**Construction of Skew-Normal Random Variables: Are They Linear Combinations of Normal and Half-Normal?**

Mohsen Pourahmadi

pourahm@math.niu.edu

Division of Statistics

Northern Illinois University

**Abstract**

Skew-normal distributions extend the normal distributions through a shape parameter $\alpha$; they reduce to the standard normal random variable $Z$ for $\alpha = 0$ and to $|Z|$ or the half-normal when $\alpha \to \infty$. In spite of the skewness they (dis)inherit some properties of normal random variables: Square of a skew-normal random variable has a chi-square distribution with one degree of freedom, but the sum of two independent skew-normal random variables is not generally skew-normal. We review and explain this lack of closure and other properties of skew-normal random variables via their representations as a special linear combination of independent normal and half-normal random variables. Analogues of such representations are used to define multivariate skew-normal distributions with a closure property similar to that of multivariate normal distributions.

*Key Words and Phrases:* Closure Property; Conditioning; Extremum; Moment Generating Functions; Multivariate Distributions.

AMS 2000 Subject Classification: 60E05, 62E15.

# 1 Introduction

The central role of general normal random variables in probability and statistics is well-known and can be traced to the simplicity of the functional forms and basic symmetry properties of the probability density function (pdf) and cumulative distribution function (cdf) of the standard normal random variable (rv) $Z$:

$$\begin{cases} \phi(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-x^2/2} & , \quad \Phi(x) = \displaystyle\int_{-\infty}^{x} \phi(u)du, \\ \\ \phi(x) = \phi(-x) & , \quad \Phi(x) + \Phi(-x) = 1. \end{cases} \tag{1}$$

Surprisingly, the product of the first two functions in (1) gives rise to another interesting class of random variables which has been the subject of intense study in the last two decades. More precisely, for any real number $\alpha$, the function

$$f_\alpha(x) = 2\ \phi(x)\Phi(\alpha x), \tag{2}$$

is a bona fide pdf of a random variable $X$, which inherits a few features of the normal random variables. Some of these features happen to be the ones that make the normal distribution the darling of statistical inference.

The class of distributions (2) were introduced by Azzalini (1985) and christened *skew-normal distributions* with the skewness parameter $\alpha$, in symbol $X \sim SN(\alpha)$. Plots of $f_\alpha(\cdot)$ for several values of $\alpha$ given in Figure 1 show the role of $\alpha$ in determining the skewness of the distribution of $X$. Note that the right-tails of these distributions are virtually indistinguishable for $\alpha > 2$. The mean, variance and the classical measure of skewness of the rv $X$ with the pdf (2), computed by Azzalini (1985), further demonstrate the role of $\alpha$ in determining the shape and hence the moments of the skew-normal random variables.

In this expository and review paper we point out in a transparent manner the deeper connections among the skew-normal, standard normal, half-normal and truncated normal random variables. It provides the necessary background material for appreciating some of technical issues which arise in the search for an ideal multivariate extension of skew-normal distributions discussed in Arellano-Valle and Azzalini (2006); Genton (2004) and Pourahmadi (2007).

The outline of the paper is as follows. In Section 2, we review some basic properties of SN random variables along with a brief history. A characterization of SN random variables and a lack of closure property are presented in Section 3. Alternative methods of constructing SN random variables based on conditioning and truncation which have better prospects for extension to the multivariate setup are discussed in Section 4, with a brief review of multivariate SN distributions in Section 5.

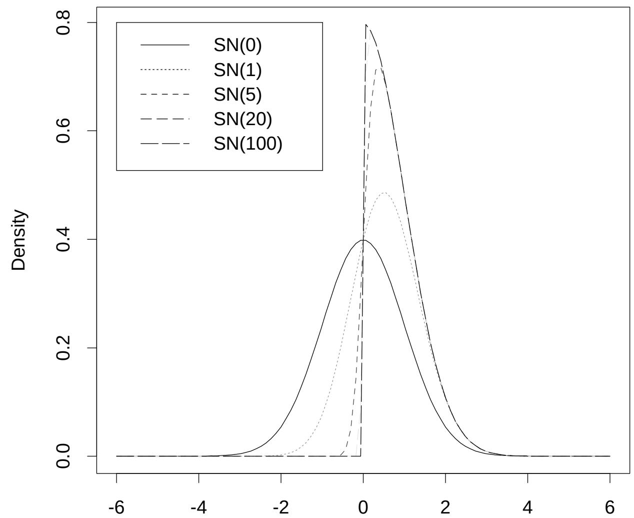


Figure 1: Plots of $f_\alpha(\cdot)$ for $\alpha = 0, 1, 5, 20, 100$.

## 2 Basic Properties and a Brief History

The following property of SN random variables is immediate from (2).

**Property** I. For $\alpha = 0, X = Z$ and for $\alpha \to \pm\,\infty, X = \pm|Z|$ where $Z \sim N(0, 1)$.

It shows that the normal and half-normal random variables lie at the *center* ($\alpha = 0$) and *boundary* ($\alpha = \pm\infty$) of the class of SN random variables, respectively. Research following the publication of Azzalini (1985, 1986), Henze (1986) and Arnold et al. (1993) has revealed that simple and common nonlinear operations such as truncation, conditioning and censoring performed on normal random variables lead invariably to versions of SN random variables. Consequently and not surprisingly, it has been revealed that the implicit appearance of SN random variables in the literature of statistics has a reasonably long pre-1985 history. The first known birthplace of SN distributions is the work of Birnbaum (1950) in the context of educational testing which involved truncation of normal variables, followed by the work of Weinstein (1964) and Nelson (1964) on finding the distribution of the sum of a normal variable and an independent truncated normal variable, see (21) below; some other early work are Roberts (1966), O'Hagan and Leonard (1976), Aigner et al. (1977) and Andel et al. (1984). For an extended review of the literature see Genton (2004), Arellano-Valle and Azzalini (2006) and Pourahmadi (2007).

The next three properties, which hold for any $\alpha$, can also be proved using (2):

**Property** II. If $X \sim SN(\alpha)$, then $-X \sim SN(-\alpha)$.

**Property** III. If $X \sim SN(\alpha)$, then $|X|$ and $|Z|$ are identically distributed.

**Property** IV. If $X \sim SN(\alpha)$, then $X^2 \sim \chi_1^2$, i.e. a chi-squared rv with *df* $= 1$.

The chi-square distribution in Property IV which is immediate from III, was first recognized and employed effectively by Roberts (1966), it implies in particular that the distributions of $|X|, X^2$, and all even functions of $X$ do not depend on the skewness parameter $\alpha$, i.e. there exists an *invariance property*, with respect to $\alpha$, that could have interesting inferential consequences (Genton et al. 2001; Loperfido, 2001. For example, all goodness-of-fit tests based on even functions of the data are incapable of distinguishing between normal and SN distributions (Loperfido, 2004).

Given the close affinity among $X, Z$ and $|Z|$, as stated in Properties I-IV, it seems natural to ask whether a rv $X$ with the pdf (2) can be expressed in terms of the standard normal rv $Z$ and the half-normal rv $|Z|$? In Section 2, we answer this question in the affirmative in the univariate case and give a characterization of $X \sim SN(\alpha)$ as a special weighted average of $Z$ and $|Z|$. The proof of the "if" part of the following result can be found in Henze (1986), see also Azzalini (1986).

**Property** V. A random variable $X$ has the pdf (2), if and only if it has the representation

$$X = \delta|Z_1| + \sqrt{1-\delta^2}\ Z_2, \tag{3}$$

where $Z_1, Z_2$ are independent $N(0,1)$ random variables, and

$$\delta = \frac{\alpha}{\sqrt{1+\alpha_2}} \in [-1, 1]. \tag{4}$$

It just happens that the new parameter $\delta$ is, indeed, the *correlation coefficient* between $X$ and $|Z_1|$. For example, the values $\alpha = 0, \pm 1, \pm\infty$, correspond to $\delta = 0, \pm\sqrt{2/2}, \pm 1$ and the transformation from $\alpha$ to $\delta$ is one-to-one and sign-preserving. Note that the coefficients in (3) do not add up to one, but their squares do:

$$1 = (\delta)^2 + \left(\sqrt{1-\delta^2}\right)^2, \tag{5}$$

so that $(X, |Z_1|, Z_2)$ in (3) are *bridged* with weights $(1, \delta, \sqrt{1-\delta^2})$ satisfying a Pythagorean identity. Note that the representation (3) is similar to the following representation of $X$ when the pair $(X, Z_1)$ is a bivariate normal random vector with standardized marginals and correlation $\delta$:

$$X = \delta Z_1 + \sqrt{1-\delta^2} Z_2, \tag{6}$$

where $Z_2 \sim N(0,1)$ and independent of $Z_1$.

The appearance of $|Z|$ or the half-normal distribution in (3) is the main cause of some difficulties in working with SN random variables. The most notable of these, the lack of closure under addition of independent SN random variables, is presented as the Property VII in Section 3.

# 3 The pdf and mgf of SN random variables

In this section, first we show that $f_\alpha(\cdot)$ in (2) is, indeed, the pdf of the rv $X$ defined via the representation (3). Then, the moment generating function (mgf) of $|Z|$ and the representation (3) are used to derive the mgf of $X$.

## 3.1 The pdf of SN random variables

In this section we give a simultaneous proof of the Property III and that $f_\alpha(\cdot)$ is a genuine pdf. Note that both follow from (2), the last two identities in (1) and the fact that the pdf of $|Z|$ is $2\phi(z), z > 0$, and zero otherwise, as outlined next.

For any $x > 0$, we have

$$
\begin{aligned}
P(|X| \le x) &= \int_{-x}^{x} 2\phi(u)\Phi(\alpha u)\, du \\
&= \int_0^x 2\phi(u)\Phi(\alpha u) du + \int_{-x}^{0} 2\phi(u)\Phi(\alpha u)\, du \\
&= \int_0^x 2\phi(u)\Phi(\alpha u) du - \int_{x}^{0} 2\phi(-u)\Phi(-\alpha u)\, du \\
&= \int_0^x 2\phi(u)\Phi(\alpha u) du + \int_{0}^{x} 2\phi(u)\Phi(-\alpha u)\, du \\
&= \int_0^x 2\phi(u)\left[\Phi(\alpha u) + \Phi(-\alpha u)\right] du \\
&= \int_0^x 2\phi(u) du = P(|Z| \le x).
\end{aligned}
$$

Moreover, the above identity for $x = \infty$ shows that (2) is, indeed, a bona fide pdf.

That $f_\alpha(\cdot)$ is the pdf of (3) was shown first by Henze (1986) using a conditioning argument. We prove the same directly by applying the standard formula for the pdf of sum of two independent random variables to

$$X = a|Z_1| + bZ_2, \tag{7}$$

where $a = \delta, b = \sqrt{1-\delta^2}, a^2 + b^2 = 1, \dfrac{a}{b} = \alpha$. Without loss of generality we take $\delta > 0$ (see Property II), and write

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f_{a|Z_1|}(x-y) f_{bZ_2}(y)\, dy = \\
&= \int_{-\infty}^{x} 2 \cdot \frac{1}{\sqrt{2\pi}a} \exp\left\{-\frac{1}{2a^2}(x-y)^2\right\} \cdot \frac{1}{\sqrt{2\pi}b} \exp\left\{-y^2/2b^2\right\} dy \\
&= \frac{2}{2\pi ab} \int_{-\infty}^{x} \exp\left\{-\frac{1}{2}\left[\frac{(x-y)^2}{a^2} + \frac{y^2}{b^2}\right]\right\} dy.
\end{aligned}
$$

Using the identity

$$\frac{(x-y)^2}{a^2} + \frac{y^2}{b^2} = x^2 + \frac{1}{a^2b^2}(y - xb)^2,$$

for the $a$ and $b$ given above, we obtain

$$\begin{aligned} f_X(x) &= 2\phi(x)\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}ab}\exp\left\{-\frac{1}{2a^2b^2}(y-xb)^2\right\}dy \\ &= 2\phi(x)\Phi\left(\frac{a}{b}x\right) = 2\phi(x)\Phi(\alpha x), \end{aligned}$$

which is the desired result.

## 3.2 The mgf of SN random variables

Next, we find the mgf of $X$ by computing the mgf of $|Z|$ and recalling that the mgf of a standard normal rv $Z$ is

$$m_Z(t) = e^{t^2/2}. \tag{8}$$

From the definition of a mgf and using the last identity in (1), it follows that

$$\begin{aligned} m_{|Z|}(t) = E\left(e^{t|Z|}\right) &= \int_{-\infty}^{+\infty} e^{t|z|}\ \phi(z)dz = 2\int_0^\infty e^{tz}\ e^{-z^2/2}\frac{dz}{\sqrt{2\pi}} \\ &= 2e^{t^2/2}\int_0^\infty e^{-\frac{1}{2}(z-t)^2}dz \\ &= 2e^{t^2/2}\int_{-t}^\infty e^{-u^2/2}du\ /\sqrt{2\pi} = 2e^{t^2/2}\Phi(t). \end{aligned} \tag{9}$$

Then, using (3), (8)-(9) and independence of $|Z_1|$ and $Z_2$, we obtain

$$\begin{aligned} m_X(t) &= E(e^{tX}) = E\left(e^{\delta t|Z_1|+\sqrt{1-\delta^2}\ tZ_2}\right) \\ &= m_{|Z_1|}(\delta t)m_{Z_2}(\sqrt{1-\delta^2}t) = 2e^{t^2/2}\Phi(\delta t). \end{aligned} \tag{10}$$

To recapitulate, so far it has been shown that a rv with representation (3) has the pdf (2) of an $SN(\alpha)$ random variable. From the uniqueness of the mgf of random variables, it follows that (10) is the mgf of an $SN(\alpha)$ random variable. Now, by reversing the steps (or the equalities) in (10), it follows that any SN random variable has the representation (3). This completes the proof of the statement in **Property V**.

More generally, the arguments leading to (10) can be used to find the mgf of $X$ in (7), for arbitrary $a$ and $b$:

$$m_X(t) = 2e^{a^2\frac{t^2}{2}}\ \Phi(at)\cdot e^{b^2\frac{t^2}{2}} = 2e^{(a^2+b^2)t^2/2}\ \Phi(at).$$

Note that after rescaling, $X$ has a skew-normal distribution of the form (3), i.e.

$$(a^2+b^2)^{-1/2}\ X = (a^2+b^2)^{-1/2}\ a|Z_1| + (a^2+b^2)^{-1/2}b\ Z_2 \sim SN\left(\frac{\delta}{\sqrt{1-\delta^2}}\right), \tag{11}$$

where $\delta = (a^2+b^2)^{-1/2}a$ is the coefficient of $|Z_1|$.

## 3.3 Lack of Closure and Bivariate SN

From (3) and (11) it is evident that the class of SN random variables remains "closed" with respect to the addition of independent $N(0,1)$ random variables, in the sense presented next.

**Property** VI. If $X \sim SN(\alpha)$ and $Z \sim N(0,1)$ are independent, then

$$\frac{X+Z}{\sqrt{2}} \sim SN\left(\frac{\alpha}{\sqrt{2+\alpha^2}}\right). \tag{12}$$

However, unlike the normal random variables, the class of SN random-variables is not closed with respect to the addition of independent copies of its members.

**Property** VII. Let $X_i \sim SN(\alpha_i)$ be independent with $\alpha_i \neq 0, i = 1,2$. Then, in general, $X_1 + X_2$ is not SN. However, if $X_1$ and $X_2$ are *dependent* sharing a common half-normal, then $X_1 + X_2$ is SN. More precisely, if

$$\begin{aligned} X_1 &= \delta_1|Z| + \sqrt{1-\delta_1^2}Z_1, \\ X_2 &= \delta_2|Z| + \sqrt{1-\delta_2^2}Z_2, \end{aligned} \tag{13}$$

where $Z, Z_1, Z_2$ are independent $N(0,1)$. Then, from (11) it follows that

$$(1+2\ \delta_1\delta_2)^{-1/2}(X_1+X_2) \sim SN\left(\frac{\delta_1+\delta_2}{\sqrt{1+2\delta_1\delta_2}}\right). \tag{14}$$

The reason for the lack of closure of general SN random variables can be traced to the fact that for independent $Z_i \sim N(0,1), i = 1,2$, sum of their absolute values cannot be written as the absolute value of another normal rv:

$$m_{|Z_1|+|Z_2|}\ (t) = 4e^{t^2}\Phi^2(t) \neq 2e^{c^2t^2/2}\ \Phi(ct), \tag{15}$$

for any $t$ and $c$.

# 4 Extensions and Constructions

In this section, several extensions and methods of construction of SN random variables are reviewed. The goal is to identify and single out the most suitable construction/extension as a stepping-stone for defining a flexible class of multivariate SN distributions with a closure property. The conditioning and truncation methods turn out to be the most suitable for the task.

## 4.1 Location-Scale Family

The relations (11)- (14) do motivate the need for scaling and generalizing the SN random variables by the inclusion of location and scale parameters. Let us denote the new parameters by $\xi$ and $\omega$, respectively, then for any $X \sim SN(\alpha)$, define a general SN random variable by

$$Y = \xi + \omega X, \tag{16}$$

and write $Y \sim SN(\xi, \omega, \alpha)$ for this rv. The pdf and mgf of $Y$, found using (2) and (10) are

$$\begin{aligned} f(y;\alpha,\xi,\omega) &= \frac{2}{\omega}\phi\left(\frac{y-\xi}{\omega}\right)\Phi\left(\alpha\frac{y-\xi}{\omega}\right), \\ m_Y(t) &= 2\ e^{\xi t+\frac{t^2}{2\omega^2}}\ \Phi(\delta\omega t). \end{aligned} \tag{17}$$

## 4.2 Conditioning on a Set of Values

It is well-known that for a bivariate normal random vector

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right], \tag{18}$$

the conditional distribution of $X_1$ given a *specific* value of $X_2$ is also normal, i.e.

$$(X_1|X_2 = x_2) \sim N\left[\mu_1 + \rho\ \frac{\sigma_1}{\sigma_2}(x_2-\mu_2),\ \sigma_1^2(1-\rho^2)\right].$$

However, the conditional distribution of $X_1$ given a subset $S$ of possible values of $X_2$, i.e. the distribution of $(X_1|X_2 \in\ S)$ is no longer normal when $S$ is not the singleton $\{x_2\}$. What is this conditional distribution? Fortunately, it has an explicit form and belongs to the class of SN distributions when $S = \{X_2 > c\}$ for a constant $c$. This leads to finding the distribution of $(X_1|X_2 > c)$ which is the distribution of $X_1$ when $X_2$ exceeds a threshold $c$ (Birnbaum, 1950; Arnold and Beaver, 2002).

The density function of $(X_1|X_2 > c)$ evaluated at a point $x_1$, denoted by $f_{X_1|X_2>c}\,(x_1)$ can be computed using the basic formula

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)},$$

with $A = \{X_1 \ \in (x_1 - h, x_1 + h),\ B = \{X_2 > c\}$ and letting $h \to 0$. Few of the results for various values of $c$ and certain circumstances encountered in the literature are summarized in the following:

**Theorem.** (a) Arnold et al. (1993). With $(X_1, X_2)$ as in (18), the conditional density of $X_1$ given that $X_2 > c$ is

$$f_{X_1|X_2>c}\,(x_1) = \frac{1}{\sigma_1}\phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right)\Phi\left(\frac{\rho\frac{x_1-\mu_1}{\sigma_1} - \theta_c}{\sqrt{1-\rho^2}}\right)/\Phi(\theta_c), \tag{19}$$

where $\theta_c = -(c - \mu_2)/\sigma_2$. This distribution is a bit more general than a $SN(\xi, \omega, \alpha)$ or (17), but it reduces to them when $\theta_c = 0$ or when the threshold $c$ is the mean $\mu_2$ of $X_2$.

(b) Azzalini and Dalla Valle (1996). If $(X_1, X_2)$ is a bivariate normal random vector with standardized marginals and correlation $\rho$, then

$$(X_1|X_2 > 0) \sim SN(\alpha),$$

where $\alpha = \dfrac{\rho}{\sqrt{1-\rho^2}}$.

(c) Azzalini and Dalla Valle (1996). If $X_1$ and $X_2$ are independent $N(0, 1)$ random variables, then

$$(X_1|\alpha X_1 > X_2) \sim SN(\alpha).$$

### 4.3 Truncation of Normal Random Variables

A slightly different extension of the class of SN random variables can be obtained by replacing the $|Z_1|$ in (3) by $Z_1(c)$, where the latter stands for $Z_1$ truncated from below at $c$. More precisely, for a given $c$ and $Z \sim N(0,1), Z(c)$ is $Z$ if $Z \geq c$ and zero otherwise. Evidently, the pdf of $Z(-c)$ is $\dfrac{\phi(z)}{\Phi(c)}$ and its mgf is

$$\begin{aligned} m_{Z(-c)}(t) = \quad & E\left(e^{tZ(-c)}\right) = \frac{1}{\Phi(c)}\int_{-c}^{\infty} e^{tz}\phi(z)dz \\ & = \frac{e^{t^2/2}}{\Phi(c)}\int_{-c}^{\infty} e^{-\frac{1}{2}(z-t)^2}dz = e^{t^2/2}\frac{\Phi(t+c)}{\Phi(c)}. \end{aligned} \tag{20}$$

Thus, the pdf and mgf of

$$X = \rho Z_1(-c) + \sqrt{1-\rho^2}\ Z_2, \tag{21}$$

found using the technique leading to (10), are

$$f_{\alpha,c}(x) = \phi(x)\Phi\left(c\sqrt{1+\alpha^2}+\alpha x\right)/\Phi(c), m_X(t) = e^{t^2/2}\frac{\Phi(\rho t + c)}{\Phi(c)}. \tag{22}$$

We note that when $c = 0$, then (20) - (22), reduce to (9), (3) and (2), respectively. This type of distributions studied by Birnbaum (1950) in the context of educational testing seems to be the first known birthplace of SN distributions. Azzalini (1985) refers to (22) as the extended SN random variable. Arnold and Beaver (2002) refer to (22) as the *additive component construction* of SN random variables. Surprisingly, though this construction method is distinct from that of conditioning on a set of values discussed in the previous section, they are equivalent in that they lead to the same class of SN random variables as can be seen by comparing the density functions in (18) and (22) or their moment generating functions. This important and unexpected equivalence were noted first by Arellano-Valle and Azzalini (2006).

## 4.4 Extremum of Normal Random Variables

The distribution of the minimum of two standardized correlated normal variables was found by Roberts (1966) in the context of twin studies. The resulting distribution is now recognized as the SN, and research on finding the distribution of extremum of normal random variables has been continued independently by Cain (1994) and Loperfido (2002). They have shown that for a general bivariate normal distribution as in (18), these distributions are $\frac{1}{2} : \frac{1}{2}$ mixtures of two SN densities. More specifically, for $\mu_1 = \mu_2 = 0$ and defining $Y_1$ and $Y_2$ as

$$\frac{Y_1}{\sigma_1} \sim SN\left(\frac{\sigma_1 - \rho\sigma_2}{\sigma_2\sqrt{1-\rho^2}}\right), \frac{Y_2}{\sigma_2} \sim SN\left(\frac{\sigma_2 - \rho\sigma_1}{\sigma_2\sqrt{1-\rho^2}}\right),$$

then, the distribution of $\max\{X_1, X_2\}$ is the mixture with equal weights of the distributions of $Y_1$ and $Y_2$. Of course, the distribution of $\min\{X_1, X_2\}$ is the same as that of $-\max\{X_1, X_2\}$.

# 5 Multivariate SN Distributions

In the last two decades several classes of multivariate SN distributions have been defined that mimic certain properties of the multivariate normal distribution. The initial proposal by Azzalini (1985) is unsatisfactory since its marginals are not SN. A class introduced later by Azzalini and Dalla Valle (1996) has proved useful in dealing with non-normal multivariate data (Azzalini and Capitanio, 1999). However, it lacks the familiar *closure property* of the multivariate normal: When $X_1$ and $X_2$ are independent univariate SN's, then $X = (X_1, X_2)$ does not have a bivariate SN distribution, see also Section 2.3. Next, by using $(X_1, X_2)$ in (13), we motivate the Azzalini and Dalla Valle (1996) class.

Recalling the close functional forms of the pdf and mgf of SN random variables, see (2) and (10), in this section we introduce multivariate SN distributions through their mgf's. To add some generality to $X_1$ and $X_2$ in (13), we take $(Z_1, Z_2)$ to be correlated, i.e.

$$(Z_1, Z_2)' \sim N_2(0, \boldsymbol{\Psi}) \text{ with } m_{Z_1,Z_2}(t) = \exp\left\{\frac{1}{2}\, t'\boldsymbol{\Psi}t\right\},$$

where $\mathbf{\Psi} = \begin{pmatrix} 1 & \psi \\ \psi & 1 \end{pmatrix}$ and $t = (t_1, t_2)$. Then, using (13), the independence of $|Z|$ and $(Z_1, Z_2)$ and the forms of their mgf's, the mgf of $(X_1, X_2)$ is given by

$$\begin{aligned}
m_{(X_1,X_2)}(t) &= E\exp\{t_1(\rho_1|Z| + \sqrt{1-\rho_1^2}\ Z_1) + t_2(\rho_2|Z| + \sqrt{1-\rho_2^2}\ Z_2\} \\
&= E\exp\{(\rho_1 t_1 + \rho_2 t_2)|Z|\}E\exp\{\sqrt{1-\rho_1^2}\ t_1 Z_1 + \sqrt{1-\rho_2^2}\ t_2 Z_2\} \\
&= m_{|Z|}(\rho_1 t_1 + \rho_2 t_2)\ m_{(Z_1,Z_2)}\left(\sqrt{1-\rho_1^2}\ t_1,\ \sqrt{1-\rho_2^2}\ t_2\right) \\
&= 2\exp\{\tfrac{1}{2}(\rho_1 t_1 + \rho_2 t_2)^2\}\ \Phi(\rho_1 t_1 + \rho_2 t_2)\ m_{(Z_1,Z_2)}\ \left(\sqrt{1-\rho_1^2}\ t_1,\ \sqrt{1-\rho_2^2}\ t_2\right) \\
&= 2\exp\{\tfrac{1}{2}t'\boldsymbol{\rho}\boldsymbol{\rho}'t + \tfrac{1}{2}\ t'D\mathbf{\Psi}Dt\}\Phi(\rho_1 t_1 + \rho_2 t_2) \\
&= 2\exp\{\tfrac{1}{2}t'\,[\boldsymbol{\rho}\boldsymbol{\rho}' + D\mathbf{\Psi}D]\,t\}\Phi(\rho_1 t_1 + \rho_2 t_2), \\
&= 2\exp\{\tfrac{1}{2}t'\Omega\ t\}\Phi(\boldsymbol{\rho}'t),
\end{aligned} \tag{23}$$

where $\boldsymbol{\rho} = (\rho_1, \rho_2), D =$ diag $\left(\sqrt{1-\rho_1^2}, \sqrt{1-\rho_2^2}\right)$ and

$$\Omega = \boldsymbol{\rho}\boldsymbol{\rho}' + D\mathbf{\Psi}D. \tag{24}$$

The simple and transparent form of the mgf (23) suggests an obvious way to define multivariate distributions that, (i) are not normal and will reduce to them when $\boldsymbol{\rho} = 0$, and (ii) have SN marginals. Thus, following Azzalini and Dalla Valle (1996) for a given $p \times p$ correlation matrix $\Omega$ and $p \times 1$ vector $\boldsymbol{\rho} = (\rho_1, \cdots, \rho_p)$ with $|\rho_i| \leq 1$, we say $X \sim SN_p(\Omega, \boldsymbol{\rho})$ if the mgf of $X$ is of the form (23) for $t \epsilon R^p$. As noted earlier, a major drawback of such definition of multivariate SN distributions is that for $X_i$'s independent univariate $SN(\rho_i)$, the mgf of $X = (X_1, \cdots, X_p)$ given by $\prod_{i=1}^{p} 2\exp\left\{\frac{1}{2}\ t_i^2\right\}\Phi(\rho_i t_i)$ is not of the form (23) and hence $X$ does not belong to the class of $SN_p(\Omega, \boldsymbol{\rho})$ distributions.

Fortunately, alternative classes of multivariate SN distributions with various closure properties have been introduced in recent years by González-Farías et al. (2004) and Gupta and Chen (2004) where the latter relies on a vectorial version of (3). Certain computational complexity issues related to using these classes of distributions in time series applications are discussed in Pourahmadi (2007). An excellent framework for unification of these classes based on truncation (see Section 4.3) is presented in Arellano-Valle and Azzalini (2006). Their formulation removes certain redundancies in parameterization and clarifies nicely the connections among various classes of multivariate SN distributions in the literature.

**Acknowledgements.** The work was supported by the NSF grants DMS-0307055 and DMS-0505696.

## References


Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function model. *J. Econometrics*, **12**, 21-37.

Andel, J., Netuka, I. and Svara, K. (1984). On threshold autoregressive processes. Kybernetika, **20**, 89-106.

Arellano-Valle, R.B. and Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*, **33**, 159-188.

Arnold, B.C. and Beaver, R.J. (2002). Skewed multivariate models related to hidden truncation. *Test*, 11, 7-54.

Arnold, B.C., Beaver, R.J., Groeneveld, R.A. and Meeker, W.Q. (1993). The nontruncated marginal of a truncated bivariate normal. *Psychometrika*, 58, 471-488.

Azzalini, A. (1985). A class of distribution which includes the normal ones. *Scan. J. Statist.*, **12**, 171-178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica*, **46**, 199-208.

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

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *J. Roy. Statist. Soc. B*, **61**, 579-602.

Birnbaum, Z.W. (1950). Effect of linear truncation on a multinormal population. *Ann. Math. Statist.*, **21**, 272-279.

Cain, M. (1994). The moment-generating function of the minimum of bivariate normal random variables. *The American Statistician*, **79**, 99-113.

Genton, M.G. (2004). *Skew-Elliptical Distributions And Their Applications: A Journey Beyond Normality.* Chapman & Hall/CRC, Boca Raton.

Genton, M.G., He, L. and Liu, X. (2001). Moments of skew-normal random vectors and their quadratic forms. *Statist. & Prob. Lett.*, **51**, 319-325.

González-Farías, G., Dominguez-Molina, J.A. and Gupta, A.K. (2004). Additive properties of skew-normal random vectors. *J. Statist. Planning and Inference*, **126**, 521-534.

Gupta, A.K. and Chen, J. T. (2004). A class of multivariate skew-normal models. *Ann. Inst. Statist. Math.*,**56**, 305-315.

Henze, N. (1986). A probabilistic representation of the 'skew-normal' distribution. *Scand. J. Statist.*, **13**, 271-275.

Loperfido, N. (2001). Quadratic forms of skew-normal random vectors. *Statist. & Prob. Lett.* **54**, 381-387.

Loperfido, N. (2002). Statistical implications of selectively reported inferential results. *Statist. & Probab. Let.*, **56**, 13-22.

Loperfido, N. (2004). Generalized skew-normal distributions. In M.G. Genton (Ed.), Skew-elliptical distributions and their applications, Chap. 4 (pp. 65-80). Chapman & Hall/CRC.

Nelson, L.S. (1964). The sum of values from a normal and a truncated normal distribution. *Technometrics*, **6**, 469-471.

O'Hagan, A. and Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, **63**, 201-202.

Pourahmadi, M. (2007). Skew-normal ARMA models with nonlinear heteroskedastic predictors. *Communication in Statistics*. To appear.

Roberts, C. (1966). A correlation model useful in the study of twins. *J. Amer. Statist. Assoc.*, **61**, 1184-1190.

Weinstein, M.A. (1964). The sum of values from a normal and a truncated normal distribution. *Technometrics*, **6**, 104-105.