A GENERAL CLASS OF MULTIVARIATE SKEW-ELLIPTICAL DISTRIBUTIONS

Márcia D. Branco
University of Sao Paulo, Brasil
E-mail:mbranco@ime.usp.br

Dipak K. Dey
University of Connecticut, USA
E-mail:dey@merlot.stat.uconn.edu

Abstract

This paper proposes a general class of multivariate skew-elliptical distributions. We extend earlier results on the so called multivariate skew-normal distribution. This family of distributions contains the multivariate normal, Student-t, exponential power and Pearson type II, but with an extra parameter to regulate skewness. We also obtain the moment generating functions and study some distributional properties. Several examples are provided.

Some key words: Elliptical distributions, exponential power family, mixture of normals, Pearson type II, skewness.

AMS 1991 subject classification: 60E05, 62H10.

1. INTRODUCTION

This article is devoted to modeling with a new class of continuous multivariate distributions that can simultaneously account for both skewness and heavy tails. Interesting special cases are discussed in Kelker(1970), Fang and Zhang(1990), Fang, Kotz and Ng(1990), Azzalini and DallaValle(1996) and Azzalini and Capitanio(1999).

The general class of multidimensional distributions that we present here will be useful to modeling multivariate random phenomena which have heavier tails than the normal as well as having some skewness. Such a rich class of distributions can be used to model multivariate regression problems with skew-elliptical error structure.

To define formally the class of multivariate skew elliptical distributions, we first need to define the multivariate elliptical distributions. Here we use the notation

$\boldsymbol{X} \sim El_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \phi)$ to indicate that $\boldsymbol{X}$ is a k-dimensional random vector, elliptically distributed with location vector $\boldsymbol{\mu} \in \mathbb{R}^k$ and a $k \times k$ (positive definite) dispersion matrix $\boldsymbol{\Sigma}$ and characteristic function (c.f.) ϕ. When $P(\boldsymbol{X} = \boldsymbol{0}) = 0$ and the density exist, this is given by

$$f(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-1/2} g^{(k)}[(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})], \tag{1.1}$$

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

$$\int_0^\infty u^{\frac{k}{2}-1} g^{(k)}(u) du = \frac{\Gamma[k/2]}{\pi^{k/2}} \tag{1.2}$$

which implies that $g^{(k)}$ is a spherical k-dimensional density. When the density function of the elliptical distribution exists, we use the density generator function $g^{(k)}$ and replace the characteristic function φ in the notation and use $X \sim El_k(\mu\Sigma; g^{(k)})$. A comprehensive review of the properties and characterizations of elliptical distributions can be found in Fang et al. (1990). The term skew elliptical (SE) will refer to parametric class of multivariate probability distributions came from the vector $\boldsymbol{Y} = [\boldsymbol{X} \mid X_0 > 0]$, where $\boldsymbol{X} \sim El_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \phi_1)$ and $X_0 \sim El_1(0, 1, \phi_2)$. The notation is $\boldsymbol{Y} \sim SE_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; \phi)$, means $\boldsymbol{Y}$ is a k dimensional skew elliptical distribution with location $\boldsymbol{\mu}$, scale $\boldsymbol{\Omega}$, characteristic function ϕ and skewness parameter $\boldsymbol{\delta}$. When the density of $\boldsymbol{Y}$ exists, we write $\boldsymbol{Y} \sim SE_k(\mu, \boldsymbol{\Omega}, \boldsymbol{\delta}; g^{(k+1)})$ with the density function

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2 f_{g^{(k)}}(\boldsymbol{y}) F_{g_{q(\boldsymbol{y})}}(\boldsymbol{\lambda}^T(\boldsymbol{y} - \boldsymbol{\mu})) \tag{1.3}$$

where $f_{g^{(k)}}(.)$ is the p.d.f. of the form (1.1) with generator function $g^{(k)}(.)$ and $F_{g_{q(\boldsymbol{y})}}$ is the c.d.f. of a univariate elliptical distribution with $g_{q(\boldsymbol{y})}$ as the generator function. The detail form will be clear in the next section. Notice that $\lambda = 0$ corresponds to elliptical density given in (1.1). However, the class of p.d.f. given by $f_{g^{(k)}}(.)$ is not the complete class of all the elliptical distributions, but a subclass, where $f_{g^{(k)}}(.)$ is the marginal density for some $(k+1)$-dimensional density.

The interest in density (1.3) comes from both theorical and applied direction. On the theoretical side it enjoys a number of formal properties which resemble those of the elliptical distributions as given in Fang and Zhang (1990). From the applied viewpoint, (1.3) is a unimodal empirical distribution with presence of skewness and possible heavy tail. Many regression and calibration problem arise in practice which can be modeled using such skewed distribution. Such model can also be used for creating skewed link function in generalized linear models as described in Chen, Dey and Shao (1999).

The format of the paper is organized as follows. In Section 2, we derive multivariate skew elliptical distribution. Section 3 is devoted to specific examples of skew elliptical distributions. This section contains a general class of skew normal scale mixtures and skewed-Pearson type II distributions. Moments calculation for such distributions are presented in Section 4. Section 5 provides important properties which resemble those of the elliptical distributions. The paper is conclued in Section 6 with brief discussions.

2. MULTIVARIATE SKEW-ELLIPTICAL DISTRIBUTION

Recently, Azzalini and Dalla Valle(1996) present a general theory for the multivariate version of skew normal distribution. Their paper suggests diferent methods to generate skew normal distributions. In this paper we extend their results to multivariate skew elliptical distribution. We consider here a conditioning method to form a new class of skew elliptical distributions. Consider $\boldsymbol{X} = (X_1, X_2, ..., X_k)^T$ a random vector. Let $\boldsymbol{X}^* = (X_0, \boldsymbol{X}^T)^T$ be a $(k+1)$-dimensional random vector, such that $\boldsymbol{X}^* \sim El_{k+1}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}; \phi)$, where $\boldsymbol{\mu}^* = (0, \boldsymbol{\mu})$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)^T, \phi$ is the characteristic function and the scale parameter matrix $\boldsymbol{\Sigma}$ has the form

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \boldsymbol{\Omega} \end{pmatrix}$$

with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)^T$. Here $\boldsymbol{\Omega}$ is the scale matrix associated to the vector $\boldsymbol{X}$. We say that the random vector $\boldsymbol{Y} = [\boldsymbol{X} \mid X_0 > 0]$ has a skew elliptical distribution and denote for $\boldsymbol{Y} \sim SE_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; \phi)$, where $\boldsymbol{\delta}$ is the skewness parameter. If the density of the random vector $\boldsymbol{X}^*$ exists and $P(\boldsymbol{X}^* = 0) = 0$, then the p.d.f. of $\boldsymbol{Y}$ will be of the form

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2 f_{g^{(k)}}(\boldsymbol{y}) F_{g_{q(\boldsymbol{y})}}(\boldsymbol{\lambda}^T(\boldsymbol{y} - \boldsymbol{\mu})), \tag{2.1}$$

where $f_{g^{(k)}}(.)$ is the p.d.f. of $El_k(\boldsymbol{\mu}, \boldsymbol{\Omega}; g^{(k)})$ and $F_{g_{q(\boldsymbol{z})}}$ is the c.d.f. of $El(0, 1; g_{q(\boldsymbol{z})})$, with

$$\boldsymbol{\lambda}^T = \frac{\boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1}}{(1 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta})^{\frac{1}{2}}}, \tag{2.2}$$

$$g^{(k)}(u) = \frac{2\pi^{\frac{k}{2}}}{\Gamma(\frac{k}{2})} \int_0^\infty g^{(k+1)}(r^2 + u) r^{k-1} dr, \quad u \geq 0, \tag{2.3}$$

$$g_{q(\boldsymbol{y})}(u) = \frac{g^{(k+1)}(u + q(\boldsymbol{y}))}{g^{(k)}(q(\boldsymbol{y}))}, \tag{2.4}$$

and $q(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$. In this case, we denote for $\boldsymbol{Y} \sim SE_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; g^{(k+1)})$, where $g^{(k+1)}$ is the generator function as given in (2.3) with k replaced by $k+1$. From (2.2) and the positive definiteness of the Σ matrix, it follows that δ and Ω must satisfy the condition $\delta^T \Omega^{-1} \delta < 1$. However such a restriction is not severe in practical problems. Note that, the class of the generator functions given in (2.3) is smaller than the general class given in (1.2). It is to be noted that Azzalini and Capitanio (1999) also obtained a class of multivariate skew elliptical distribution of the form similar to (2.1). However, their approach was slightly different than ours.

To derive (2.1), we consider that the p.d.f. of $\boldsymbol{Y}$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) \propto P(X_0 > 0 \mid \boldsymbol{y}) f_{\boldsymbol{X}}(\boldsymbol{y}).$$

Using the fact $\boldsymbol{X}^* \sim El_{(k+1)}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}; g_{k+1})$ and the properties of the elliptical distributions (see Fang and Zhang, 1990) we obtain

$$\boldsymbol{X} \sim El_k(\boldsymbol{\mu}, \boldsymbol{\Omega}; g^{(k)})$$

and

$$X_0 \mid \boldsymbol{y} \sim El(\boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), 1 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta}; g_{q(\boldsymbol{y})}).$$

Observe that $f_{\boldsymbol{X}}(\boldsymbol{y}) = f_{g^{(k)}}(\boldsymbol{y})$. Now, considering $Z_0 = (X_0 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}))/(1 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta})^{1/2}$ and the symmetric property of the elliptical distribution, it follows that $P(X_0 > 0 \mid \boldsymbol{y}) = P(Z_0 > -\frac{(\boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}))}{(1 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta})^{\frac{1}{2}}} \mid \boldsymbol{y}) = F_{g_{q(\boldsymbol{y})}}\left(\boldsymbol{\lambda}^T(\boldsymbol{y} - \boldsymbol{\mu})\right)$. $\diamond$

Using the relation between the elliptical generator function we can get an alternative and convenient expression for the p.d.f of the skew-elliptical distribution as

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2|\boldsymbol{\Omega}|^{-1/2} \int_{-\infty}^{\boldsymbol{\lambda}^T(\boldsymbol{y} - \boldsymbol{\mu})} g^{(k+1)}(r^2 + (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}))dr. \tag{2.5}$$

3. SPECIAL CASES

In this section we develop many examples of multivariate skew elliptical distributions as special cases.

3.1. Skew scale mixture of normal distribution

The generator function for a $(k+1)$-variate scale mixture of normal is

$$g^{(k+1)}(u) = \int_0^\infty (2\pi K(\eta))^{-(k+1)/2} exp\{-u/2K(\eta)\} dH(\eta),$$

where η is a mixing variable with c.d.f. $H(\eta)$ and $K(\eta)$ is a weight function. Then, from (2.5), $f_{\boldsymbol{Y}}(\boldsymbol{y})$ is given as

$$2|\boldsymbol{\Omega}|^{-1/2} \int_{-\infty}^{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})} \int_0^\infty (2\pi K(\eta))^{-(k+1)/2} exp\{-\frac{(r^2 + (\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}))}{2K(\eta)}\} dH(\eta) dr$$

$$= 2\int_0^\infty |\boldsymbol{\Omega}|^{-1/2} (2\pi K(\eta))^{-k/2} exp\{-\frac{((\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}))}{2K(\eta)}\} \Phi\left(\frac{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})}{K(\eta)^{1/2}}\right) dH(\eta).$$

Thus,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2\int_0^\infty \phi_k(\boldsymbol{y}; \boldsymbol{\mu}, K(\eta)\boldsymbol{\Omega}) \Phi\left(\frac{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})}{K(\eta)^{1/2}}\right) dH(\eta), \tag{3.1}$$

where $\phi_k(\boldsymbol{y}; \boldsymbol{\mu}, K(\eta)\boldsymbol{\Omega})$ is the p.d.f. of $N_k(\boldsymbol{\mu}, K(\eta)\boldsymbol{\Omega})$ and Φ is the c.d.f. of the standard normal distribution.

One particular case of this distribution is the skew normal distribution, for which H is degenerate, with $K(\eta) = 1$. In this case the density is given as

$$2\phi_k(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Omega}) \Phi\left(\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})\right).$$

This result was obtained in Azzalini and Dalla Valle(1996). Therefore, if $g^{(k+1)}(.)$ is a scale mixture of normal, it follows from (3.1) that $f_{\boldsymbol{Y}}(.)$ is again a scale mixture of skew normal distribution.

The next examples are special cases of the skew-scale mixture of normal.

Example 3.1.1: Skew finite mixture of normal.

The generator function of the finite mixture of normal is

$$g^{(k+1)}(u) = \sum_{i=1}^n p_i (2\pi K(\eta_i))^{-(k+1)/2} exp\{-u/2K(\eta_i)\},$$

with $0 \leq p_i \leq 1$ and $\sum_{i=1}^{n} p_i = 1$. Then, the distribution H is a discrete measure on $\{\eta_1, \eta_2, \ldots, \eta_n\}$ with probabilities $p_1, p_2, \ldots, p_n$, respectively. The density of the skew finite mixture of normal is given as

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2\sum_{i=1}^{n} p_i \phi_k(\boldsymbol{y}; \boldsymbol{\mu}, K(\eta_i)\boldsymbol{\Omega})\Phi\left(\frac{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})}{K(\eta_i)^{1/2}}\right)$$

which is again a finite mixture of skew-normal distribution. In this case, we often take $K(\eta_i) = 1,\ i = 1, 2, \ldots, n$, for simplicity.

Example 3.1.2: Skew logistic distribution.

The generator function of the logistic distribution is

$$g^{(k+1)}(u) = \frac{exp\{-u\}}{1 + exp\{-u\}}, \qquad u > 0.$$

As pointed out by Choy(1995), the logistic distribution is a special case of the scale mixture of normal, when $K(\eta) = 4\eta^2$ and η follows an asymptotic Kolmogorov distribution with density

$$f(\eta) = 8\sum_{k=1}^{\infty}(-1)^{k+1}k^2\eta exp\{-2k^2\eta^2\}.$$

However, this density is not computationally attractive. Recently, Chen and Dey(1998) present an alternative way to work with this density.

Example 3.1.3: Skew stable distribution.

A skew stable distribution can be obtained as a scale mixture of skew-normal as given in (3.1) with $K(\eta) = 2\eta$ and the mixture distribution $dH(\eta) = S^P(\alpha, 1)$, where the p.d.f. of the positive stable distribution $S^P(\alpha, 1)$ in the polar form is given by

$$h_{SP}(\eta \mid \alpha, 1) = \frac{\alpha}{1-\alpha}\eta^{-(\alpha/(1-\alpha)+1)} \int_0^1 s(u) exp\{-\frac{s(u)}{\eta^{\alpha/1-\alpha}}\}du, \tag{3.2}$$

for $0 < \alpha < 1$ with

$$s(u) = \left\{\frac{sin(\alpha\pi u)}{sin(\pi u)}\right\}^{\frac{\alpha}{1-\alpha}} \left\{\frac{sin[(1-a)\pi u]}{sin(\pi u)}\right\}.$$

See Samorodnitsky and Taqqu(1994) for more details on the positive stable distribution. Note that when $\alpha = 1$, we get a skew Cauchy distribution. Clearly skew normal distribution can also be obtained from skew stable by taking $\alpha \to 1$.

Example 3.1.4: Skew Exponential Power distribution.

A skew exponential power distribution can also be obtained as scale mixture of skew normal as given in (3.1) by choosing $K(\eta) = 1/(2c_0\eta)$ and $h(\eta) = (1/\eta)^{\frac{k+1}{2}} h_{SP}(\eta \mid \alpha, 1)$ where $h_{SP}(\eta \mid \alpha, 1)$ is given in (3.2), and c_0 is defined as

$$c_0 = \frac{\Gamma[3/2\alpha]}{\Gamma[1/2\alpha]} \quad \text{and} \quad \frac{1}{2} \leq \alpha \leq 1.$$

The parameter α is called the kurtosis parameter. See Andrews and Mallows(1974) and Choy(1995) for further references on symmetric exponential power family of distributions. Again skew normal and skew double exponential distributions can be obtained by taking $\alpha = 1$ and $\alpha = 1/2$ respectively.

3.2. Skew t distribution

Recall hat the t-distribution is a special case of scale mixture of normal distribution. To develop the skew t distribution, we can again use (3.1) by considering $K(\eta) = 1/\eta$ with H as the c.d.f. of a gamma distribution, i.e., $\eta \sim G(\nu/2, \nu/2)$, where $\eta \sim G(a, b)$ means that η has the p.d.f. $f(\eta) = b^a exp\{-b\eta\}/\Gamma[a],\ a > 0, b > 0$. However, we will study this case separately, since in this case we have a nice analytic expression for the density function. A particular case of the skew t distribution is the skew Cauchy distribution, when $\nu = 1$. Also when $\nu \to \infty$, we get skew normal distribution as the limiting case.

Let us consider here the generalized version of Student's t distribution. This generalized version can be obtained by considering a new parameter $\tau > 0$, such that, $\eta \sim G(\nu/2, \tau/2)$. To get the usual version of the multivariate t, it is enough to consider $\tau = \nu$ (see Arellano-Valle and Bolfarine, 1995). In this case, the generator function is given as

$$g_{\nu,\tau}(u) = C(\nu, \tau)[\tau + u]^{-\frac{(\nu+k+1)}{2}},$$

where

$$C(\nu, \tau) = \frac{\Gamma[\frac{\nu+k+1}{2}]\tau^{\nu/2}}{\Gamma[\nu/2]\pi^{\frac{k+1}{2}}}.$$

It follows from (2.5) that,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2|\boldsymbol{\Omega}|^{-1/2} \int_{-\infty}^{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})} C(\nu, \tau)[\tau + r^2 + (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{y} - \boldsymbol{\mu})]^{-(\nu+k+1)/2} dr.$$

Considering $\tau^* = \tau + (\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$, $\quad \nu^* = \nu + k$ and

$$C^*(\nu,\tau) = \frac{\Gamma[\frac{\nu+k+1}{2}](\tau + (\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}))^{(\nu+k)/2}}{\Gamma[(\nu+k)/2]\pi^{\frac{1}{2}}} = \frac{\Gamma[\frac{\nu^*+1}{2}](\tau^*)^{\nu^*/2}}{\Gamma[\frac{\nu^*}{2}]\pi^{1/2}},$$

we obtain $f_{\boldsymbol{Y}}(\boldsymbol{y})$ equals to

$$2|\boldsymbol{\Omega}|^{\frac{-1}{2}}\frac{\tau^{\nu/2}[\tau + (\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})]^{-(\nu+k)/2}\Gamma[\frac{\nu+k}{2}]}{\Gamma[\frac{\nu}{2}]\pi^{\frac{k}{2}}}\int_{-\infty}^{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})} C^*(\nu,\tau)[\tau^* + r^2]^{-(\nu^*+1)/2}dr$$

So, that the p.d.f. of the multivariate skew t is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2f_{\nu,\tau}(\boldsymbol{y};\boldsymbol{\mu},\boldsymbol{\Omega})F_{\nu^*,\tau^*}(\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu}))$$

where $f_{\nu,\tau}(.;\boldsymbol{\mu},\boldsymbol{\Omega})$ is the p.d.f. of a k-variate generalized Student's t distribution with location and scale parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$, respectively, and $F_{\nu^*,\tau^*}(.)$ is the c.d.f. of an univariate standard generalized t distribution.

3.3. Skewed Pearson type II distribution

The usual density function of the Pearson type II distribution (see Fang et al.,1990) has the following generator function

$$g^{(k+1)}(u) = \frac{\Gamma[m+1+(k+1)/2]}{\Gamma[m+1]\pi^{(k+1)/2}}(1-u)^m, \quad 0 \le u \le 1, \quad m > -1.$$

We are considering here a generalized version of the Pearson type II distribution with an additional parameter $r > 0$. The generator function for this is

(3.2)
$$g^{(k+1)}(u) = \frac{\Gamma[m+1+(k+1)/2]}{\Gamma[m+1]\pi^{(k+1)/2}}s^{-(m+(k+1)/2)}(s-u)^m, \quad 0 \le u \le s, \quad m > -1, \; s > 0.$$

Using (2.2) and (3.2) we obtain that $f_{\boldsymbol{Y}}(\boldsymbol{y})$ equals to

$$2|\boldsymbol{\Omega}|^{-1/2}\int_{-\sqrt{s}}^{\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})}\frac{\Gamma[m+1+(k+1)/2]}{\Gamma[m+1]\pi^{(k+1)/2}}s^{-(m+(k+1)/2)}[s-(r^2+(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}))]^m dr$$

$$= 2|\boldsymbol{\Omega}|^{-1/2}\frac{\Gamma[m+1+(k+1)/2]s^{-(m+1/2+k/2)}}{\Gamma[m+1+1/2]\pi^{k/2}(s-(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}))^{-(m+1/2)}}F_{m,s^*}(\boldsymbol{\lambda}^T(\boldsymbol{y}-\boldsymbol{\mu})),$$

where $s^* = s - (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Omega} (\boldsymbol{y} - \boldsymbol{\mu})$.

Then, considering $m^* = m + 1/2$, it follows that

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2 f_{m^*,s}(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Omega}) F_{m,s^*}(\boldsymbol{\lambda}^T (\boldsymbol{y} - \boldsymbol{\mu}))$$

where $f_{m^*,s}(., \boldsymbol{\mu}, \boldsymbol{\Omega})$ is the p.d.f of the k-variate generalized Pearson type II distribution with location and scale parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$, respectivly, and F_{m,s^*} is the c.d.f. of the standard univariate generalized Pearson II.

4. MOMENTS OF SKEW-ELLIPTICAL DISTRIBUTION

4.1. Moment Generating Function

In this section we derive a general expression for the moment generating function(mgf) for skew scale mixture of normal distribution. Without loss of generality we consider here $\boldsymbol{\mu} = \boldsymbol{0}$, i.e., $\boldsymbol{Z} \sim SE_k(\boldsymbol{0}, \boldsymbol{\Omega}, \boldsymbol{\delta}; g^{(k+1)})$, where $g^{(k+1)}$ is the generator function for a $(k+1)$-variate scale mixture of normal distribution.

Considering the equation (3.1) and $\boldsymbol{t} \in \mathbb{R}^k$, we have

$$M(\boldsymbol{t}) = E[exp\{\boldsymbol{t}^T \boldsymbol{Z}\}]$$

$$= 2|\boldsymbol{\Omega}|^{-1/2} \int_{\mathbb{R}^k} \int_0^\infty (2\pi K(\eta))^{-k/2} exp\{-[(\boldsymbol{z}^T \boldsymbol{\Omega}^{-1} \boldsymbol{z})/2K(\eta) - \boldsymbol{t}^T \boldsymbol{z}]\} \Phi\left(\frac{\boldsymbol{\lambda}^T \boldsymbol{z}}{K(\eta)^{1/2}}\right) dH(\eta) d\boldsymbol{z}.$$

Using

$$\boldsymbol{z}^T \boldsymbol{\Omega}^{-1} \boldsymbol{z} - 2K(\eta) \boldsymbol{t}^T \boldsymbol{z} = (\boldsymbol{z} - K(\eta)\boldsymbol{\Omega}\boldsymbol{t})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{z} - K(\eta)\boldsymbol{\Omega}\boldsymbol{t}) - K(\eta)^2 \boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t},$$

the moment generating function $M(\boldsymbol{t})$ can be expresed as

$$2 \int_0^\infty exp\{\frac{(\boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t}) K(\eta)}{2}\} \int_{\mathbb{R}^k} |\boldsymbol{\Omega}|^{-1/2} (2\pi K(\eta))^{-k/2} exp\{-\frac{(\boldsymbol{z} - K(\eta)\boldsymbol{\Omega}\boldsymbol{t})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{z} - K(\eta)\boldsymbol{\Omega}\boldsymbol{t})}{2K(\eta)}\} \times$$

$$\times \Phi\left(\frac{\boldsymbol{\lambda}^T \boldsymbol{z}}{K(\eta)^{1/2}}\right) d\boldsymbol{z} dH(\eta)$$

$$= 2 \int_0^\infty exp\{\frac{(\boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t}) K(\eta)}{2}\} E\left[\Phi_{\boldsymbol{\Omega}}\left(\frac{\boldsymbol{\lambda}^T (\boldsymbol{z}^* + K(\eta)\boldsymbol{\Omega}\boldsymbol{t})}{K(\eta)^{1/2}}\right)\right] dH(\eta),$$

where, $\boldsymbol{z}^* = \boldsymbol{z} - K(\eta)\boldsymbol{\Omega t}$ and $\Phi_{\boldsymbol{\Omega}}$ is the c.d.f. of the $N(\boldsymbol{0}, K(\eta)\boldsymbol{\Omega})$.

Now, using proposition 4 (in Azzalini and Dalla Valle, 1996, page 719), it follows that, the mgf of the skew scale mixture of normal is

$$M_{SMN}(\boldsymbol{t}) = 2\int_0^\infty exp\{\frac{(\boldsymbol{t}^T K(\eta)^{1/2}\boldsymbol{\Omega} K(\eta)^{1/2}\boldsymbol{t})}{2}\}\Phi_{\boldsymbol{\Omega}}\left[\frac{\boldsymbol{\lambda}^T\boldsymbol{\Omega} K(\eta)\boldsymbol{t}}{(1+\boldsymbol{\lambda}^T\boldsymbol{\Omega\lambda}/K(\eta))^{1/2}}\right] dH(\eta)$$

$$= \int_0^\infty M_{SN}(K(\eta)^{1/2}\boldsymbol{t})dH(\eta), \tag{4.1}$$

where $M_{SMN}(.)$ is the moment generating function for the skew scale mixture of normal distribution and $M_{SN}(.)$ is the moment generating function for skew normal distribution.

4.2. Mean vector and covariance matrix

We have the following expressions for the mean, variance and covariance of a univariate skew normal distribution:

$$E_{SN}[Z] = \left(\frac{2}{\pi}\right)^{1/2}\delta, \qquad V_{SN}[Z] = 1 - \frac{2\delta^2}{\pi} \quad \text{and} \quad Cov_{SN}[Z_i, Z_j] = w_{ij} - \frac{2}{\pi}\delta_i\delta_j, \tag{4.2}$$

where w_{ij} is the $(i,j)^{th}$ element of the matrix $\boldsymbol{\Omega}$.

Using the moment generating function given in (4.1) and differentiating it with respect to t, we have

$$M'_{SMN}(t) = \int_0^\infty M'_{SN}(tK(\eta)^{1/2})K(\eta)^{1/2}dH(\eta).$$

Thus,

$$M'_{SMN}(0) = \int_0^\infty M'_{SN}(0)K(\eta)^{1/2}dH(\eta) = \int_0^\infty \left(\frac{2}{\pi}\right)^{1/2}\delta K(\eta)dH(\eta),$$

that is the mean of the univariate scale mixture of normal distribution is

$$E_{SMN}[Z] = \left(\frac{2}{\pi}\right)^{1/2}\delta E[K(\eta)^{1/2}].$$

The existence of this expectation depends on the existence of $E[K(\eta)^{1/2}]$, where this last expression is obtained by considering the mixture measure H. We can do the similar calculation to obtain the second derivative as

$$M''_{SMN}(t) = \int_0^\infty M''_{SN}(tK(\eta)^{1/2})K(\eta)dH(\eta).$$

Thus,

$$M''_{SMN}(0) = \int_0^\infty K(\eta)dH(\eta), \quad \text{since} \quad M''_{SN}(0) = 1.$$

It follows that the existence of the second moment for the skew scale mixture of normal distribution depends on the existence of $E[K(\eta)]$. Then, if $E[K(\eta)] < \infty$

$$V_{SMN}[Z] = E[K(\eta)] - \frac{2\delta^2}{\pi}E[K(\eta)^{1/2}].$$

It follows from (4.2) that $E[Z_i Z_j] = w_{ij}$. Using this fact and the moment generating function given in (4.1), we have

$$E_{SMN}[Z_i Z_j] = \int_0^\infty E_{SN}[Z_i Z_j]K(\eta)dH(\eta) = \int_0^\infty w_{ij}K(\eta)dH(\eta).$$

Therefore, if $E[K(\eta)] < \infty$, we have

$$Cov_{SMN}[Z_i, Z_j] = w_{ij}E[K(\eta)] - \frac{2}{\pi}\delta_i\delta_j E^2[K(\eta)^{1/2}].$$

In general, if we consider the mean vector $\boldsymbol{m}$ and the covariance matrix $\boldsymbol{M}$ for the skew-scale mixture of normal, we have

$$\boldsymbol{m} = (\frac{2}{\pi})^{1/2}E[K(\eta)^{1/2}]\boldsymbol{\delta}$$

and

$$\boldsymbol{M} = E[K(\eta)]\boldsymbol{\Omega} - \frac{2E^2[K(\eta)^{1/2}]}{\pi}\boldsymbol{\delta}^T\boldsymbol{\delta}.$$

Example 4.1: The skew t distribution has $K(\eta) = 1/\eta$ with $\eta \sim G(\nu/2; \nu/2)$. If $\nu > 1$ then $E[K(\eta)^{1/2}] < \infty$ and

$$E[K(\eta)^{1/2}] = \frac{\Gamma[\frac{\nu-1}{2}]\sqrt{\nu}}{\Gamma[\frac{\nu}{2}]\sqrt{2}}.$$

Thus,

$$E_{ST}[Z] = \frac{\delta\Gamma[\frac{\nu-1}{2}]}{\Gamma[\frac{\nu}{2}]}\sqrt{\frac{\nu}{\pi}}.$$

If $\nu > 2$ then $E[1/\eta] < \infty$ and $E[1/\eta] = \frac{\nu}{\nu-2}$, so that

$$V_{ST}[Z] = \frac{\nu}{\nu-2} - \frac{\delta^2\nu}{\pi}\left[\frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}\right]^2.$$

5. PROPERTIES OF THE SKEW ELLIPTICAL DISTRIBUTIONS

First we present two corollaries, given in Fang and Zhang(1990, page 65), in the form of propositions that will help us to prove following results concerning properties of the skew elliptical distributions. The notation $X \stackrel{d}{=} Y$ means that X and Y have the same distribution.

Proposition 5.1: $\boldsymbol{X} \sim El_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ *if and only if*

$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + R\boldsymbol{A}^T\boldsymbol{u}^{(k)},$$

where $R \geq 0$ is a random variable independent of $\boldsymbol{u}^{(k)}$, R is one to one with the generator function g, $\boldsymbol{A}$ is a k x n matrix such that $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{\Sigma}$, and $\boldsymbol{u}^{(k)}$ is a uniform distribution on the unit sphere in $\mathbb{R}^k$. This is called stochastic representation for $\boldsymbol{X}$, where R is the radial variable.

Proposition 5.2: *Assume $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + R\boldsymbol{A}^T\boldsymbol{u}^{(k)} \sim El_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$. Then*

$$(\boldsymbol{X} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \stackrel{d}{=} R^2.$$

Now we will state two similar propositions for the skew elliptical distribution.

Proposition 5.3: *If $\boldsymbol{Y} \sim SE(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; g)$ then $(\boldsymbol{Y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu})$ has the same distribution as R^2, where R is the radial variable in the stochastic representation of $\boldsymbol{X}$.*

Proof: An alternative and equivalent form to define the skew-elliptical $\boldsymbol{Y}$ is

$$\boldsymbol{Y} = \begin{cases} (\boldsymbol{X}-\boldsymbol{\mu}) & if \quad X_0 > 0 \\ -(\boldsymbol{X}-\boldsymbol{\mu}) & if \quad X_0 \leq 0 \end{cases}.$$

Using this definition it is easy to see that

$$(\boldsymbol{Y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}) = (\boldsymbol{X}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{X}-\boldsymbol{\mu}).$$

Now, using the Proposition 5.2, the proof is complete.

Proposition 5.4: *Let $\boldsymbol{C}$ be a non-singular matrix with dimension k and $\boldsymbol{b}$ is a $k \times 1$ vector and $\boldsymbol{Y}_* = \boldsymbol{b} + \boldsymbol{C}\boldsymbol{Y}$ where $\boldsymbol{Y} \sim SE_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; g)$ then,*

$$\boldsymbol{Y}_* \sim SE_k(\boldsymbol{\mu}_*, \boldsymbol{\Omega}_*, \boldsymbol{\delta}_*; g),$$

with $\boldsymbol{\mu}_* = \boldsymbol{b} + \boldsymbol{C}\boldsymbol{\mu}$, $\boldsymbol{\Omega}_* = \boldsymbol{C}^T\boldsymbol{\Omega}\boldsymbol{C}$ and $\boldsymbol{\delta}_* = \boldsymbol{C}\boldsymbol{\delta}$.

Proof: Since $\boldsymbol{C}$ is non-singular, $\boldsymbol{Y} = \boldsymbol{C}^{-1}(\boldsymbol{Y}_* - \boldsymbol{b})$. Using the jacobian method

$$f_{\boldsymbol{Y}}(y) = 2f_{g^{(k)}}(\boldsymbol{C}^{-1}(\boldsymbol{y}_* - \boldsymbol{b}))F_{g_{q(C^{-1}(y_*-\boldsymbol{b}))}}(\boldsymbol{\lambda}^T\boldsymbol{C}^{-1}(\boldsymbol{y}_* - \boldsymbol{b}))|\boldsymbol{C}|^{-1}.$$

However,

$$|\boldsymbol{C}|^{-1}f_{g^{(k)}}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{b})) = |\boldsymbol{C}|^{-1}|\boldsymbol{\Omega}|^{-1/2}g^{(k)}[(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{b})-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{b})-\boldsymbol{\mu}]$$

$$= |\boldsymbol{C}^T\boldsymbol{\Omega}\boldsymbol{C}|^{-1/2}g^{(k)}[(\boldsymbol{y}-\boldsymbol{\mu}_*)^T(\boldsymbol{C}^T\boldsymbol{\Omega}\boldsymbol{C})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_*)],$$

and the last expression is a p.d.f. of $El_k(\boldsymbol{\mu}_*, \boldsymbol{\Omega}_*; g^{(k)})$.

After some algebraic manipulations, we can see that

$$F_{g_{q(C^{-1}(y-b))}}(\boldsymbol{\lambda}^T\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{b})) = F_{g_{q_*(\boldsymbol{y}_*)}}(\boldsymbol{\lambda}_*^T(\boldsymbol{y}_* - \boldsymbol{\mu}_*),$$

where $\boldsymbol{\lambda}_*^T = \frac{\boldsymbol{\delta}_*^T\boldsymbol{\Omega}_*}{(1-\boldsymbol{\delta}_*^T\boldsymbol{\Omega}_*)^{-1}\boldsymbol{\delta}_*)^{1/2}}$ and $q_*(y_*) = (\boldsymbol{y}_* - \boldsymbol{\mu}_*)^T\boldsymbol{\Omega}_*^{-1}(\boldsymbol{y}_* - \boldsymbol{\mu}_*)$.

In the next proposition, we obtain the marginal distribution of the skew elliptical distribution. For that we consider the following partitions:

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_A \\ \boldsymbol{Y}_B \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_A \\ \boldsymbol{X}_B \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix},$$

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \boldsymbol{\delta}_A^T & \boldsymbol{\delta}_B^T \\ \boldsymbol{\delta}_A & \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\delta}_B & \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{pmatrix},$$

where $\boldsymbol{\delta}_A, \boldsymbol{Y}_A, \boldsymbol{X}_A$ and $\boldsymbol{\mu}_A$ ($\boldsymbol{\delta}_B, \boldsymbol{Y}_B, \boldsymbol{X}_B$ and $\boldsymbol{\mu}_B$) are vectors with the first m (the last $k-m$) elements of the $\boldsymbol{\delta}, \boldsymbol{Y}, \boldsymbol{X}$ and $\boldsymbol{\mu}$, respectively. Further $\boldsymbol{\Omega}_{11}$ is a $m \times m$ matrix, $\boldsymbol{\Omega}_{12} = \boldsymbol{\Omega}_{21}^T$ is a $m \times (k-m)$ matrix and $\boldsymbol{\Omega}_{22}$ is a $(k-m) \times (k-m)$ matrix.

Proposition 5.5: *If* $\boldsymbol{Y} \sim SE_k(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}; \phi)$ *then* $\boldsymbol{Y}_A \sim SE_m(\boldsymbol{\mu}_A, \boldsymbol{\Omega}_{11}, \boldsymbol{\delta}_A; \phi)$.

Proof: Using properties of the elliptical distributions, it follows that

$$\boldsymbol{X}_A \sim El(\boldsymbol{\mu}_A, \boldsymbol{\Omega}_{11}; \phi) \quad \text{and} \quad (X_0, \boldsymbol{X}_A) \sim El_{m+1}(\boldsymbol{\mu}_A^*, \boldsymbol{\Sigma}_{11}, \phi),$$

with $\boldsymbol{\mu}_A^* = \begin{pmatrix} 0 \\ \boldsymbol{\mu}_A \end{pmatrix}$ and $\boldsymbol{\Sigma}_{11} = \begin{pmatrix} 1 & \boldsymbol{\delta}_A^T \\ \boldsymbol{\delta}_A & \boldsymbol{\Omega}_{11} \end{pmatrix}$.

The proof follows using the definition of a skew-elliptical distribution and $\boldsymbol{Y}_A = [\boldsymbol{X}_A \mid X_0 > 0]$.

6. CONCLUSION

This article has presented a new class of multivariate skew-elliptical distributions which include several unimodal elliptical and spherical distributions. By introducing a skewness parameter, the new distribution brings additional flexibility of modeling skewed data. Proposed classes of distributions are useful in regression and calibration problem when the corresponding error distribution exhibit presence of skewness. The results obtained in this paper extend many properties of the elliptical and spherical distributions in a nontrivial way.

ACKNOWLEDGEMENT

The first author thanks to *Fundação de Amparo a Pesquisa do Estado de São Paulo* (FAPESP) for the financial support in her visit to the University of Connecticut. The authors wish to thank two referees for many helpful suggestions which improved the presentation of this paper.

REFERENCES

1. Andrews, D.F. and Mallows, C.L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society,* Series B, **36**, 99-102.

2. Arellano-Valle, R.B. and Bolfarine, H. (1995). On some characterizations of the t-distribution. *Statistics and Probability Letters,* **25**, 79-85.

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. Chen, M-H. and Dey, D.K. (1998). Bayesian Modeling of correlated binary response via scale mixture of multivariate normal link functions. *Sankhyā,* **60**, 322-343.

6. Choy, S.T.B. (1995). Robust Bayesian analysis using scale mixture of normals distributions, *Ph.D. Dissertation*, Department of Mathematics, Imperial College, London.

7. Fang, K.T., Kotz, S. and Ng, K-W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall.

8. Fang, K.T. and Zhang, Y.T. (1990).*Generalized Multivariate Analysis.* Springer-Verlag.

9. Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā*,**32**, 419-430.

10. Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* Chapman and Hall: New York.