An. Şt. Univ. Ovidius Constanţa Vol. 13(2), 2005, 83–96

On the multivariate Skew-Normal distribution and its scale mixtures

Raluca Vernic

Abstract

In this paper we study the multivariate skew-normal distribution and its scale mixtures, as extensions of the similar non-skewed distributions. Different parameterizations and some properties are investigated.

*

Subject Classification: 60E05.

1 Introduction

Although popular and easy to handle, the classical normal distribution is not always so adequate to model random phenomena. For example, it is well known that insurance risks have skewed distributions (see e.g. Lane, 2000), and the extensive use of the classical normal distribution to model this kind of losses was questioned.

Introduced by Azzalini (1985), the skew-normal distribution is a skewed extension of the normal distribution. Arnold and Beaver (2002) noticed that skew normal distributions may be encountered in situations in which the observations obey a normal law, but they have been truncated with respect to some hidden covariable. They exemplified this by the joint distribution of height and waist measurements of the selected individuals for elite troops. More models involving the skew-normal distribution in different scientific disciplines can be found in the discussions on Arnold and Beaver (2002).

Key Words: Multivariate Normal distributions; Multivariate Skew-Normal distributions; Scale Mixtures of Multivariate Skew-Normal distributions.

*The author acknowledges the support from the Dutch Organization for Scientific Research (NWO 048.031.2003.001).

Inferential aspects and other statistical issues of the skew-normal distributions are investigated by Azzalini and Capitanio (1999), and are illustrated by numerical examples with data from biomedical measurements on a group of athletes, on a group of individuals affected by hepatitis, and on a group of patients affected by diabetes.

In this paper we study a specific form of the multivariate skew-normal distribution and its scale mixtures. We start by recalling a first form of the density of the skew-normal distribution and some of its properties (section 2), studied by Arnold and Beaver (2002). We then introduce a more general form for the density with three different parameterizations, and we prove some properties for this general form. Some of these properties were outlined without details in Arnold and Beaver (2002). In section 3 we define a scale mixture of the multivariate skew-normal distribution and state some properties for it. Some examples are also given.

In the following, we denote an $n \times 1$ column vector by a bold-face letter and its elements by the corresponding italic with a subscript denoting the number of the element, i.e. $\mathbf{x} = (x_1, \ldots, x_n)'$. By $\mathbf{0}$ we denote the zero vector, by I_n the $n \times n$ identity matrix, and we let $\mathbf{e}' = (1, 1, ..., 1)$. Also, if $\mathbf{B}$ is a symmetric and positive definite $n \times n$ matrix, we denote by $\mathbf{B}^{1/2}$ the unique nonsingular $n \times n$ matrix that satisfies $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2}$, and by $\mathbf{B}^{-1/2}$ the inverse of $\mathbf{B}^{1/2}$. As a remark, $\mathbf{B}^{1/2}$ is also symmetric.

2 Multivariate Skew-Normal distributions

As mentioned in section 1, the univariate skew-normal distribution was introduced by Azzalini (1985) as a natural extension of the classical normal distribution to accommodate asymmetry. In conjunction with coauthors, he also extended this class to include the multivariate analog of the skew-normal. A survey of such models is provided by Arnold and Beaver (2002). More recently, Gupta et al. (2004) also studied a form of the skew-normal distribution slightly different of the general one introduced by Arnold and Beaver (2002).

The general n-variate distribution can be developed in several ways. One method consists of starting with the independent and identically distributed standard normal random variables $W_1, W_2, \ldots, W_n, U$ and considering the distribution of $\mathbf{W} = (W_1, W_2, ..., W_n)'$ given that $\lambda_0 + \lambda_1'\mathbf{W} > U$, where $\lambda_0 \in \mathbb{R}$ and $\lambda_1 \in \mathbb{R}^n$. This formulation involves a linear transformation of a hidden truncation. Denoting $A = \{\lambda_0 + \lambda_1'\mathbf{W} > U\}$ and letting $\mathbf{X}$ be the random vector with the same distribution as the conditional distribution of $\mathbf{W}$ given A, then $\mathbf{X}$ follows an *n-variate skew-normal distribution* denoted by $SN_n(\lambda_0, \lambda_1)$, with the density

$$f(\mathbf{x}) = \frac{\Phi(\lambda_0 + \lambda_1'\mathbf{x})}{\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)} \prod_{j=1}^{n} \varphi(x_j), \tag{1}$$

where φ and Φ are the standard normal $N(0,1)$ density and distribution function, respectively.

A particular case of this density was obtained by Azzalini and Dalla Valle (1996) for the choice $\lambda_0 = 0$. The resulting density takes the form

$$f(\mathbf{x}) = 2\Phi(\lambda_1'\mathbf{x}) \prod_{j=1}^{n} \varphi(x_j).$$

A useful reparameterization of (1) is obtained introducing $\delta_0 = \dfrac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}$ and $\delta_1 = \dfrac{\lambda_1}{\sqrt{1+\lambda_1'\lambda_1}}$. Then $\lambda_0 = \dfrac{\delta_0}{\sqrt{1-\delta_1'\delta_1}}$, $\lambda_1 = \dfrac{\delta_1}{\sqrt{1-\delta_1'\delta_1}}$, and the density (1) can be written as

$$f(\mathbf{x}) = \frac{1}{\Phi(\delta_0)} \Phi\left(\frac{\delta_0 + \delta_1'\mathbf{x}}{\sqrt{1-\delta_1'\delta_1}}\right) \prod_{j=1}^{n} \varphi(x_j).$$

This will also be denoted by $\widetilde{SN}_n(\delta_0, \delta_1)$. As we will see in the following, this reparameterization can simplify the writing of some formulas.

Let us now recall some properties of this skew-normal distribution (see e.g. Arnold and Beaver, 2002). Its moment generating function (mgf) is given by

$$M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{\frac{\mathbf{t}'\mathbf{t}}{2}\right\} \frac{\Phi\left(\frac{\lambda_0+\lambda_1'\mathbf{t}}{\sqrt{1+\lambda_1'\lambda_1}}\right)}{\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)} = \exp\left\{\frac{\mathbf{t}'\mathbf{t}}{2}\right\} \frac{\Phi(\delta_0 + \delta_1'\mathbf{t})}{\Phi(\delta_0)}. \tag{2}$$

It was also shown that all conditionals as well as all marginals of the density (1) are of the same type. If we partition $\mathbf{X} = \begin{pmatrix} \dot{\mathbf{X}} \\ \ddot{\mathbf{X}} \end{pmatrix}$ into two subvectors of dimensions m and $n-m$ respectively, we need to similarly partition $\lambda_1 = \begin{pmatrix} \dot{\lambda}_1 \\ \ddot{\lambda}_1 \end{pmatrix}$ and, of course, $\mathbf{x} = \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix}$. Then the conditional distribution of $\dot{\mathbf{X}}$ given $\ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ is $SN_m\left(\lambda_0 + \ddot{\lambda}_1'\ddot{\mathbf{x}}, \dot{\lambda}_1\right)$ and its unconditional distribution is $\dot{\mathbf{X}} \sim SN_m\left(\frac{\lambda_0}{\sqrt{1+\ddot{\lambda}_1'\ddot{\lambda}_1}}, \frac{\dot{\lambda}_1}{\sqrt{1+\ddot{\lambda}_1'\ddot{\lambda}_1}}\right)$.

The expected value of $\mathbf{X}$ is given by

$$\mathbb{E}X_i = \frac{\lambda_{1i}}{\sqrt{1+\lambda_1'\lambda_1}} \frac{\varphi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)}{\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)} = \delta_{1i}\frac{\varphi(\delta_0)}{\Phi(\delta_0)}. \tag{3}$$

Considerable simplification occurs when $\lambda_0 = 0$, case in which $\frac{\varphi(\delta_0)}{\Phi(\delta_0)} = \sqrt{\frac{2}{\pi}}$.

A more general form of skew-normal distribution is obtained by introducing a location parameter μ and scale parameter $\mathbf{\Sigma}$ in model (1). Here $\mu \in \mathbb{R}^n$ and $\mathbf{\Sigma}$ is an $n \times n$ symmetric and positive definite matrix with $\mathbf{\Sigma} = \mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2}$, as stated in the introduction. We define

$$\mathbf{X} = \mu + \mathbf{\Sigma}^{1/2}\mathbf{V},$$

where $\mathbf{V}$ has density (1). Then, from Arnold and Beaver (2002), the density of $\mathbf{X}$ is of the form

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) \quad &\propto \quad \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right)\Phi\left(\lambda_0 + \lambda_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)\right) = \qquad (4) \\ & \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right)\Phi\left(\frac{\delta_0 + \delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)}{\sqrt{1-\delta_1'\delta_1}}\right), \end{aligned}$$

where "$\propto$" means "proportional with". We denote this by $\mathbf{X} \sim SN_n(\mu, \mathbf{\Sigma}; \lambda_0, \lambda_1)$ or alternatively by $\widetilde{SN}_n(\mu, \mathbf{\Sigma}; \delta_0, \delta_1)$.

Introducing $\gamma = \mathbf{\Sigma}^{1/2}\delta_1$, hence $\delta_1 = \mathbf{\Sigma}^{-1/2}\gamma$, we obtain a second reparameterization, denoted by $SN_n^*(\mu, \mathbf{\Sigma}; \delta_0, \gamma)$. This second reparameterization derives from the skew-elliptical distributions (see Branco and Dey, 2001), of which the skew-normal distribution is a particular case, and it is useful for the presentation of some properties.

We will now give the exact form of the density $f_{\mathbf{X}}$ for all three parameterizations.

Proposition 1 *The exact form of the density of the above skew-normal distributed random variable is*

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \left[\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)\right]^{-1}\varphi_n(\mathbf{x};\mu,\mathbf{\Sigma})\,\Phi\left(\lambda_0 + \lambda_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)\right) \qquad (5) \\ &= \frac{1}{\Phi(\delta_0)}\varphi_n(\mathbf{x};\mu,\mathbf{\Sigma})\,\Phi\left(\frac{\delta_0 + \delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)}{\sqrt{1-\delta_1'\delta_1}}\right) = \qquad (6) \\ &= \frac{1}{\Phi(\delta_0)}\varphi_n(\mathbf{x};\mu,\mathbf{\Sigma})\,\Phi\left(\frac{\delta_0 + \gamma'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)}{\sqrt{1-\gamma'\mathbf{\Sigma}^{-1}\gamma}}\right), \qquad (7) \end{aligned}$$

where $\varphi_n(.;\mu,\mathbf{\Sigma})$ *is the n-dimensional normal* $N_n(\mu,\mathbf{\Sigma})$ *density.*

Proof. From Proposition 4 in Azzalini and Dalla Valle (1996) we know that if $a \in \mathbb{R}, \mathbf{b} \in \mathbb{R}^n$ and $\mathbf{Y}$ is an $n \times 1$ vector of independent standard normal random variables, then

$$\mathbb{E}\left[\Phi\left(a+\mathbf{b}'\mathbf{Y}\right)\right] = \Phi\left(\frac{a}{\sqrt{1+\mathbf{b}'\mathbf{b}}}\right).$$

Applying this result to the density condition $\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\,d\mathbf{x} = 1$, we have for example for the form (4) of the density $f_{\mathbf{X}}$,

$$1 = c\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right)\Phi\left(\lambda_0+\lambda_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)\right)d\mathbf{x} =$$
$$= c\sqrt{(2\pi)^n|\mathbf{\Sigma}|}\int_{-\infty}^{\infty}\varphi_n(\mathbf{y};\mathbf{0},I_n)\,\Phi(\lambda_0+\lambda_1'\mathbf{y})\,d\mathbf{y} = c\sqrt{(2\pi)^n|\mathbf{\Sigma}|}\,\mathbb{E}\left[\Phi\left(\lambda_0+\lambda_1'\mathbf{Y}\right)\right] =$$
$$= c\sqrt{(2\pi)^n|\mathbf{\Sigma}|}\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right),$$

hence

$$c = \frac{1}{\sqrt{(2\pi)^n|\mathbf{\Sigma}|}\Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}\right)}.$$

Introducing this in (4) we obtain (5). The other two forms of $f_{\mathbf{X}}$, (6) and (7), result in a similar way.

It is now easy to see that if we take $\lambda_1 = \delta_1 = \gamma = \mathbf{0}$ we obtain the well-known density of the multivariate normal distribution $N_n(\mu,\mathbf{\Sigma})$.

We will now present some important properties of this general form of skew-normal distribution. Some of these properties were just stated by Arnold and Beaver (2002), without details or proofs.

First, the mgf of $\mathbf{X}$ follows easily from the definition of $\mathbf{X}$ and from (2), as

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= e^{\mathbf{t}'\mu}M_{\mathbf{V}}\left(\mathbf{\Sigma}^{1/2}\mathbf{t}\right) = \exp\left\{\mathbf{t}'\mu+\frac{\mathbf{t}'\mathbf{\Sigma}\mathbf{t}}{2}\right\}\frac{\Phi\left(\delta_0+\delta_1'\mathbf{\Sigma}^{1/2}\mathbf{t}\right)}{\Phi(\delta_0)} = \\ &= \exp\left\{\mathbf{t}'\mu+\frac{\mathbf{t}'\mathbf{\Sigma}\mathbf{t}}{2}\right\}\frac{\Phi(\delta_0+\gamma'\mathbf{t})}{\Phi(\delta_0)}. \end{aligned} \tag{8}$$

The property of having marginals and conditionals of the same type continues to hold. In order to prove this, we partition as before $\mathbf{X} = \begin{pmatrix} \dot{\mathbf{X}} \\ \ddot{\mathbf{X}} \end{pmatrix}$

into two subvectors of dimensions m and $n-m$ respectively, and similarly $\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}$, $\mu = \begin{pmatrix} \dot{\mu} \\ \ddot{\mu} \end{pmatrix}$, $\delta_1 = \begin{pmatrix} \dot{\delta}_1 \\ \ddot{\delta}_1 \end{pmatrix}$, $\gamma = \begin{pmatrix} \dot{\gamma} \\ \ddot{\gamma} \end{pmatrix}$ and $\mathbf{t} = \begin{pmatrix} \dot{\mathbf{t}} \\ \ddot{\mathbf{t}} \end{pmatrix}$. We have the following proposition.

Proposition 2 *With the above notations, if $\mathbf{X} \sim SN_n^*(\mu, \mathbf{\Sigma}; \delta_0, \gamma)$, then*
(i) $\dot{\mathbf{X}} \sim SN_m^*(\dot{\mu}, \mathbf{\Sigma}_{11}; \delta_0, \dot{\gamma})$;
(ii) The conditional distribution of $\dot{\mathbf{X}}$ given $\ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ is SN_m $\left(\dot{\mu}(\ddot{\mathbf{x}}), \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}; l_0, \mathbf{l}_1\right)$, where

$$\dot{\mu}(\ddot{\mathbf{x}}) = \dot{\mu} + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\ddot{\mathbf{x}} - \ddot{\mu}),$$

$$l_0 = \frac{\delta_0 + \ddot{\gamma}'\mathbf{\Sigma}_{22}^{-1}(\ddot{\mathbf{x}} - \ddot{\mu})}{\sqrt{1 - \gamma'\mathbf{\Sigma}^{-1}\gamma}},\ \mathbf{l}_1 = \frac{\left(\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1/2}\left(\dot{\gamma} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\ddot{\gamma}\right)}{\sqrt{1 - \gamma'\mathbf{\Sigma}^{-1}\gamma}}.$$

Proof. (i) We will use the mgf. Taking $\ddot{\mathbf{t}} = \mathbf{0}$ in (8) gives

$$M_{\mathbf{X}}\begin{pmatrix} \dot{\mathbf{t}} \\ \mathbf{0} \end{pmatrix} = \exp\left\{\dot{\mathbf{t}}'\dot{\mu} + \frac{\dot{\mathbf{t}}'\mathbf{\Sigma}_{11}\dot{\mathbf{t}}}{2}\right\}\frac{\Phi\left(\delta_0 + \dot{\gamma}'\dot{\mathbf{t}}\right)}{\Phi(\delta_0)}.$$

We then have $\dot{\mathbf{X}} \sim SN_m^*(\dot{\mu}, \mathbf{\Sigma}_{11}; \delta_0, \dot{\gamma})$.
(ii) Arnold and Beaver (2002) noticed that the conditional density of $\dot{\mathbf{X}}$ given $\ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ satisfies

$$f_{\dot{\mathbf{X}}}(\dot{\mathbf{x}}\,|\ddot{\mathbf{x}}) \propto \exp\left\{-\frac{1}{2}(\dot{\mathbf{x}} - \dot{\mu}(\ddot{\mathbf{x}}))'\left(\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1}(\dot{\mathbf{x}} - \dot{\mu}(\ddot{\mathbf{x}}))\right\}$$
$$\Phi\left(\frac{\delta_0 + \delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \mu)}{\sqrt{1 - \delta_1'\delta_1}}\right).$$

We will now prove that this is a general skew-normal density with the location and scale parameters equal to $\dot{\mu}(\ddot{\mathbf{x}})$ and $\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$, respectively. Based on the expression between the brackets of Φ, we need to find the form of the two other parameters. For this purpose, we consider the partition

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix}.$$

Then

$$\begin{aligned} \delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \mu) = \gamma'\mathbf{\Sigma}^{-1}(\mathbf{x} - \mu) &= (\dot{\gamma}'\mathbf{T}_{11} + \ddot{\gamma}'\mathbf{T}_{21})(\dot{\mathbf{x}} - \dot{\mu}) + \\ &\quad (\dot{\gamma}'\mathbf{T}_{12} + \ddot{\gamma}'\mathbf{T}_{22})(\ddot{\mathbf{x}} - \ddot{\mu}) = \\ &= (\dot{\gamma}'\mathbf{T}_{11} + \ddot{\gamma}'\mathbf{T}_{21})(\dot{\mathbf{x}} - \dot{\mu}(\ddot{\mathbf{x}})) + \\ &+ \left[(\dot{\gamma}'\mathbf{T}_{11} + \ddot{\gamma}'\mathbf{T}_{21})\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1} + (\dot{\gamma}'\mathbf{T}_{12} + \ddot{\gamma}'\mathbf{T}_{22})\right](\ddot{\mathbf{x}} - \ddot{\mu}). \end{aligned}$$

It can be shown that

$$\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1} = -\mathbf{T}_{11}^{-1}\mathbf{T}_{12},\ \mathbf{\Sigma}_{22}^{-1} = \mathbf{T}_{22} - \mathbf{T}_{21}\mathbf{T}_{11}^{-1}\mathbf{T}_{12},$$
$$\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} = \mathbf{T}_{11}^{-1},$$

so that

$$\begin{aligned}\delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}\text{-}\mu) &= \left(\dot{\gamma}'\mathbf{T}_{11}+\ddot{\gamma}'\mathbf{T}_{21}\right)\mathbf{T}_{11}^{-1}\mathbf{T}_{11}\left(\dot{\mathbf{x}}\text{-}\dot{\mu}\left(\ddot{\mathbf{x}}\right)\right)+\ddot{\gamma}'\left(\mathbf{T}_{22}\text{-}\mathbf{T}_{21}\mathbf{T}_{11}^{-1}\mathbf{T}_{12}\right)\left(\ddot{\mathbf{x}}\text{-}\ddot{\mu}\right)= \\ &= \left(\dot{\gamma}'+\ddot{\gamma}'\mathbf{T}_{21}\mathbf{T}_{11}^{-1}\right)\left(\mathbf{\Sigma}_{11}\text{-}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1}\left(\dot{\mathbf{x}}\text{-}\dot{\mu}\left(\ddot{\mathbf{x}}\right)\right)+\ddot{\gamma}'\mathbf{\Sigma}_{22}^{-1}\left(\ddot{\mathbf{x}}\text{-}\ddot{\mu}\right).\end{aligned}$$

Hence,

$$\Phi\left(\frac{\delta_0+\delta_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)}{\sqrt{1-\delta_1'\delta_1}}\right) =$$
$$\Phi\left(\frac{\delta_0+\ddot{\gamma}'\mathbf{\Sigma}_{22}^{-1}\left(\ddot{\mathbf{x}}-\ddot{\mu}\right)}{\sqrt{1-\delta_1'\delta_1}}+\frac{\left(\dot{\gamma}'-\ddot{\gamma}'\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1}}{\sqrt{1-\delta_1'\delta_1}}\left(\dot{\mathbf{x}}-\dot{\mu}\left(\ddot{\mathbf{x}}\right)\right)\right).$$

From this and (5), it is easy to find the expressions of the last two parameters l_0 and $\mathbf{l}_1$ given in (ii).

Remark 1 *With the second parameterizations, (i) from Proposition 2 can also be written as* $\dot{\mathbf{X}} \sim \widetilde{SN}_m\left(\dot{\mu},\mathbf{\Sigma}_{11};\delta_0,\delta_1^{(m)}\right)$, *where* $\delta_1^{(m)} = \mathbf{\Sigma}_{11}^{-1/2}\dot{\gamma}$ *and generally* $\delta_1^{(m)} \neq \dot{\delta}_1$. *To be more specific, if we accordingly partition* $\mathbf{\Sigma}^{1/2}=\begin{pmatrix}\mathbf{\Omega}_{11} & \mathbf{\Omega}_{12}\\ \mathbf{\Omega}_{21} & \mathbf{\Omega}_{22}\end{pmatrix}$, *we notice that* $\dot{\gamma} = \mathbf{\Omega}_{11}\dot{\delta}_1+\mathbf{\Omega}_{12}\ddot{\delta}_1$, *so that* $\delta_1^{(m)} = \mathbf{\Sigma}_{11}^{-1/2}\left(\mathbf{\Omega}_{11}\dot{\delta}_1+\mathbf{\Omega}_{12}\ddot{\delta}_1\right)$.

Corollary 1 *In particular, the marginal distributions of* $\mathbf{X} \sim SN_n^*(\mu,\mathbf{\Sigma};\delta_0,\gamma)$ *are given by* $X_j \sim SN_1^*\left(\mu_j,\sigma_j^2;\delta_0,\gamma_j\right)$, *where* $\sigma_j^2 = \sigma_{jj}$. *We also have that*

$$\mathbb{E}X_j = \mu_j + \gamma_j\frac{\varphi(\delta_0)}{\Phi(\delta_0)}. \tag{9}$$

Proof. The first affirmation of the corollary is immediate from (i) in Proposition 2.

Using now the marginal distribution, it is easy to determine the expected value of X_j by writing $X_j = \mu_j + \sigma_j V_j$, where $V_j \sim \widetilde{SN}_1\left(\delta_0,\sigma_j^{-1}\gamma_j\right)$. Applying also (3), we get

$$\mathbb{E}X_j = \mu_j + \sigma_j\sigma_j^{-1}\gamma_j\frac{\varphi(\delta_0)}{\Phi(\delta_0)} = \mu_j + \gamma_j\frac{\varphi(\delta_0)}{\Phi(\delta_0)}.$$

The following corollary is an immediate consequence of (ii) in Proposition 2.

Corollary 2 *For the particular case* $n = 2$ *and* $m = 1$, *the conditional distribution of* X_1 *given* $X_2 = x_2$ *is*

$$SN_1\left(\mu_1+\frac{\sigma_{12}}{\sigma_2^2}\left(x_2\text{-}\mu_2\right), \sigma_1^2\text{-}\frac{\sigma_{12}^2}{\sigma_2^2}; \lambda_0+\frac{\gamma_2}{\sigma_2^2}\frac{(x_2\text{-}\mu_2)}{\sqrt{1\text{-}\gamma'\mathbf{\Sigma}^{-1}\gamma}}, \frac{\gamma_1\text{-}\sigma_{12}\sigma_2^{-2}\gamma_2}{\sqrt{\left(\sigma_1^2\text{-}\sigma_{12}^2\sigma_2^{-2}\right)\left(1\text{-}\gamma'\mathbf{\Sigma}^{-1}\gamma\right)}}\right).$$

Another important property of the skew-normal is that any linear combination of skew-normal distributed random vectors is still skew-normal.

Proposition 3 *Let* $\mathbf{b}$ *be an* $n \times 1$ *real vector and* $\mathbf{C}$ *an* $m \times n$ *matrix of rang* m, *where* $m \leq n$. *We define* $\mathbf{Y} = \mathbf{b} + \mathbf{CX}$, *where* $\mathbf{X} \sim SN_n^*\left(\mu, \mathbf{\Sigma}; \delta_0, \gamma\right)$. *Then* $\mathbf{Y} \sim SN_n^*\left(\mathbf{b} + \mathbf{C}\mu, \mathbf{C\Sigma C}'; \delta_0, \mathbf{C}\gamma\right)$.

Proof. We will use the mgf function. From (8),

$$M_{\mathbf{Y}}(\mathbf{t}) = e^{\mathbf{t}'\mathbf{b}} M_{\mathbf{X}}\left(\mathbf{C}'\mathbf{t}\right) = \exp\left\{\mathbf{t}'\left(\mathbf{b} + \mathbf{C}\mu\right) + \frac{\mathbf{t}'\mathbf{C\Sigma C}'\mathbf{t}}{2}\right\}\frac{\Phi\left(\delta_0 + \gamma'\mathbf{C}'\mathbf{t}\right)}{\Phi\left(\delta_0\right)}.$$

Since $\mathbf{C\Sigma C}'$ is also a positive definite matrix, it follows that
$\mathbf{Y} \sim SN_n^*\left(\mathbf{b} + \mathbf{C}\mu, \mathbf{C\Sigma C}'; \delta_0, \mathbf{C}\gamma\right)$.

Remark 2 *With the second parameterization, the result in Proposition 3 can also be written as*
$\mathbf{Y} \sim \widetilde{SN}_n\left(\mathbf{b} + \mathbf{C}\mu, \mathbf{C\Sigma C}'; \delta_0, \delta_1\right)$, *while with the third parameterization and* $\mathbf{C\Sigma}^{1/2}\delta_1 = \mathbf{C}\gamma$ *it becomes* $\mathbf{Y} \sim SN_n\left(\mathbf{b} + \mathbf{C}\mu, \mathbf{C\Sigma C}'; \delta_0, \mathbf{C}\gamma\right)$.

Corollary 3 *If* $\mathbf{X} \sim SN_n^*\left(\mu, \mathbf{\Sigma}; \delta_0, \gamma\right)$, *then* $S = \sum_{i=1}^n X_i \sim SN_1^*\left(\mu_S, \sigma_S^2; \delta_0, \gamma_S\right)$, *where*
$\mu_S = \mathbf{e}'\mu = \sum_{j=1}^n \mu_j$, $\sigma_S^2 = \mathbf{e}'\mathbf{\Sigma e} = \sum_{i,j=1}^n \sigma_{ij}$, $\gamma_S = \mathbf{e}'\gamma = \sum_{j=1}^n \gamma_j$.

Proof. We apply the linear property from Proposition 3 by taking $\mathbf{b} = \mathbf{0}$ and $\mathbf{C} = \begin{pmatrix} 1 & 1 & ... & 1 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{pmatrix}$, and also Corollary 1 to obtain the marginal distribution of S.

We will now briefly recall the *general p-multivariate skew normal distribution* (GMSN), introduced by Gupta et al. (2004) as a generalization of the form of the multivariate skew normal distribution studied in detail in their paper. Although they didn't make a detailed study of this GMSN distribution, Gupta et al. (2004) defined it in order to have a closed family, in the sense that it contains its marginal and conditional distributions. Its density has the form

$$f_{p,q}\left(\mathbf{y};\mu,\mathbf{\Sigma},\mathbf{D},\nu,\mathbf{\Delta}\right)=\Phi_q^{-1}\left(\mathbf{D}\mu;\nu,\mathbf{\Delta}+\mathbf{D\Sigma D}'\right)\varphi_p\left(\mathbf{y};\mu,\mathbf{\Sigma}\right)\Phi_q\left(\mathbf{D}\mathbf{y};\nu,\mathbf{\Delta}\right),$$

where $\mu,\mathbf{y}\in\mathbb{R}^p,\nu\in\mathbb{R}^q,\mathbf{\Sigma}\,(p\times p)$ and $\mathbf{\Delta}\,(q\times q)$ are two covariance matrices, $\mathbf{D}\,(q\times p)$ is an arbitrary matrix and $\Phi_q\,(.;\nu,\mathbf{\Delta})$ denotes the distribution function of the q-dimensional normal distribution $N_q\,(\nu,\mathbf{\Delta})$. We notice that the multivariate skew-normal distribution studied in this paper can be obtained as a particular case of the GMSN taking $q=1$.

3 Scale Mixtures of Multivariate Skew-Normal distributions

Branco and Dey (2001) defined the scale mixture of a skew-normal distribution starting from the skew-elliptical distributions. In the following, we will define it directly from the skew-normal distribution, and based on this definition we will deduce some of its properties.

Let Θ be a positive random variable with distribution function H, and let $K:(0,\infty)\to(0,\infty)$ be a weight function. Then we define the *scale H-mixture of the multivariate skew-normal distribution* as the distribution of an n-dimensional random vector $\mathbf{X}$ that, given $\Theta=\theta$, follows a multivariate skew-normal $SN_n\,(\mu,K\,(\theta)\,\mathbf{\Sigma};\lambda_0,\lambda_1)$ distribution. We denote this by $\mathbf{X}\sim SN_n-H\,(\mu,\mathbf{\Sigma};\lambda_0,\lambda_1)$ or alternatively by $\widetilde{SN}_n-H\,(\mu,\mathbf{\Sigma};\delta_0,\delta_1)$, where, as in section 1, $\delta_0=\dfrac{\lambda_0}{\sqrt{1+\lambda_1'\lambda_1}}$ and $\delta_1=\dfrac{\lambda_1}{\sqrt{1+\lambda_1'\lambda_1}}$.

We notice that if the distribution of $\mathbf{X}$ given $\Theta=\theta$ is, with the second parameterization, $\widetilde{SN}_n\,(\mu,K\,(\theta)\,\mathbf{\Sigma};\delta_0,\delta_1)$, with the third parameterization it will be $SN_n^*\left(\mu,K\,(\theta)\,\mathbf{\Sigma};\delta_0,\sqrt{K\,(\theta)}\gamma\right)$, where $\gamma=\mathbf{\Sigma}^{1/2}\delta_1$. Hence, we will also use the notation $\mathbf{X}\sim SN_n^*-H\,(\mu,\mathbf{\Sigma};\delta_0,\gamma)$, but keep in mind that this means that the distribution of $\mathbf{X}$ given $\Theta=\theta$ is $SN_n^*\left(\mu,K\,(\theta)\,\mathbf{\Sigma};\delta_0,\sqrt{K\,(\theta)}\gamma\right)$ and *not* $SN_n^*\,(\mu,K\,(\theta)\,\mathbf{\Sigma};\delta_0,\gamma)$.

The density of this distribution is then given by

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X}}(\mathbf{x}\,|\Theta = \theta)\,dH(\theta) =$$

$$\overset{(5)}{=} [\Phi(\delta_0)]^{-1} \int_0^\infty \varphi_n(\mathbf{x};\mu,K(\theta)\boldsymbol{\Sigma})\,\Phi\left(\lambda_0 + \lambda_1'(K(\theta)\boldsymbol{\Sigma})^{-1/2}(\mathbf{x}-\mu)\right) dH(\theta) =$$

$$\overset{(6)}{=} [\Phi(\delta_0)]^{-1} \int_0^\infty \varphi_n(\mathbf{x};\mu,K(\theta)\boldsymbol{\Sigma})\,\Phi\left(\frac{\delta_0 + \delta_1'(K(\theta)\boldsymbol{\Sigma})^{-1/2}(\mathbf{x}-\mu)}{\sqrt{1-\delta_1'\delta_1}}\right) dH(\theta) =$$

$$\overset{(7)}{=} [\Phi(\delta_0)]^{-1} \int_0^\infty \varphi_n(\mathbf{x};\mu,K(\theta)\boldsymbol{\Sigma})\,\Phi\left(\frac{\delta_0 + \sqrt{K(\theta)}\gamma'(K(\theta)\boldsymbol{\Sigma})^{-1}(\mathbf{x}-\mu)}{\sqrt{1-\gamma'\boldsymbol{\Sigma}^{-1}\gamma}}\right) dH(\theta).$$

We will now present some properties of this scale mixture of a multivariate skew normal distribution.

Its mgf is given by

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E}\left[\mathbb{E}\left(e^{\mathbf{t}'\mathbf{X}}\,|\Theta=\theta\right)\right] \overset{(8)}{=} \\ &= \mathbb{E}\left[\exp\left\{\mathbf{t}'\mu + \frac{\mathbf{t}'K(\Theta)\boldsymbol{\Sigma}\mathbf{t}}{2}\right\}\frac{\Phi\left(\delta_0 + \delta_1'\sqrt{K(\Theta)}\boldsymbol{\Sigma}^{1/2}\mathbf{t}\right)}{\Phi(\delta_0)}\right] = \\ &= \frac{\exp\{\mathbf{t}'\mu\}}{\Phi(\delta_0)}\mathbb{E}\left[\exp\left\{\frac{K(\Theta)}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right\}\Phi\left(\delta_0 + \sqrt{K(\Theta)}\gamma'\mathbf{t}\right)\right]. \qquad (10) \end{aligned}$$

We will now prove that the marginals and the linear combinations of these distributions are of the same type, while their conditionals are not. For this purpose, just as in the previous section, we partition $\mathbf{X} = \begin{pmatrix} \dot{\mathbf{X}} \\ \ddot{\mathbf{X}} \end{pmatrix}$ into two subvectors of dimensions m and $n-m$ respectively, and similarly $\boldsymbol{\Sigma}$, μ, δ_1, γ and $\mathbf{t}$. The following proposition holds.

Proposition 4 *With the above notations, if* $\mathbf{X} \sim SN_n^* - H(\mu,\boldsymbol{\Sigma};\delta_0,\gamma)$, *then*
(i) $\dot{\mathbf{X}} \sim SN_m^* - H(\dot{\mu},\boldsymbol{\Sigma}_{11};\delta_0,\dot{\gamma})$;
(ii) $X_j \sim SN_1^* - H\left(\mu_j,\sigma_j^2;\delta_0,\gamma_j\right)$ *or, equivalently,* $X_j \sim \widetilde{SN}_1 - H\left(\mu_j,\sigma_j^2;\delta_0,\gamma_j\sigma_j^{-1}\right)$;
(iii) Let $\mathbf{b} \in \mathbb{R}^n$ *and* $\mathbf{C}$ *be an* $n \times n$ *non-singular matrix. If we define* $\mathbf{Y} = \mathbf{b} + \mathbf{C}\mathbf{X}$, *then*
$\mathbf{Y} \sim SN_n^* - H\left(\mathbf{b}+\mathbf{C}\mu,\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}';\delta_0,\mathbf{C}\gamma\right)$;
(iv) $S = \sum_{i=1}^n X_i \sim SN_1^* - H\left(\mu_S,\sigma_S^2;\delta_0,\gamma_S\right)$, *where as before,* $\mu_S = \mathbf{e}'\mu$, $\sigma_S^2 = \mathbf{e}'\boldsymbol{\Sigma}\mathbf{e}$, $\gamma_S = \mathbf{e}'\gamma$;
(v) The conditional distribution of $\dot{\mathbf{X}}$ *given* $\ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ *and* $\Theta = \theta$ *is*

$SN_m\left(\dot{\mu}(\ddot{\mathbf{x}}), K(\theta)\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right); l_0(\theta), \mathbf{l}_1\right)$, *where*

$$\dot{\mu}(\ddot{\mathbf{x}}) = \dot{\mu} + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\ddot{\mathbf{x}} - \ddot{\mu}),$$

$$l_0(\theta) = \lambda_0 + \frac{\ddot{\gamma}'\mathbf{\Sigma}_{22}^{-1}(\ddot{\mathbf{x}} - \ddot{\mu})}{\sqrt{K(\theta)(1-\gamma'\mathbf{\Sigma}^{-1}\gamma)}}, \ \mathbf{l}_1 = \frac{\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1/2}\left(\dot{\gamma} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\ddot{\gamma}\right)}{\sqrt{1-\gamma'\mathbf{\Sigma}^{-1}\gamma}}.$$

Proof. (i) Results immediately from the mgf (10) taking $\ddot{\mathbf{t}} = \mathbf{0}$.
(ii) Is a consequence of (i).
(iii) From Proposition 3, the distribution of $\mathbf{Y}$ given $\Theta = \theta$ is $SN_n^*\left(\mathbf{b} + \mathbf{C}\mu, K(\theta)\mathbf{C}\mathbf{\Sigma}\mathbf{C}'; \delta_0, \sqrt{K(\theta)}\mathbf{C}\gamma\right)$, and hence the result.
(iv) Results from Corollary 3, knowing that the distribution of S given $\Theta = \theta$ is $SN_1^*\left(\mu_S, K(\theta)\sigma_S^2; \delta_0, \sqrt{K(\theta)}\gamma_S\right)$.
(v) Results from (ii) in Proposition 2, where the parameters are

$$\dot{\mu} + K(\theta)\mathbf{\Sigma}_{12}\left(K(\theta)\mathbf{\Sigma}_{22}\right)^{-1}(\ddot{\mathbf{x}} - \ddot{\mu}) = \dot{\mu}(\ddot{\mathbf{x}}),$$

$$K(\theta)\mathbf{\Sigma}_{11} - K(\theta)\mathbf{\Sigma}_{12}\left(K(\theta)\mathbf{\Sigma}_{22}\right)^{-1}K(\theta)\mathbf{\Sigma}_{21} = K(\theta)\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right),$$

$$\lambda_0 + \frac{\sqrt{K(\theta)}\ddot{\gamma}'\left(K(\theta)\mathbf{\Sigma}_{22}\right)^{-1}(\ddot{\mathbf{x}} - \ddot{\mu})}{\sqrt{1-\gamma'\mathbf{\Sigma}^{-1}\gamma}} = l_0(\theta),$$

$$\frac{\left[\sqrt{K(\theta)}\right]^{-1}\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)^{-1/2}\sqrt{K(\theta)}\left(\dot{\gamma} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\ddot{\gamma}\right)}{\sqrt{1-\gamma'\mathbf{\Sigma}^{-1}\gamma}} = \mathbf{l}_1.$$

Remark 3 *From (v) in Proposition 4 we see that because the parameter $l_0(\theta)$ depends on θ, the conditional distribution of $\dot{\mathbf{X}}$ given $\ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ is not a scale mixture of a skew-normal distribution anymore.*

Examples of scale mixtures of skew-normal distributions

1. Finite scale mixture of skew-normal. This distribution can be obtained by taking Θ to be a finite discrete random variable given as $\Theta\begin{pmatrix}\theta_1 & \dots & \theta_m \\ p_1 & \dots & p_m\end{pmatrix}$, with $0 \le p_i \le 1$ and $\sum_{i=1}^m p_i = 1$. The density of the finite scale mixture of skew-normal is then given by

$$f_{\mathbf{X}}(\mathbf{x}) = \left[\Phi(\delta_0)\right]^{-1}\sum_{i=1}^{m} p_i\varphi_n\left(\mathbf{x}; \mu, K(\theta_i)\mathbf{\Sigma}\right)\Phi\left(\lambda_0 + \frac{1}{\sqrt{K(\theta_i)}}\lambda_1'\mathbf{\Sigma}^{-1/2}(\mathbf{x}-\mu)\right).$$

In the particular case when Θ is degenerate in θ_0 and $K(\theta_0) = 1$, we recover the skew-normal distribution.

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

$$f_\Theta(\theta) = 8\sum_{k=1}^{\infty} (-1)^{k+1} k^2\theta \exp\left\{-2k^2\theta^2\right\}.$$

However, this density is not computational attractive, but Chen and Dey (1998) overcome this problem by finding a t-approximation to the logistic distribution.

3. Skew Stable distribution. This distribution results by taking $K(\theta) = 2\theta$, where Θ follows a positive stable distribution $S^p(\alpha, 1)$, with density given by

$$f_\Theta(\theta\,|\alpha, 1) = \frac{\alpha}{1-\alpha}\theta^{-\frac{1}{1-\alpha}} \int_0^1 s(u) \exp\left\{-s(u)\,\theta^{-\frac{\alpha}{1-\alpha}}\right\} du,$$

for $0 < \alpha < 1$, with

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

We notice that the skew-normal distribution can also be obtained from the skew-stable by taking $\alpha \to 1$.

4. Skew Exponential Power distribution. A skew exponential power distribution can be obtained as a scale mixture of skew normal by choosing $K(\theta) = \dfrac{1}{2c_0\theta}$ and $f_\Theta(\theta) = \dfrac{1}{\theta^{(n+1)/2}} f_\Theta(\theta\,|\alpha, 1)$,where $f_\Theta(.\,|\alpha, 1)$ is the one given above, $c_0 = \dfrac{\Gamma[3/(2\alpha)]}{\Gamma[1/(2\alpha)]}$ and $\dfrac{1}{2} < \alpha < 1$. Here α is called the kurtosis parameter. Further references on the symmetric exponential power family of distributions can be found in West (1987) and Choy (1995).

5. Skew *t* distribution. This distribution can be obtained taking $K(\theta) = \dfrac{1}{\theta}$ and $\Theta \sim Gamma\left(\dfrac{\nu}{2}, \dfrac{\nu}{2}\right)$. As two of its particular cases, we have the skew Cauchy distribution for $\nu = 1$, and again the skew-normal distribution as the limiting case when $\nu \to \infty$.

We can also consider the generalized version of Student's t distribution by taking $\Theta \sim Gamma\left(\frac{\nu}{2}, \frac{\tau}{2}\right)$, $\nu, \theta > 0$, with the density given by

$$f_{\Theta}(\theta) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\tau}{2}\right)^{\nu/2} \theta^{\nu/2-1} \exp\left\{-\frac{\tau}{2}\theta\right\}.$$

Branco and Dey (2001) showed that the density of the multivariate skew generalized t distribution is given for $\lambda_0 = 0$ by

$$f_{\mathbf{X}}(\mathbf{x}) = 2 f_{\nu,\tau}(\mathbf{x}; \mu, \mathbf{\Sigma}) F_{\nu^*,\tau^*}\left(\lambda_1'(\mathbf{x}-\mu)\right), \tag{11}$$

where $f_{\nu,\tau}(.; \mu, \mathbf{\Sigma})$ is the density of an n-dimensional generalized Student's t distribution with location parameter μ and scale $\mathbf{\Sigma}$, while $F_{\nu^*,\tau^*}(.)$ is the distribution function of an univariate standard generalized t distribution with $\nu^* = \nu + n$ and $\tau^* = \tau + (\mathbf{x}-\mu)' \mathbf{\Sigma}^{-1} (\mathbf{x}-\mu)$. Formula (11) is in fact another way to define a skew distribution starting from its symmetric form, see e.g. Arnold and Beaver (2002).

References

[1] Arnold, B.C. and Beaver, R.J. (2002) - *Skewed multivariate models related to hidden truncation and/or selective reporting*, Sociedad de Estadistica e Investigacion Operativa Test 11, no.1, 7-54.

[2] Azzalini, A. (1985) - *A class of distributions which includes the normal ones*, Scandinavian Journal of Statistics 12, 171-178.

[3] Azzalini, A. and Capitanio, A. (1999) - *Statistical applications of the multivariate skew-normal distribution*, Journal of the Royal Statistical Society, Series B 61, no.3, 579-602.

[4] Azzalini, A. and Dalla Valle, A. (1996) - *The multivariate skew-normal distribution*, Biometrika 83, 715-726.

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

[7] Choy, S.T.B. (1995) - *Robust Bayesian Analysis Using Scale Mixture of Normals Distributions*, Ph.D. dissertation, Department of Mathematics, Imperial College, London.

[8] Gupta, A., Gonzalez-Farias, G. and Dominguez-Molina, J.A. (2004) - *A multivariate skew normal distribution*, Journal of Multivariate Analysis 89, 181-190.

[9] Lane, M.N. (2000) - *Pricing risk transfer transactions*, ASTIN Bulletin 30, 2, 259-293.

[10] West, M. (1987) - *On scale mixtures of normal distributions*, Biometrika 74, 646-648.

"Ovidius" University of Constanta
Department of Mathematics and Informatics,
900527 Constanta, Bd. Mamaia 124
Romania
e-mail: rvernic@univ-ovidius.ro