Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT

Journal of Multivariate Analysis 89 (2004) 181–190

Journal of Multivariate Analysis

http://www.elsevier.com/locate/jmva

A multivariate skew normal distribution

Arjun K. Gupta,[a,*] Graciela González-Farías,[b] and J. Armando Domínguez-Molina[c]

[a] *Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403-0221, USA*
[b] *Centro de Investigación en Matemáticas, Mexico*
[c] *Universidad de Guanajuato, Mexico*

Received 12 February 2002

Abstract

In this paper, we define a new class of multivariate skew-normal distributions. Its properties are studied. In particular we derive its density, moment generating function, the first two moments and marginal and conditional distributions. We illustrate the contours of a bivariate density as well as conditional expectations. We also give an extension to construct a general multivariate skew normal distribution.

AMS 2000 subject classifications: Primary 62H10; Secondary 62E15

Keywords: Non-normal models; Density; Marginal; Conditional; Regression; Moments; Moment generating function; Contours

1. Introduction

The univariate skew normal distribution was introduced by Azzalini [3,4] and its multivariate version by Azzalini and Dalla Valle [6] and Azzalini and Capitanio [5]. These classes of distributions include the normal and have some properties like the normal and yet are skew. They are useful in studying robustness.

*Corresponding author.
E-mail addresses: gupta@bgnet.bgsu.edu (A.K. Gupta), farias@cimat.mx (G. González-Farías), jadguez@cimat.mx (J.A. Domínguez-Molina).

doi:10.1016/S0047-259X(03)00131-3

It is a well-established fact, see Arnold and Beaver [2], that hidden truncation will lead to skew distribution, including the skew normal distributions and, in that sense, it is always nice to cope with the problem using a more suitable distribution.

Azzalini and Dalla Valle [6] and Azzalini and Capitanio [5] obtained the multivariate distribution by conditioning on one random variable being positive; we condition on the same number of random variables being positive. Hence, by construction, in the univariate case the two families are the same.

Arellano-Valle et al. [1] present two stochastic representations that are equivalent, one based on conditioning on the positiveness of a random vector and another one using a vector of absolute values. However, only for the univariate case these stochastic representations will be valid for our distribution and, it has not yet been possible to extend these representations without at least, having to modify the dimension of the conditional vector. The deficiency of stochastic representation comes from the fact that if $X \sim N_p(0, AA')$, then $X < 0$ does not imply that $AX < 0$ even when A is positive definite.

We study this new family of multivariate skew normal (MSN) distributions. This class of distributions includes the normal distribution and has some properties like the normal family and yet are skewed. In the next section we give the definition of this distribution, some of its properties and a method for simulating random vectors with this distribution. In Section 3, we compute the first two moments of the MSN distribution. We discuss in detail the bivariate skew normal (BSN) distribution in Section 4. Finally, in Section 5 we discuss a general version of the MSN distribution.

2. The MSN distribution

A continuous random vector $Y(p \times 1)$ is said to have a *multivariate skew normal distribution* if its p.d.f. is given by

$$f_p(y; \mu, \Sigma, D) = \frac{1}{\Phi_p(0; I + D\Sigma D')} \phi_p(y; \mu, \Sigma) \Phi_p[D(y - \mu)], \tag{2.1}$$

where $\mu \in \mathbb{R}^p$, $\Sigma > 0$, $D(p \times p)$, $\phi_p(\cdot; \mu, \Sigma)$ and $\Phi_p(\cdot; \Sigma)$ denote the p.d.f. and the c.d.f. of p-dimensional normal distribution with mean μ and covariance matrix $\Sigma > 0$. However, we will denote $\phi_p(\cdot; 0, I)$ by $\phi_p(\cdot)$, and $\Phi_p(\cdot; 0, \Sigma)$ by $\Phi_p(\cdot; \Sigma)$. We will denote that a random vector is distributed according to MSN with parameters μ, Σ, D by writing $Y \sim SN_p(\mu, \Sigma, D)$.

2.1. Some properties of $SN_p(\mu, \Sigma, D)$

The distribution function can be derived explicitly in terms of the multivariate normal distribution, the computation follows Azzalini and Dalla Valle.

Remark. If we have a collection of n independent random variables with skew normal distribution then the joint distribution of the n random variables is MSN.

This property does not hold for the MSN distribution defined by Azzalini and Dalla Valle [6, Eq. 1.1]. Thus, $\Sigma = \sigma^2 I$ and $D = \delta I$ give us the i.i.d. case.

2.1.1. The moment generating function

The following lemma is useful for evaluating some integrals that we will use in the rest of the paper, the proof is in the appendix.

Lemma 2.1. *Let B be a constant $(p \times p)$ matrix, and $a \in \mathbb{R}^p$. If $V \sim N_p(\mu_1, \Sigma)$ then*

$$E_V[\Phi_p(a + BV; \mu_2, \Omega)] = \Phi_p(a - \mu_2 + B\mu_1; \Omega + B\Sigma B').$$

Another useful result is the following,

Proposition 2.2. *If $Y \sim SN_p(\mu, \Sigma, D)$ then its m.g.f. is given by*

$$M_Y(t) = \frac{\Phi_p(D\Sigma t; I + D\Sigma D')}{\Phi_p(0; I + D\Sigma D')} e^{\mu' t + \frac{1}{2} t' \Sigma t}, \quad t \in \mathbb{R}^p. \tag{2.2}$$

The proof of this proposition is given in the appendix.

Note that if $D = diag(\delta_1, \ldots, \delta_p)$ and $\Sigma = diag(\sigma_1^2, \ldots, \sigma_p^2)$ then

$$M_Y(t) = 2^p \prod_{i=1}^{p} \Phi\left(\frac{\delta_i \sigma_i^2 t_i}{\sqrt{1 + \delta_i^2 \sigma_i^2}}\right) e^{t_i \mu_i + \frac{1}{2} \sigma_i^2 t_i^2},$$

which is the independent case.

From Proposition 2.2, we give the distribution of a linear transform of Y.

Proposition 2.3. *Let $Y \sim SN_p(\mu, \Sigma, D)$. Let $A(p \times p)$ a non-singular matrix and $b \in \mathbb{R}^p$ be constants, then $AY + b \sim SN_p(b + A\mu; A\Sigma A', DA^{-1})$.*

2.2. Construction of the MSN

In this section, we give a derivation of the MSN distribution based on a partitioned-conditional method. This procedure is useful for simulating random vectors with this distribution.

Proposition 2.4. *Let*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} \xi \\ \mu \end{pmatrix}, \begin{pmatrix} I + D\Sigma D' & D\Sigma \\ \Sigma D' & \Sigma \end{pmatrix}\right].$$

Then the distribution of the random vector $Y | \{X \geqslant \xi\}$ is $SN_p(\mu, \Sigma, D)$.

Proof. By using the fact

$$f_{Y|X}(y|X>\xi) = \frac{f_Y(y)Pr(X>\xi|y)}{Pr(X>\xi)},$$

and the conditional density of X given $Y=y$,

$$X|\{Y=y\} \sim N_p(\xi + D(y-\mu), I),$$

we conclude that

$$\begin{aligned} f_{Y|X}(y|X>\xi) &= \frac{\phi_p(y;\mu,\Sigma)}{Pr(\xi - X<0)} Pr(\xi - X<0|y) \\ &= \frac{\phi_p(y;\mu,\Sigma)}{\Phi_p(0; I + D\Sigma D')} \Phi_p[D(y-\mu)]. \qquad \square \end{aligned}$$

With the help of Proposition 2.4 we can prove the following result that generalized Proposition 1 of Azzalini and Dalla Valle [6].

Proposition 2.5. *Consider $W \sim N_p(0; I)$ and $X \sim N_p(0, \Sigma)$ two independent random vectors, and let D be an arbitrary $p \times p$ matrix, then*

$$Y = X|\{DX \geqslant W\}$$

has the density function given by (2.1).

3. Mean and variance of the MSN distribution

Let Y be a random vector with distribution $SN_p(\mu, \Sigma, D)$. In order to compute the first and second moments of the MSN distribution we consider the derivatives of the m.g.f. given in Eq. (2.2). Thus, the mean and the variance of Y are

$$EY = \mu + \frac{G^p(0; D\Sigma, I + D\Sigma D')}{\Phi_p(0; I + D\Sigma D')},$$

$$\begin{aligned} Var(Y) = \Sigma &- \frac{G^p(0; D\Sigma, I + D\Sigma D')G^{p'}(0; D\Sigma, I + D\Sigma D')}{\Phi_p^2(0; I + D\Sigma D')} \\ &+ \frac{G^p_{[2]}(0; D\Sigma, I + D\Sigma D')}{\Phi_p(0; I + D\Sigma D)}, \end{aligned}$$

where

$$G^p(0; A, \Omega) = \frac{\partial}{\partial t} \Phi_p(At, \Omega)\bigg|_{t=0}$$

and

$$G^p_{[2]}(0; A, \Omega) = \frac{\partial}{\partial t \partial t'} \Phi_p(At; \Omega)\bigg|_{t=0},$$

where $A(p \times p)$ is an arbitrary matrix and let $\Omega(p \times p)$ be a positive definite matrix.

The ith element of $G^p(0; A, \Omega)$ is

$$G^p_i(0; A, \Omega) = \frac{1}{\sqrt{2\pi}|\Omega|^{1/2}} \sum_{j=1}^{p} (A)_{ji} |H^{-1}_{(j)}|^{1/2} \Phi_{p-1}[0; H^{-1}_{(j)}],$$

where $H_{(j)}$ is the matrix obtained by deleting the jth row and the jth column from Ω^{-1}.

It is tedious to calculate $G^p_{[2]}(0; A, \Omega)$. An expression for $G^p_{[2]}(0; A, \Omega)$ is available in Gupta et al. [10, Appendix C.2].

In the special case when $D = diag(\delta_1, \ldots, \delta_p)$ and $\Sigma = diag(\sigma_1^2, \ldots, \sigma_p^2)$ the ith element of $G^p_{[2]}(0; A, \Omega)$ is

$$G^p_i(0; D\Sigma, I + D\Sigma D') = \frac{1}{2^{p-1}\sqrt{2\pi}} \frac{\delta_i \sigma_i^2}{1 + \delta_i^2 \sigma_i^2},$$

and the (i,j)th element of $G^p_{[2]}(0; D\Sigma, I + D\Sigma D')$ is

$$G^p_{i,j}(0; D\Sigma, I + D\Sigma D') = \begin{cases} \dfrac{\delta_i \delta_j \sigma_i^2 \sigma_j^2}{2^{p-1}\pi(1 + \delta_i^2\sigma_i^2)(1 + \delta_j^2\sigma_j^2)}, & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

Thus, the mean and variance of Y are given by

$$EY = \mu + \sqrt{\frac{2}{\pi}} v, \quad Var(Y) = \Sigma - \frac{2}{\pi} \Delta^2,$$

where $v = (\frac{\delta_1\sigma_1^2}{1+\delta_1^2\sigma_1^2}, \ldots, \frac{\delta_p\sigma_p^2}{1+\delta_p^2\sigma_p^2})'$ and $\Delta = diag(\frac{\delta_1\sigma_1^2}{1+\delta_1^2\sigma_1^2}, \ldots, \frac{\delta_p\sigma_p^2}{1+\delta_p^2\sigma_p^2})$.

4. The BSN distribution

In this section, we obtain a closed form for the density of a random vector with distribution $SN_2(\mu, \Sigma, D)$.

When $D = \begin{pmatrix} \delta_1 & \delta_2 \\ 0 & 0 \end{pmatrix}$ or $D = \begin{pmatrix} 0 & 0 \\ \delta_1 & \delta_2 \end{pmatrix}$, the BSN density reduces to

$$f_2(x, y; \mu, \Sigma, D) = 2\phi_2(x, y; \mu, \Sigma)\Phi[\delta_1(x - \mu_1) + \delta_2(y - \mu_2)]$$

which is the same as given by Azzalini and Dalla Valle [6, Eq. 1.1]. This case will not be discussed in this paper.

We will consider the general case where $D = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix}$. The next lemma helps us to evaluate the constant of the density (2.1).

Lemma 4.1. *If* $R = \begin{pmatrix} r_1^2 & r_{12} \\ r_{12} & r_2^2 \end{pmatrix}$ *is a positive definite matrix, then* $\Phi_2(0; R) = \frac{1}{2} - \frac{1}{2\pi} \arccos \frac{r_{12}}{r_1 r_2}$.

Note that if $r_{12} = 0$ we get $\Phi_2(0; R) = \frac{1}{4}$, because $\arccos(0) = \frac{1}{2}\pi$, this result coincides with the fact that when $R = \mathrm{diag}(r_1^2, r_2^2)$ it follows that

$$\Phi_2(0; R) = \Phi(0, r_1^2)\Phi(0, r_2^2) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}.$$

Corollary 4.2. *If* $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$ *is a positive definite matrix, and* $D = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix}$ *is an arbitrary matrix, then*

$$\Phi_2(0; I + D\Sigma D') = \frac{1}{2} - \frac{1}{2\pi} \arccos(\rho_{D\Sigma}), \tag{4.1}$$

where

$$\rho_{D\Sigma} = \frac{\delta_{21}\delta_{11}\sigma_1^2 + \delta_{22}\delta_{12}\sigma_2^2 + (\delta_{12}\delta_{21} + \delta_{22}\delta_{11})\sigma_1\sigma_2\rho}{\sqrt{(1 + \delta_{11}^2\sigma_1^2 + 2\delta_{11}\delta_{12}\sigma_1\sigma_2\rho + \delta_{12}^2\sigma_2^2)(1 + \delta_{21}^2\sigma_1^2 + 2\delta_{21}\delta_{22}\sigma_1\sigma_2\rho + \delta_{22}^2\sigma_2^2)}}.$$

If $\rho = 0$

$$\rho_{D\Sigma} = \frac{\delta_{21}\delta_{11}\sigma_1^2 + \delta_{22}\delta_{12}\sigma_2^2}{\sqrt{(1 + \delta_{11}^2\sigma_1^2 + \delta_{12}^2\sigma_2^2)(1 + \delta_{21}^2\sigma_1^2 + \delta_{22}^2\sigma_2^2)}}$$

and if $D = diag(\delta_1, \delta_2)$

$$\rho_{D\Sigma} = \frac{\delta_2\delta_1\sigma_1\sigma_2\rho}{\sqrt{(1 + \delta_1^2\sigma_1^2)(1 + \delta_2^2\sigma_2^2)}}, \tag{4.2}$$

and if $D = diag(\delta_1, \delta_2)$ and $\delta_1 = 0$ and/or $\delta_2 = 0$ then $\rho_{D\Sigma} = 0$.

If $\delta_{21} = \delta_{11}$ and $\delta_{12} = \delta_{22}$, then

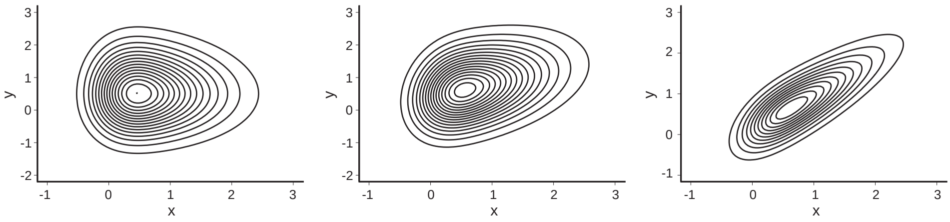
$$\rho_{D\Sigma} = \frac{\delta_{11}^2\sigma_1^2 + \delta_{22}^2\sigma_2^2 + 2\delta_{22}\delta_{11}\sigma_1\sigma_2\rho}{1 + \delta_{11}^2\sigma_1^2 + 2\delta_{11}\delta_{22}\sigma_1\sigma_2\rho + \delta_{22}^2\sigma_2^2}.$$

Let $\mu = (\mu_1, \mu_2)'$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$. From (2.1) and (4.1) we have that the density of the *bivariate skew normal distribution* is given by

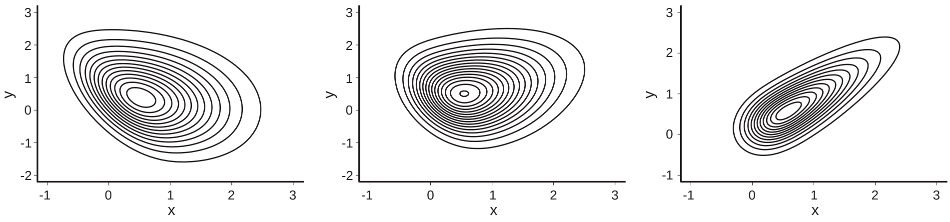
$$f_2(x, y; \mu, \Sigma, D) = \frac{\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\frac{\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}]\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}[\frac{1}{2} - \frac{1}{2\pi}\arccos(\rho_{D\Sigma})]}$$
$$\times \Phi[\delta_{11}(x-\mu_1) + \delta_{12}(y-\mu_2)]\Phi[\delta_{21}(x-\mu_1) + \delta_{22}(y-\mu_2)].$$

With the explicit expression for the density of the BSN distribution we can draw some contours of this density. Note that these contours are not elliptical.

(i) For $\mu = \binom{0}{0}, \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},\ D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$ and $\rho = 0, \frac{1}{2}, \frac{9}{10}$, respectively:



(ii) For $\mu = \binom{0}{0},\ \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},\ D = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$ and $\rho = 0, \frac{1}{2}, \frac{9}{10}$, respectively:



4.1. The marginal and conditional density of the BSN

Writing the m.g.f. of (X, Y) given in Proposition 2.2 as

$$M_{X,Y}(t_1, t_2) = \frac{\Phi_p(D\Sigma t; I + D\Sigma D')}{\Phi_p(0; I + D\Sigma D')} e^{\mu' t + \frac{1}{2} t' \Sigma t},$$

where $t = (t_1, t_2)'$, we get that the m.g.f. of X is given by

$$\begin{aligned} M_X(t_1) &= M_{X,Y}(t_1, 0) \\ &= \frac{\Phi_2[(D_1\Sigma_{11} + D_2\Sigma_{21})t_1; I + D\Sigma D']}{\Phi_2(0; I + D\Sigma D')} e^{\mu_1 t_1 + \frac{1}{2}\sigma_1^2 t_1^2}, \end{aligned} \tag{4.3}$$

where D was partitioned as $D = (D_1(2 \times 1) \quad D_2(2 \times 1))$. It is easy to verify that (4.3) corresponds to a random variable with distribution GMSN given in Section 5,

$$\begin{aligned} &GSN_{1,2}(\mu_1, \Sigma_{11}, D_1 + D_2\Sigma_{21}\Sigma_{11}^{-1}, D_1\mu_1 + D_2\Sigma_{21}\Sigma_{11}^{-1}\mu_1, I \\ &\quad + D_2(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})D_2'). \end{aligned}$$

Now, by direct computations, the conditional density of Y given $X = x$ is given by

$$GSN_{1,2}(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x-\mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, D_2, D_2\mu_2 - D_1(x-\mu_1), I).$$

Using (5.2) we get that the m.g.f. of the conditional distribution of Y given $X = x$ is

$$\begin{aligned} M_{Y|X}(t) \\ &= \frac{\Phi_2[D_2(\Sigma_{21}\Sigma_{11}^{-1}(x-\mu_1-\Sigma_{12}t)+\Sigma_{22}t); -D_1(x-\mu_1), I + D_2(\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})D_2']}{\Phi_2[D_2\Sigma_{21}\Sigma_{11}^{-1}(x-\mu_1); -D_1(x-\mu_1); I + D_2(\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})D_2']} \\ &\quad \times e^{\mu_2 t + \frac{\sigma_2}{\sigma_1}\rho(x-\mu_1)t + \frac{1}{2}\sigma_2^2(1-\rho^2)t^2}. \end{aligned}$$

When $D = diag(\delta_1, \delta_2)$ from (5.1) the conditional density of Y given $X = x$ is

$$f_{Y|X}(y|x; \mu, \Sigma, D) = \frac{\phi[y; \mu_2 + \frac{\sigma_2}{\sigma_1}\rho(x-\mu_1), \sigma_2^2(1-\rho^2)]}{\Phi[\frac{\delta_2\sigma_2\rho}{\sigma_1\sqrt{1+\delta_2^2\sigma_2^2(1-\rho^2)}}(x-\mu_1)]}\Phi[\delta_2(y-\mu_2)]$$

and the m.g.f. of $Y|\{X = x\}$ reduces to

$$M_{Y|X}(t) = \frac{\Phi[\frac{\delta_2\frac{\sigma_2}{\sigma_1}\rho(x-\mu_1)+\delta_2\sigma_2^2 t(1-\rho^2)}{\sqrt{1+\delta_2^2\sigma_2^2(1-\rho^2)}}]}{\Phi(\frac{\delta_2\sigma_2\,\rho(x-\mu_1)}{\sigma_1\sqrt{1+\delta_2^2\sigma_2^2(1-\rho^2)}})} e^{\mu_2 t + \frac{\sigma_2}{\sigma_1}\rho(x-\mu_1)t + \frac{1}{2}\sigma_2^2(1-\rho^2)t^2}. \tag{4.4}$$

We get the conditional expectation of $Y|X$ by differentiating the m.g.f. (4.4), thus

$$\begin{aligned} E(Y|x) &= \frac{\mathrm{d}}{\mathrm{d}t}E(e^{tY}|X)\Big|_{t=0} \\ &= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1) + \frac{(1-\rho^2)\delta_2\sigma_2^2}{\sqrt{1+(1-\rho^2)\delta_2^2\sigma_2^2}} \frac{\phi\left[\frac{\rho\delta_2\sigma_2(x-\mu_1)}{\sigma_1\sqrt{1+\delta_2^2\sigma_2^2(1-\rho^2)}}\right]}{\Phi\left[\frac{\rho\delta_2\sigma_2(x-\mu_1)}{\sigma_1\sqrt{1+\delta_2^2\sigma_2^2(1-\rho^2)}}\right]}. \end{aligned}$$

If $\delta_2 = 0$ we get that the marginal densities of X and Y are $SN_1(\mu_1, \sigma_1^2, \delta_1)$ and $SN_1(\mu_2, \sigma_2^2, 0) = N(\mu_2, \sigma_2^2)$, respectively. Thus, when $D = diag(\delta_1, 0)$ the conditional expectation of Y, given $X = x$, is the usual regression line of Y on X. Note that in this case the BSN density of $(X, Y)'$ reduces to

$$f_2(x, y; \mu, \Sigma, D) = 2\phi_2(x, y; \mu, \Sigma)\Phi[\delta_1(x-\mu_1)],$$

which is the Azzalini–Dalla Valle density with $\alpha = (\delta_1, 0)$.

5. General MSN distribution

In order to have a closed family such that it contains its marginal and conditional densities it is necessary to define a general version of the MSN distribution. We define the *general multivariate skew normal distribution* (GMSN) as a distribution whose density is of the form

$$f_{p,q}(y;\mu,\Sigma,D,v,\Delta)=\Phi_q^{-1}(D\mu;v,\Delta+D\Sigma D')\phi_p(y;\mu,\Sigma)\Phi_q(Dy;v,\Delta), \tag{5.1}$$

where $\mu\in\mathbb{R}^p$, $v\in\mathbb{R}^q$, and $\Sigma(p\times p)$ and $\Delta(q\times q)$ are two covariance matrices and $D(q\times p)$ is an arbitrary matrix. We say that a random variable W has distribution GMSN by writing $W\sim GSN_{p,q}(\mu,\Sigma,D,v,\Delta)$.

The fact that (5.1) is a density is readily verified with the help of Lemma 2.2.

The m.g.f. of this density is given by

Proposition 5.1. *If $Y\sim SN_{p,q}(\mu,\Sigma,D,v,\Delta)$ then its m.g.f. is given by*

$$M_Y(t)=\frac{\Phi_q[D(\mu+\Sigma t);v,\Delta+D\Sigma D']}{\Phi_q(D\mu;v;\Delta+D\Sigma D')}e^{\mu't+\frac{1}{2}t'\Sigma t}. \tag{5.2}$$

6. Concluding remarks

The skew normal distribution is very useful in applied statistics for modelling the skewness, see for example the works of Azzalini and Capitanio [5] and Genton et al. [9].

For the bivariate case we saw that the contours of the BSN density are not elliptical as we showed in Section 4. When the contours of a bivariate density are not elliptical the correlation coefficient is not a good measure of association between the two bivariate variables. Bjerve and Doksum [7] suggest to use a correlation curve which is a natural local measure of the strength of the association between Y and X near $Y=x$. We had worked the computation of the correlation curve for the BSN distribution and properties and applications of the general MSN distribution in Domínguez-Molina et al. [8].

Acknowledgments

This research has been supported by Consejo Nacional de Ciencia y Tecnología Grant 39017 E. The authors also wish to express their gratitude to the editor and referees for many useful suggestion which improved the presentation in this paper.

Appendix

Proof of Lemma 2.1. The proof is similar to the one given in the scalar case by Zacks [12, pp. 53–54]. Marsaglia [11] derives a similar result for the multivariate case.

According to the law of the iterated expectation, we have that

$$E_V[\Phi_p(a+BV;\mu_2,\Omega)] = E_V[Pr(U \leqslant a+BV|V)]$$
$$= Pr(U \leqslant a+BV),$$

where $U \sim N_p(\mu_2, \Omega)$. Then

$$E_V[\Phi_p(a+BV;\mu_2,\Omega)] = Pr(U - BV \leqslant a).$$

Given that $U - BV \sim N_p(\mu_2 - B\mu_1, \Omega + B\Sigma B')$ we get

$$Pr(U - BV \leqslant a) = \Phi_p(a;\mu_2 - B\mu_1, \Omega + B\Sigma B'). \qquad \square$$

Proof of Proposition 2.2. By definition of m.g.f. we have

$$M_Y(t) = Ee^{t'Y} = \Phi_p^{-1}(0; I + D\Sigma D') \int_{\mathbb{R}^p} e^{t'y}\phi_p(y;\mu,\Sigma)\Phi_p[D(y-\mu)]\,\mathrm{d}y$$
$$= \Phi_p^{-1}(0; I + D\Sigma D')e^{t'\mu+\frac{1}{2}t'\Sigma t} \int_{\mathbb{R}^p} \phi_p(y;\mu+\Sigma t,\Sigma)\Phi_p[D(y-\mu)]\,\mathrm{d}y$$
$$= \frac{E_W[\Phi_p(DW - D\mu)]}{\Phi_p(0; I + D\Sigma D')} e^{t'\mu+\frac{1}{2}t'\Sigma t},$$

where $W \sim N_p(\mu + \Sigma t, \Sigma)$. The result follows from Lemma 2.1. $\square$

References

[1] R. Arellano-Valle, G. Del Pino, E. San Martin, A general class of skew-distributions, Publicaciones técnicas, Centro de modelamiento matemático, Universidad de Chile UMR-CNRS, 2001.
[2] B.C. Arnold, R.J. Beaver, Skew multivariate models related to hidden truncation and/or selective reporting, Test. Sociedad de Estadística e Investigación Operativa. 11 (1) (2002) 7–54.
[3] A. Azzalini, A class of distributions which includes the normal ones, Scand. J. Statist. 12 (1985) 171–178.
[4] A. Azzalini, Further results on a class of distributions which includes the normal ones, Statistica 46 (1986) 199–208.
[5] A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew normal distribution, J. Roy. Statist. Soc. B 61 (1999) 579–602.
[6] A. Azzalini, A. Dalla-Valle, The multivariate skew-normal distribution, Biometrika 83 (1996) 715–726.
[7] S. Bjerve, K. Doksum, Correlation curves: measures of association as function of covariate values, Ann. Statist. 21 (2) (1993) 890–902.
[8] J.A. Domínguez-Molina, G. González-Farías, A.K. Gupta, A General Multivariate Skew Normal Distribution, Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 01-09, 2001.
[9] M.G. Genton, L. He, X. Liu, Moments of skew-normal random vectors and their quadratic forms, Statist. Probab. Lett. 51 (2001) 319–325.
[10] A.K. Gupta, G. Gonzáles-Farías, Domínguez-Molina, A Multivariate Skew Normal Distribution, Department of Mathematics and Statistics, Bowling Green State University, Technical Report No. 01-10, 2001.
[11] G. Marsaglia, Expressing the normal distribution with covariance matrix $A+B$ in terms of one with covariance matrix A, Biometrika 50 (1963) 535–538.
[12] S. Zacks, Parametric Statistical Inference, Oxford, Pergamon, 1981.