\} / \sigma^3_{T^*}$, where $\sigma^2_{T^*} = (1+\lambda^2)Var(T) = \lambda^2 \sigma^2_{U_{(\alpha,\beta)}} + \sigma^2_V$. One can observe that $U_{(\alpha,\beta)} \sim Tt_\nu(\alpha, \beta)$, a truncated t_ν-distribution, and $V \sim t_\nu$. Since $\sigma_V^2 = \nu/(\nu - 2)$, we have $\sigma^2_{U_{(\alpha,\beta)}} = \{(1 + \lambda^2)Var(T) - \nu/(\nu - 2)\}/\lambda^2$. After some algebra using the moments of a truncated normal distribution (Johnson *et al.*, 1994, p. 156), we obtain the expression of $K(\nu)$ given by

$$
\begin{aligned}
K(\nu) &= 3\{E_\eta E[U_{(\alpha,\beta)} V^2 \mid \eta] - E_\eta E[U_{(\alpha,\beta)} \mid \eta] E_\eta E[V^2 \mid \eta]\} \\
&= \frac{3}{F_\nu(\beta) - F_\nu(\alpha)} \Big\{ E_\eta[\eta^{-3/2}(\phi(\eta^{1/2}\alpha) - \phi(\eta^{1/2}\beta))] \\
&\quad - \frac{\nu}{\nu - 2} E_\eta[\eta^{-1/2}\phi(\eta^{1/2}\alpha) - \phi(\eta^{1/2}\beta)] \Big\}.
\end{aligned}
$$

Evaluating the expectation with respect to $\eta \sim Gamma(\nu/2, 2/\nu)$ variable gives the expression for $K(\nu)$. The condition for $K(\nu) = 0$ is trivial. For $\beta^2 > \alpha^2$, we see that $\nu\{(\nu+\alpha^2)^{-(\nu-1)/2} - (\nu+\beta^2)^{-(\nu-1)/2}\} < \{\nu/(\nu+\alpha^2)\}\{(\nu+\alpha^2)^{-(\nu-3)/2} - (\nu+\beta^2)^{-(\nu-3)/2}\} < \{(\nu+\alpha^2)^{-(\nu-3)/2} - (\nu+\beta^2)^{-(\nu-3)/2}\}$. Thus the inequality, $K(\nu) > 0$ holds for $\beta^2 > \alpha^2$. Similar argument gives the condition for $K(\nu) < 0$. □

COROLLARY 4.1.

(i) For $|u(a)| < u(b)$, the $PT_{(a,b)}(\theta, \Sigma, \nu)$ distribution is skewed to the right (the left) when $\rho > 0$ ($\rho < 0$).

(ii) For the case $|u(a)| > u(b)$, the distribution is skewed to the right (the left) when $\rho < 0$ ($\rho > 0$).

(iii) Finally, for the case $|u(a)| = u(b)$, the distribution is symmetric.

PROOF. Since $U_{(u(a),\, u(b))} \sim Tt_\nu(u(a), u(b))$, it is obvious that $\mu^{(3)}_{U_{(u(a),\, u(b))}} > 0$ for $|u(a)| < u(b)$, $\mu^{(3)}_{U_{(u(a),\, u(b))}} < 0$ for $|u(a)| > u(b)$, and $\mu^{(3)}_{U_{(u(a),\, u(b))}} = 0$ for $|u(a)| = u(b)$. This fact and Theorem 4.2 immediately give the results. □

We see that Figure 3.1 and Property 3.5 coincide with the results of Corollary 4.1. The class of $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution includes well-known skew-elliptical distributions as in the following two examples.

EXAMPLE 4.1. One member of the $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution is the skew-normal distribution, $\mathcal{SN}(\theta_1, \sigma_1, \lambda)$, by Azzalini (1985). This is obtained from (3.6) by setting $a = \theta_1$, $b = \infty$ and η being degenerate with $\eta = 1$ in (3.7). Its *pdf* reduces to

$$f_Z^*(z) = 2/\sigma_1 \phi(u(z))\Phi(\lambda u_1(z)), \quad z \in \mathbb{R}. \tag{4.8}$$

From (4.6), one finds the moments:

$$E[Z] = \theta_1 + \sigma_1 \rho \sqrt{2/\pi}, \quad Var(Z) = \sigma_1^2(1 - 2\rho^2/\pi)$$

and

$$\mu_Z^{(3)} = (4/\pi - 1)\sqrt{2/\pi}\rho^3(1 - 2\rho^2/\pi)^{-3/2}. \tag{4.9}$$

These values of Z agree with those given in Arnold *et al.* (1993). See Azzalini (1985) and Henze (1986) for the other properties of the distribution.

EXAMPLE 4.2. Another member of the $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution is the generalized skew-t_ν distribution. When $a = \theta_1$ and $b = \infty$, the *pdf* in (3.4) reduces to

$$f_Z^*(z) = 2/\sigma_1 f_\nu(u_1(z)) F_{\nu+1}\left(\frac{\lambda u_1(z)\sqrt{\nu + 1}}{\sqrt{\nu + u_1(z)^2}}\right), \quad z \in \mathbb{R}. \tag{4.10}$$

This distribution is equivalent to the generalized skew-t_ν distribution, written $\mathcal{ST}(\theta_1, \sigma_1, \nu, \lambda)$, considered by Kim (2002) and Branco and Dey (2001) . From (4.6), we obtain, for $\nu > 1$,

$$\begin{aligned} E[Z] &= \theta_1 + \sigma_1 \rho \sqrt{\nu/\pi}\,\Gamma((\nu - 1)/2)/\Gamma(\nu/2), \\ Var(Z) &= \sigma_1^2 \nu/(\nu - 2) - E[Z]^2 \ \text{ for } \nu > 2, \\ \mu_Z^{(3)} &= A/Var(Z)^{3/2} \ \text{ for } \nu > 3, \end{aligned} \tag{4.11}$$

where

$$\begin{aligned} A &= \sigma_1^3 \lambda \left\{ \left(\frac{\nu}{\pi}\right)^{1/2} \frac{3\nu\Gamma((\nu - 1)/2)}{(1 + \lambda^2)^{3/2}(\nu - 2)\Gamma(\nu/2)} \left(\frac{1}{\nu - 3} + \lambda^2 B\right)\right\}, \\ B &= \left\{\frac{2(\nu - 2)}{3(\nu - 3)} + \frac{2(\nu - 2)\Gamma\{(\nu - 1)/2\}^2}{3\pi\Gamma\{\nu/2\}^2} - 1\right\} \ \text{ and } \lambda = \frac{\rho}{\sqrt{1 - \rho^2}}. \end{aligned}$$

These two examples show that $\mathcal{PT}_{(u(a),\,u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution extends the skew-normal and skew-t distributions where only the case of $(\theta_2 = 0, \sigma_2 = 1)$ and $(a = \theta_1, b = \infty)$ is considered. See Kim (2002) and Branco and Dey (2001) for the other properties and applications of the distribution (4.10).

5. MCMC METHOD

In this section we develop computational procedure for the estimation of the perturbed t_ν-distribution by using a Markov chain Monte Carlo (MCMC). In order to specify the model (3.4) for MCMC computation we use the hierarchical setup of the Z distribution.

5.1. *A Metropolis-Hastings Algorithm*

Let $Z_1, Z_2, \ldots, Z_n$ be a random sample of size n obtained from the $\mathcal{PT}_{(u(a),\,u(b))}$ $(\theta_1, \sigma_1, \rho, \nu)$ distribution. Then, using Theorem 3.1 and Property 3.5 we have the following full likelihood specification.

$$\begin{aligned}
Z_i|(y_i, \theta_1, \sigma_1, \rho, u(b), \eta_i) &\sim N(\rho\sigma_1 y_i + \theta_1, \sigma_1^2(1-\rho^2)/\eta_i) \text{ for } i = 1, \ldots, n,\\
Y_i &\sim N(0, \eta_i^{-1}) I(u(a) < y_i < u(b)) \text{ for } i = 1, \ldots, n,\\
\theta_1 &\sim N(\theta_0, \tau^2),\\
\sigma_1 &\sim \sqrt{S}\chi_m^{-1},\\
\rho &\sim Uniform(-1, 1),\\
u(b) &\sim Uniform(F_\nu^{-1}(p), 4),\\
\eta_i &\sim Gamma(\nu/2, 2/\nu) \text{ for } i = 1, \ldots, n,\\
\nu &\sim \pi(\nu) I(\nu > 3),
\end{aligned}$$

where $p = F_\nu(u(b)) - F_\nu(u(a))$, $u(a) = F_\nu^{-1}(F_\nu(u(b)) - p)$. In the the perturbed normal case, the last two distributional specifications are omitted and F_ν is changed to Φ. Although the parameter ν is traditionally taken as an integral, it can be treated as a continuous parameter taking positive values since the associated densities are well defined in this case. We assume the distribution ν is constrained $(\nu > 3)$. The constraint assures the finiteness of the mean and variance of the associated perturbed t_ν distribution. As usual in the case of selection modeling, we assume that p_ν is priori known.

All of the full conditional distributions of the parameters are given by

$$Y_i|(z_i, \theta_1, \sigma_1, \rho, u(b), \eta_i, \nu) \sim N(\rho x_i, (1-\rho^2)/\eta_i) I(u(a) < y_i < u(b)),$$

$$
\begin{aligned}
\theta_1|(\mathbf{y},\mathbf{z},\sigma_1,\rho,u(b),\eta,\nu) &\sim N\left(\theta^*, \frac{\tau^2\sigma_1^2(1-\rho^2)}{\tau^2\sum_{i=1}^n \eta_i + \sigma_1^2(1-\rho^2)}\right),\\
\eta_i|(\mathbf{y},\mathbf{z},\theta_1,\sigma_1,\rho,u(b),\nu) &\sim Gamma\left(\frac{\nu+2}{2}, \frac{2(1-\rho^2)}{x_i^2-2\rho x_i y_i + y_i^2 + (1-\rho^2)\nu}\right),\\
p(\sigma_1|\mathbf{y},\mathbf{z},\theta_1,\rho,u(b),\eta,\nu) &\propto \sigma_1^{-(n+m+1)}\exp(-Q^*),\\
p(\rho|\mathbf{y},\mathbf{z},\theta_1,\sigma_1,u(b),\eta,\nu) &\propto (1-\rho^2)^{-n/2}\exp\left\{-\frac{\sum_{i=1}^n \eta_i(y_i^2-2\rho y_i x_i + x_i^2)}{2(1-\rho^2)}\right\},\\
p(u(b)|\mathbf{y},\mathbf{z},\theta_1,\sigma_1,\rho,\eta,\nu) &\propto \prod_{i=1}^n \left(\Phi(\eta_i^{1/2}u(b)) - \Phi(\eta_i^{1/2}u(a))\right)^{-1}\\
&\quad \times I(F_\nu^{-1}(p) < u(b) < 4),
\end{aligned}
$$

where $u(a) = F_\nu^{-1}(F_\nu(u(b)) - p)$,

$$
x_i = (z_i - \theta_1)/\sigma_1, \quad \theta^* = \frac{\tau^2\sum_{i=1}^n \eta_i(z_i - \rho\sigma_1 y_i) + (1-\rho^2)\sigma_1^2\theta_0}{\tau^2\sum_{i=1}^n \eta_i + (1-\rho^2)\sigma_1^2}
$$

and

$$
Q^* = \frac{(1-\rho^2)S + \sum_{i=1}^n \eta_i(z_i-\theta_1)^2 - 2\rho\sigma_1\sum_{i=1}^n \eta_i y_i(z_i-\theta_1)}{2\sigma_1^2(1-\rho^2)}.
$$

Finally, we sample ν from its full conditional distribution. The conditional likelihood function $L(\nu|\mathbf{z},\theta_1,\sigma_1,\rho,u(b))$ is $\prod_{i=1}^n f_Z^*(z_i)$, where $f_Z^*(z)$ is defined in (3.4). The full conditional posterior density is proportional to $L(\nu|\mathbf{z},\theta_1,\sigma_1,\rho,u(b))\pi(\nu)$ $I(\nu > 3)$. Since $u(a)$ and $u(b)$ appearing in the full conditional distributions of Y_i and $u(b)$ involve ν, this is a simple way of obtaining the full conditional

Because of the complexity of the full conditional distributions, we use Metropolis steps to generate σ_1, ρ, $u(b)$, and ν. Then the Metropolis-Hastings algorithm is obtained by drawing from all the full conditionals of the parameters in turn, proceeding until convergence.

A sample of Y_i variable is easily extracted from the truncated normal distribution by using the method documented by Devroye (1986). We have, however, little knowledge about the shapes of the full conditional distributions of σ_1, ρ, $u(b)$, and ν. This fact suggests using the Metropolis-Hastings sampling algorithm (Gustafson, 1998). For the Metropolis step for sampling σ_1 and ρ, Kim (2005) elaborated the Random Walk Metropolis algorithm by applying parameter deconstraint transformations to σ_1 and ρ. Therefore, to complete computational procedure for the Metropolis-Hastings sampling algorithm under the above likelihood specification, we need to develop algorithms to sample $u(b)$ and ν from their respective full conditional distributions.

5.2. Random Variate Generations of ν and $u(b)$

It is convenient to transform ν to ψ where $\psi = \nu^{-1}$. Then the full conditional posterior density for ψ with a $Uniform(0, 1/3)$ prior density $\pi(\psi)$ is

$$\pi(\psi|\mathbf{y}, \mathbf{z}, \theta_1, \sigma_1, \rho, u(b), \eta) \propto \pi(\psi)h(\psi|\mathbf{y}, \mathbf{z}, \theta_1, \sigma_1, \rho, u(b), \eta), \tag{5.1}$$

where

$$h = \psi^{-2}L(\nu|\mathbf{z}, \theta_1, \sigma_1, \rho, u(b))I(\nu > 3),$$

$h = h(\psi|\mathbf{z}, \theta_1, \sigma_1, \rho, u(b))$. A way of generating ν is to use a Metropolis step (Chib and Greenberg, 1995) using the $Uniform$ prior on ψ. We set a proposal density $q(\psi, \psi^*) = \pi(\psi^*)$ which supplies candidate values ψ^* given the current value of ψ. In this case, the probability of move requires only the computation of h function. Thus the $(k+1)$th iteration of the Metropolis step is given by

Step 1. Generate ψ^* from a $Uniform(0, 1/3)$.

Step 2. Generate u from a $Uniform(0, 1)$.

Step 3. If $u < h(\psi^*|\theta_1, \sigma_1, \rho, p_\nu, u(b), \mathbf{z}, \mathbf{y})/h(\psi^{(k)}|\theta_1, \sigma_1, \rho, p_\nu, u(b), \mathbf{z}, \mathbf{y})$ then $\psi^{(k+1)} = \psi^*$; otherwise, $\psi^{(k+1)} = \psi^{(k)}$.

After obtaining $\psi^{(k+1)}$, we compute $\nu^{(k+1)}$ by using the relation $\nu = 1/\psi$. Note that the proposal density need not to enforce the interval constraint, because it is a $Uniform$ distribution on $0 < \psi \leq 1/3$.

The same algorithm applies for sampling $u(b)$ form its full conditional posterior density:

$$\pi(u(b)|\theta_1, \sigma_1, \rho, \mathbf{z}, \mathbf{y}) \propto \pi(u(b)) \prod_{i=1}^{n} \left[\{\Phi(\eta_i^{1/2}u(b)) - \Phi(\eta_i^{1/2}u(a))\}\right]^{-1}, \tag{5.2}$$

where $u(a) = F_\nu^{-1}(F_\nu(u(b)) - p)$ and $\pi(u(b))$ is the $Uniform(F_\nu^{-1}(p), 4)$ prior density for $u(b)$. Note that, for generating $u(b)$, Step 1 of the above Metropolis algorithm need to be changed to "Generate $u(b)^*$ from a $Uniform(F_\nu^{-1}(p_\nu), 4)$", and $u(a)^*$ can be calculated from $u(a)^* = F_\nu^{-1}(F_\nu(u(b)^*) - p)$. Thus we use a Metropolis step to draw σ_1, ρ, ν and $u(b)$ and the Gibbs sample is obtained by drawing $\mathbf{y}$, η, θ_1, σ_1, ρ, $u(b)$ and ν in turn, after convergence.

6. A Simulation Study

Our example is an illustration of extensive studies we have undertaken to validate the MCMC method. We generated n observations from $\mathcal{PT}_{(u(a),\,u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution using the algorithm in Section 3 and then ran the Gibbs sampler for 60,000 iterations. We repeated this procedure 100 times. For the starting points of the sampler, it appears that the sample mean and the sample standard deviation are reasonable starting points for θ_1 and σ_1. In an attempt to test the robustness of the sampler, we started ρ, b, and ν well away from their true values, *i.e.* the true value of ρ plus 0.2, the true value of ν plus 2, and the true value of b plus 0.3. The hyperparameter specification was defined by $\delta = 0$, $\tau = \sqrt{10}$, $m = 3$ and $S = 10$, reflecting rather vague initial information relative to that to be provided by the data. Further, we set the hyperparameter $p = 0.6$, which usually be provided in the selection modeling case. A simulation study with various sample sizes and set of true parameter values is conducted and estimation results by 100 repetitions of the MCMC method are listed in Table 6.1. As given in Table 6.1, we see that the method produces accurate estimates for the parameters of $\mathcal{PT}_{(u(a),\,u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution.

For calculating the estimates, the iterative process was monitored by observing trace of Gibbs samples. The diagnostic we used are described in Cowles and Carlin (1996). For each data set, we used 40,000 iterations to "burn in" the sampler; the decision is based on trace plots of the ergodic averages of the trace of the parameters, $(\theta_1, \sigma_1, \rho, \nu, u(b))$, leading us to believe that convergence has been attained before 40,000 iterations. By adjusting the tuning constant (standard deviations of the transition density in the RW Metropolis algorithm), we were able to keep the jumping probabilities between 0.23 and 0.5 (Gelman *et al.*, 1996; Robert *et al.*, 1997).

7. Conclusion

This paper has presented a new class of perturbed t_ν-distributions and its Bayesian estimation. To form the class of distributions, we considered a conditioning method to the bivariate t_ν-distributions. This introduces yet other weight function, inducing perturbation of the symmetry with univariate t_ν-distribution, that brings additional flexibility of modeling skewed and heavy tailed distribution as well as high kurtosis distribution. Several properties of the class are studied. The study shows that we have at hand a class of distributions with following

TABLE 6.1 *Estimation Results for* $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$ *distribution. Standard Error of the Estimate is in the Parenthesis*

True Value		*Mean of 100 Repeated Posterior Means (Standard Error)*				
$(\rho,\ \nu,\ \theta_1,\ \sigma_1,\ u(b))$	n	$\hat{\rho}$	$\hat{\nu}$	$\hat{\theta}_1$	$\hat{\sigma}_1$	$\hat{u(b)}$
(0.3, 10, 5, 2, 1.5)	20	0.351(0.077)	9.418(1.601)	4.551(0.263)	2.318(0.295)	1.422(0.122)
	50	0.341(0.078)	9.387(1.451)	4.696(0.190)	2.156(0.208)	1.410(0.090)
	100	0.313(0.075)	9.833(1.328)	4.875(0.049)	2.063(0.127)	1.399(0.108)
(0.3, 10, -1, 2, 1.5)	20	0.417(0.115)	12.164(1.760)	-1.043(0.268)	1.880(0.196)	1.336(0.197)
	50	0.256(0.092)	11.123(1.186)	-1.041(0.262)	1.913(0.181)	1.436(0.148)
	100	0.262(0.058)	9.116(1.325)	-0.973(0.239)	2.073(0.189)	1.415(0.067)
(0.5, 10, 5, 2, 1.5)	20	0.548(0.073)	9.468(1.843)	4.699(0.275)	2.215(0.280)	1.423(0.166)
	50	0.465(0.078)	9.401(1.528)	4.896(0.265)	2.040(0.296)	1.405(0.165)
	100	0.468(0.077)	9.877(1.316)	5.047(0.135)	1.973(0.146)	1.401(0.059)
(0.5, 5, 10, 2, 1.5)	20	0.547(0.064)	6.164(1.756)	9.463(0.339)	2.179(0.217)	1.403(0.162)
	50	0.541(0.056)	6.195(1.546)	9.915(0.177)	2.150(0.140)	1.419(0.136)
	100	0.537(0.039)	5.621(1.315)	9.958(0.113)	2.152(0.121)	1.460(0.098)
(0.8, 10, 5, 2, 1.5)	20	0.686(0.074)	9.562(1.347)	4.984(0.207)	1.903(0.229)	1.425(0.051)
	50	0.732(0.093)	9.950(1.168)	5.178(0.156)	1.894(0.161)	1.418(0.056)
	100	0.741(0.075)	10.280(1.143)	5.209(0.116)	2.105(0.107)	1.408(0.069)
(-0.8, 10, 5, 2, 1.5)	20	-0.729(0.071)	12.263(1.594)	4.857(0.301)	1.849(0.180)	1.412(0.059)
	50	-0.738(0.062)	11.373(1.443)	4.893(0.279)	1.894(0.292)	1.452(0.057)
	100	-0.823(0.047)	11.249(1.171)	4.889(0.208)	1.979(0.257)	1.447(0.048)

properties:

(i) inclusion of the normal, skewed-normal, t_ν, skewed-t_ν, and a perturbed normal,

(ii) mathematical tractability,

(iii) wide range of indices of skewness and kurtosis,

(iv) applicability in solving a screening problem in statistical inference.

Form the applied view point, the properties (i), (ii) and (iii) imply that the class is useful for modeling random phenomena which have heavier (or righter) tails than the normal as well as having some skewness. Furthermore, the simulation study in Section 6 shows that the class of distributions is useful for the robust estimation of θ_1 and σ_1 form a sample with asymmetric outliers in the observed sample from a true t_ν (or normal) distribution with mean θ_1 and scale parameter σ_1.

A different point of view is given by a selection modeling. Property 3.1 implies that (iv) is immediate. Property 3.1 says that, by using $\mathcal{PT}_{(u(a),\, u(b))}(\theta_1, \sigma_1, \rho, \nu)$ distribution, one can solve the following screening problem: Consider the case in which Y represents the or screening variable (permitting upper and lower truncation on Y) and X represents the variable that is measured following initial screening. We assume that X values are available only for the nontruncated Y values, while the values of the random variable Y are not available. In this case the screening problem is to estimate θ_1, σ_1 and ν of the marginal distribution of X as well as the correlation ρ between X and Y. The MCMC method provides estimates of θ_1, σ_1 and ν of X in the original unscreened population. Moreover, an estimate of the correlation between the two variables in original unscreened population can be also obtained, even though no Y observations are available.

REFERENCES

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.

AZZALINI, A. (1985). "A class of distributions which includes the normal ones". *Scandinavian Journal of Statistics*, **12**, 171–178.

BRANCO, M. D. AND DEY, D. K. (2001). "A general class of multivariate skew-elliptical distributions", *Journal of the Multivariate Analysis*, **79**, 99–113.

BAYARRI, M. J. AND DEGROOT, H. M. (1992). "A BAD view of weighted distribution and selection models", *Bayesian Statistics*, Oxford University Press, Oxford.

CHIB, S. AND GREENBERG, E. (1995). "Understanding the Metropolis-Hastings Algorithm", *The American Statistician*, **49**, 327–335.

COWLES, M. AND CARLIN, B. (1996). "Markov chain Monte Carlo diagnostics: A comparative review", *Journal of the American Statistical Association*, **91**, 883–904.

DEVROYE, L. (1986). *Non-Uniform Random Variate Generator*, Springer Verlag, New York.

DUNNETT, C. W. AND SOBEL, M. (1954). "A bivariate generalization of Student's t-distribution, with tables for certain special cases", *Biometrika*, **41**, 153–169.

FANG, K. T., KOTZ, S. AND NG, K. W. (1990). *Symmetric Multivariate and Related Distributions*, Chapman & Hall, New York.

FANG, K. T. AND ZHANG, Y. T. (1990). *Generalized Multivariate Analysis*, Springer-Verlag, New York.

GELMAN, A. ROBERT, G. O. AND GILKS, W. R. (1996). "Efficient Metropolis jumping rules", *Bayesian Statistics*, Oxford University Press, Oxford.

GUSTAFSON, P. (1998). "A guided walk Metropolis algorithm", *Statistics and Computing*, **8**, 357–364.

HENZE, N. (1986). "A probabilistic representation of the Skew-normal distribution", *Scandinavian Journal of Statistics*, **13**, 271–275.

JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, John Wiley & Sons, New York.

KIM, H. J. (2002). "Binary regression with a class of skewed t link models", *Communications in Statistics: Theory and Methods*, **31**, 1863–1886.

KIM, H. J. (2005). "On Bayesian estimation and properties of the marginal distribution of a truncated bivariate t-distribution", *Journal of the Korean Statistical Society*, **34**, 245–261.

MA, Y., GENTON, M. G. AND TSIATIS, A. (2005). "Locally efficient semiparametric estimators for generalized skew-elliptical distributions", *Journal of the American Statistical Association*, **100**, 980–989.

RAO, C. R. (1985). "Weighted distributions arising out of methods of ascertainment: What population does a sample represent ?", *A Celebration of Statistics: The ISI Centenary Volume* (A. G. Atkinson and S. E. Fienberg, eds.), Springer, New York.

ROBERT, G. O., GELMAN, A. AND GILKS, W. R. (1997). "Weak convergence and optimal scaling of random walk Metropolis algorithm", *Annals of Applied Probability*, **7**, 110–120.

ZACKS, S (1981). *Parametric Statistical Inference*, Pergamon Press, Oxford.