Ann. Inst. Statist. Math.
Vol. 56, No. 2, 351–360 (2004)

CHARACTERIZATION OF THE SKEW-NORMAL DISTRIBUTION

ARJUN K. GUPTA[1], TRUC T. NGUYEN[1] AND JOSE ALMER T. SANQUI[2]*

[1] *Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403-0221, U.S.A.*

[2] *Department of Mathematical Sciences, Appalachian State University, ASU Box 32092, Boone, NC 28608-2092, U.S.A.*

(Received January 25, 2002; revised July 23, 2003)

Abstract. Two characterization results for the skew-normal distribution based on quadratic statistics have been obtained. The results specialize to known characterizations of the standard normal distribution and generalize to the characterizations of members of a larger family of distributions. Results on the decomposition of the family of distributions of random variables whose square is distributed as χ_1^2 are obtained.

Key words and phrases: Non-normal distribution, chi-square distribution, half-normal distribution, skew-symmetric distribution, sequence of moments, induction, decomposition, characteristic function.

1. Introduction

A random variable Z has a skew-normal distribution with parameter λ, denoted by $Z \sim SN(\lambda)$, if its density is given by $f(z, \lambda) = 2\Phi(\lambda z)\phi(z)$, where Φ and ϕ are the standard normal cumulative distribution function and the standard normal probability density function, respectively, and z and λ are real numbers (Azzalini (1985)).

Some basic properties of the $SN(\lambda)$ distribution given in Azzalini (1985) are:

1. $SN(0) = N(0,1)$;
2. If $Z \sim SN(\lambda)$ then $-Z \sim SN(-\lambda)$;
3. As $\lambda \to \pm\infty$, $SN(\lambda)$ tends to the half-normal distribution, i.e., the distribution of $\pm|X|$ when $X \sim N(0,1)$; and
4. If $Z \sim SN(\lambda)$ then $Z^2 \sim \chi_1^2$.

Properties 1, 2, and 3 follow directly from the definition while Property 4 follows immediately from

LEMMA 1.1. (Roberts and Geisser (1966)) $W^2 \sim \chi_1^2$ *if and only if the p.d.f. of* W *has the form* $f(w) = h(w)\exp(-w^2/2)$ *where* $h(w) + h(-w) = \sqrt{2/\pi}$.

In terms of characteristic functions, Lemma 1.1 can be restated as

LEMMA 1.2. (Roberts (1971)) $W^2 \sim \chi_1^2$ *if and only if the characteristic function* Ψ_W *of* W *satisfies* $\Psi_W(t) + \Psi_W(-t) = 2\exp(-t^2/2)$.

*Research supported by a non-service fellowship at Bowling Green State University.

That the skew-normal density is a proper density follows directly from the following lemma:

LEMMA 1.3. (Azzalini (1985)) *Let f be a density function symmetric about* 0, *and G an absolutely continuous distribution such that G' is symmetric about* 0. *Then $2G(\lambda y)f(y)$, where $-\infty < y < \infty$, is a density function for any real λ.*

A probabilistic representation of a skew-normal random variable is given in

LEMMA 1.4. (Henze (1986)) *If U and V are identically and independently distributed $N(0,1)$ random variables then $\frac{\lambda}{\sqrt{1+\lambda^2}}|U| + \frac{1}{\sqrt{1+\lambda^2}}V \sim SN(\lambda)$.*

The characteristic function of the $SN(\lambda)$ distribution is given by

LEMMA 1.5. (Pewsey (2000*b*)) *If $Z \sim SN(\lambda)$ then its characteristic function is $\Psi_Z(t) = \exp(-t^2/2)(1+i\tau(\delta t))$ where for $x \geq 0$, $\tau(x) = \int_0^x \sqrt{2/\pi}\exp(u^2/2)du$, $\tau(-x) = -\tau(x)$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.*

The skew-normal distribution, due to its mathematical tractability and inclusion of the standard normal distribution, has attracted a lot of attention in the literature. Azzalini (1985, 1986), Chiogna (1998) and Henze (1986) discussed basic mathematical and probabilistic properties of the $SN(\lambda)$ family. The works of Azzalini and Dalla Valle (1996), Azzalini and Capitanio (2003), Arnold *et al.* (1993), Arnold and Beaver (2002), Gupta *et al.* (2002*a*) and Branco and Dey (2001) focused on the theoretical developments of various extensions and multivariate generalizations of the model. Loperfido (2001), Genton *et al.* (2001) and Gupta and Huang (2002) focused on probabilistic properties of quadratic skew-normal variates. The statistical inference aspect for this distribution is partially addressed in Azzalini and Capitanio (1999), Pewsey (2000*a*), Salvan (1986) and Liseo (1990). Gupta and Chen (2001) tabulated the c.d.f. of the $SN(\lambda)$ distribution and illustrated the use of their table in goodness-of-fit testing for this distribution. Applications in reliability studies was discussed in Gupta and Brown (2001). Very few, however, tackled the problem of characterizing this seemingly important distribution. It is to fill this void in the literature that this paper came about.

In this paper, we give two characterization results for the $SN(\lambda)$ distribution. We first give the results in more general form and then state the results in the context of the skew-normal and standard normal distributions as corollaries.

In Section 2, we give a generalization of a characterization of the normal distribution based on quadratic statistics given in Roberts and Geisser (1966). In their paper, they showed that, if X_1 and X_2 are independently and identically distributed (i.i.d.) random variables, then X_1^2, X_2^2, and $\frac{1}{2}(X_1+X_2)^2$ are all χ_1^2 distributed if and only if X_1 and X_2 are standard normal random variables. Since the standard normal distribution belongs to the skew-normal class, a natural question to ask is whether a similar characterization holds true for the skew-normal distribution. The answer is given by Corollary 2.2 which generalizes the result of Roberts and Geisser.

Another characterization based on the quadratic statistics X^2 and $(X+a)^2$ for some constant $a \neq 0$ will be given.

In Section 3, the decomposition of a larger family of distributions which we will refer to as the $SN3$ family, will be discussed.

2. Characterization results

The characterization results in this section are closely tied up with the so called Hamburger moment problem and uniqueness problem which basically ask the questions "Given a sequence of real numbers, does there exist a distribution whose sequence of moments coincide with the given sequence and if so, is the distribution unique?". A solution to the uniqueness problem is given in the following corollary:

COROLLARY 2.1. (Shohat and Tamarkin (1943) p. 20) *If the Hamburger moment problem has a solution* $F(t) = \int_{-\infty}^{t} f(t)dt$ *where* $f(t) \geq 0$ *and* $\int_{-\infty}^{\infty} f(t)^q e^{s|t|}dt < \infty$ *for some* $q \geq 1$ *and* $s > 0$, *then the solution is unique.*

An immediate consequence of the previous corollary is the following result:

LEMMA 2.1. *The skew-normal distribution is uniquely determined by its sequence of moments.*

PROOF. We only need to note that the conditions of the previous corollary are satisfied by the standard normal distribution (i.e. take $f(t)$ = standard normal p.d.f.) with $q = 1$ and $s = 1$. Now, since one tail of the $SN(\lambda)$ distribution, when $\lambda \neq 0$, is shorter than that of the standard normal distribution and the other tail has the same rate of convergence to 0 as the standard normal distribution, it follows that the conditions of Corollary 2.1 are also satisfied by the skew-normal distribution. That is, taking $q = 1$, $s = 1$, $f(t)$ = standard normal p.d.f. and $g(t) = SN(\lambda)$ p.d.f., we have $\int_{-\infty}^{\infty} g(t)e^{|t|}dt \leq 2\int_{-\infty}^{\infty} f(t)e^{|t|}dt < \infty$.

We are now ready to give our main result.

THEOREM 2.1. *Let* X *and* Y *be i.i.d.* F_0, a *given distribution that is uniquely determined by its sequence of moments* $\{\mu_i^0 : i = 1, 2, 3, \ldots\}$ *which all exist. Denote by* G_0 *the distribution of* X^2 *and* Y^2 *and by* H_0 *the distribution of* $\frac{1}{2}(X+Y)^2$. *Let* X_1, X_2 *be i.i.d.* F, *an unspecified distribution with sequence of moments* $\{\mu_i : i = 1, 2, 3, \ldots\}$ *which all exist. Then* $X_1^2 \sim G_0$, $X_2^2 \sim G_0$, *and* $\frac{1}{2}(X_1 + X_2)^2 \sim H_0$ *if and only if* $F(x) = F_0(x)$ *or* $F(x) = \bar{F}_0(x) = 1 - F_0(-x)$.

PROOF. The sufficiency follows directly from the definition of F_0, G_0 and H_0 and by noting that if $X_1 \sim F = \bar{F}_0$ then $-X_1 \sim F_0$.

To prove the necessity, first note that since all moments of F_0 exist and since X and Y are independent, it follows that all moments of G_0 and H_0 exist. Now let X_1, X_2 be i.i.d. F, X, Y be i.i.d. F_0 and define the following for $i = 1, 2, 3, \ldots$

μ_i is the i-th moment of F.

μ_i^0 is the i-th moment of F_0.

η_i is the i-th moment of H_0.

Since $\frac{1}{2}(X_1 + X_2)^2 \sim H_0$, we have

$$E\left[\frac{1}{2}(X_1 + X_2)^2\right]^k = \eta_k = E\left[\frac{1}{2}(X + Y)^2\right]^k \quad \forall k. \tag{2.1}$$

As in Nguyen *et al.* (2000), we will proceed by induction to show that $\mu_{2i+1} = \epsilon\mu^0_{2i+1}$ where ϵ is either $+1$ or -1, i.e., we will show that the odd moments of F are the odd moments of either F_0 or $\bar{F}_0$. The even moments of F coincide with the even moments of F_0 (which are also the even moments of $\bar{F}_0$), i.e.,

$$\mu_{2i} = \mu^0_{2i} \quad \forall i, \tag{2.2}$$

since $X_1^2 \sim G_0$ by the hypothesis and $Y^2 \sim G_0$ if $Y \sim F_0$.

Next, note that either all the odd moments of F_0 are zero or $\exists$ positive odd integer $\hat{j} \geq 1$ such that $\mu^0_{\hat{j}} \neq 0$ and $\mu^0_h = 0$ for positive odd integer $h < \hat{j}$, i.e., $\mu^0_{\hat{j}}$ is the first non-zero odd moment of F_0. We will consider the latter case first.

First, we will show by induction that

$$\mu_h = 0 \quad \text{for} \quad h = 1, 3, 5, \ldots, \hat{j} - 2. \tag{2.3}$$

Taking $k = 1$ in (2.1) and using (2.2), we get $\mu_1 = \epsilon\mu^0_1$. Thus $\mu_1 = 0$ since we are assuming that $\mu^0_1 = 0$. Hence, the induction statement (2.3) is true when $h = 1$.

Now suppose $\mu_k = 0$ for $k = 1, 3, 5, \ldots, 2i-3$ for some i in $\{2, 3, 4, \ldots, \frac{\hat{j}-1}{2}\}$. Again, from (2.1) the equation

$$\sum_{k=0}^{2l} \binom{2l}{k} \mu_k \mu_{2l-k} = \sum_{k=0}^{2l} \binom{2l}{k} \mu^0_k \mu^0_{2l-k} \tag{2.4}$$

holds $\forall$ integer l, because the left hand side is the l-th moment of $(X_1 + X_2)^2$ which is equal to the l-th moment of $(X + Y)^2$, the right hand side of (2.4). Take $l = 2i - 1$. Then (2.4) becomes

$$\begin{aligned} &\binom{4i-2}{2i-1} \mu^2_{2i-1} + \sum_{k=0}^{4i-2} \binom{4i-2}{k} \mu_k \mu_{4i-2-k} \\ &\quad = \binom{4i-2}{2i-1} (\mu^0_{2i-1})^2 + \sum_{k=0}^{4i-2} \binom{4i-2}{k} \mu^0_k \mu^0_{4i-2-k} \end{aligned} \tag{2.5}$$

where the indices on both summations cannot take the value $2i - 1$.

From (2.2), all terms with even moments in the left-hand side cancel with the corresponding terms in the right-hand side. By the induction hypothesis, $\mu_k = 0$ for $k = 1, 3, 5, \ldots, 2i - 3$. Since we are also assuming that $\mu^0_h = 0$ for all odd integer $h < \hat{j}$, then it follows that (2.5) would give $\mu_{2i-1} = 0$. Hence, our induction is complete which proves that $\mu_h = 0 = \mu^0_h$ for all odd $h < \hat{j}$.

Next we will show by induction that $\mu_k = \epsilon\mu^0_k$ for all odd $k \geq \hat{j}$. From (2.4), take $l = \hat{j}$ to get

$$\binom{2\hat{j}}{\hat{j}} \mu^2_{\hat{j}} + \sum_{k=0}^{2\hat{j}} \binom{2\hat{j}}{k} \mu_k \mu_{2\hat{j}-k} = \binom{2\hat{j}}{\hat{j}} (\mu^0_{\hat{j}})^2 + \sum_{k=0}^{2\hat{j}} \binom{2\hat{j}}{\hat{j}} \mu^0_k \mu^0_{2\hat{j}-k}$$

where the indices of the two summations cannot take the value $\hat{j}$.

Since $\mu_h = \mu^0_h = 0 \forall$ odd $h < \hat{j}$ and $\mu_{2i} = \mu^0_{2i} \forall i$, the above equation reduces to $\mu^2_{\hat{j}} = (\mu^0_{\hat{j}})^2$ or equivalently $\mu_{\hat{j}} = \epsilon\mu^0_{\hat{j}}$ where ϵ is either $+1$ or -1.

Now suppose $\mu_k = \epsilon\mu_k^0$ for $k = \hat{j}, \hat{j}+1, \ldots, 2n-3$ for some n in $\{\frac{\hat{j}+3}{2}, \frac{\hat{j}+5}{2}, \frac{\hat{j}+7}{2}, \ldots\}$. Take $l = \frac{2n-1+\hat{j}}{2}$ in (2.4) to get

$$\sum_{k=0}^{2n-1+\hat{j}} \binom{2n-1+\hat{j}}{k} \mu_k \mu_{2n-1+\hat{j}-k} = \sum_{k=0}^{2n-1+\hat{j}} \binom{2n-1+\hat{j}}{k} \mu_k^0 \mu_{2n-1+\hat{j}-k}$$

or equivalently

$$\begin{aligned}
&\left[\binom{2n-1+\hat{j}}{\hat{j}} + \binom{2n-1+\hat{j}}{2n-1}\right] \mu_{\hat{j}} \mu_{2n-1} + \sum_{k=0}^{\hat{j}-1} \binom{2n-1+\hat{j}}{k} \mu_k \mu_{2n-1+\hat{j}-k} \\
&\quad + \sum_{k=\hat{j}+1}^{2n-1+\hat{j}} \binom{2n-1+\hat{j}}{k} \mu_k \mu_{2n-1+\hat{j}-k} \\
&\quad = \left[\binom{2n-1+\hat{j}}{\hat{j}} + \binom{2n-1+\hat{j}}{2n-1}\right] \mu_{\hat{j}}^0 \mu_{2n-1}^0 + \sum_{k=0}^{\hat{j}-1} \binom{2n-1+\hat{j}}{k} \mu_k^0 \mu_{2n-1+\hat{j}-k}^0 \\
&\quad + \sum_{k=\hat{j}+1}^{2n-1+\hat{j}} \binom{2n-1+\hat{j}}{k} \mu_k^0 \mu_{2n-1+\hat{j}-k}^0
\end{aligned}$$

where the indices on all summations cannot take the value $2n-1$.

Because $\mu_{2i} = \mu_{2i}^0 \forall i$, all terms with even moments vanish. Also, since $\mu_h = \mu_h^0 = 0 \forall$ odd $h < \hat{j}$, the first summations in the two sides vanish.

The second summations on the two sides also vanish since for $k = \hat{j}+2, \ldots, 2n-3$, $\mu_k \mu_{2n-1+\hat{j}-k} = \epsilon^2 \mu_k^0 \mu_{2n-1+\hat{j}-k} = \mu_k^0 \mu_{2n-1+\hat{j}}$ by the induction hypothesis that $\mu_k = \epsilon\mu_k^0$ for $k = \hat{j}, \hat{j}+2, \ldots, 2n-3$ and since $\epsilon^2 = 1$. Also, for $k = 2n+1, 2n+3, \ldots, 2n-2+\hat{j}$, $\mu_{2n-1+\hat{j}-k} = 0 = \mu_{2n-1+\hat{j}-k}^0$ since $\mu_h = 0 = \mu_h^0$ for odd $h < \hat{j}$. Thus, for $k = 2n+1, 2n+3, \ldots, 2n-2+\hat{j}$, $\mu_k \mu_{2n-1+\hat{j}-k} = 0 = \mu_k^0 \mu_{2n-1+\hat{j}-k}^0$.

Hence, after all the cancellations, we are left with $\mu_{\hat{j}} \mu_{2n-1} = \mu_{\hat{j}}^0 \mu_{2n-1}^0$. But since $\mu_{\hat{j}} = \epsilon\mu_{\hat{j}}^0 \neq 0$, we get $\mu_{2n-1} = \frac{1}{\epsilon}\mu_{2n-1}^0$. Thus the induction is complete which shows that $\mu_k = \epsilon\mu_k^0$ for all odd $k \geq \hat{j}$.

We have shown that $\mu_h = 0 = \mu_h^0$ for odd $h < \hat{j}$ and $\mu_k = \epsilon\mu_k^0$ for odd $k \geq \hat{j}$ which is equivalent to saying that $\mu_k = \epsilon\mu_k^0$ for all odd k. Since $\mu_i = \mu_i^0$ for all even i and F_0 is uniquely determined by its sequence of moments, it follows that $F = F_0$ or $F = \bar{F}_0$.

The only remaining case we need to consider is the case when all odd moments of F_0 are zero. But the same induction argument we used in proving that $\mu_h = 0 = \mu_h^0$ for odd $h < \hat{j}$ holds by changing only the induction hypothesis $\mu_k = 0$ for $k = 1, 3, 5, \ldots, 2i-3$ for some i in $\{2, 3, 4, \ldots, \frac{\hat{j}-1}{2}\}$ to the new induction hypothesis $\mu_k = 0$ for $k = 1, 3, 5, \ldots, 2i-3$ for some i in $\{2, 3, 4, \ldots\}$.

Thus, $\mu_k = 0$ for all odd k, and again since $\mu_{2i} = \mu_{2i}^0$ for all i, it again follows that $F = F_0$. (Note that in this case $\bar{F}_0 = F_0$ since F_0 is symmetric about the origin.)

The induction proof of Theorem 2.1 is quite involved but it gives two immediate corollaries.

COROLLARY 2.2. *Let X_1, X_2 be i.i.d. F, an unspecified distribution which admits moments of all order. Then $X_1^2 \sim \chi_1^2$, $X_2^2 \sim \chi_1^2$, and $\frac{1}{2}(X_1+X_2)^2 \sim H_0(\lambda)$ if and only if $F = SN(\lambda)$ or $F = SN(-\lambda)$ where $H_0(\lambda)$ is the distribution of $\frac{1}{2}(X+Y)^2$ when X and Y are i.i.d. $SN(\lambda)$.*

PROOF. Take $F_0 = SN(\lambda)$, so $G_0 = \chi_1^2$, $H_0 = H_0(\lambda)$ and $\bar{F}_0 = SN(-\lambda)$. Apply Theorem 2.1 and note that the $SN(\lambda)$ distribution is uniquely determined by its moments by Lemma 2.1.

Corollary 2.2 characterizes the skew-normal distribution based on the distribution of the quadratic statistics X_1^2, X_2^2 and $\frac{1}{2}(X_1+X_2)^2$. As mentioned in the introduction, a similar result is obtained by Roberts and Geisser which characterizes the standard normal distribution. We give their result as another corollary.

COROLLARY 2.3. (Roberts and Geisser (1966)) *Let X_1 and X_2 be i.i.d. random variables from a distribution which admits moments of all order. Then X_1^2, X_2^2 and $\frac{1}{2}(X_1+X_2)^2$ are all χ_1^2 if and only if X_1 and X_2 are both $N(0,1)$ r.v.*

PROOF. The result is obtained from Corollary 2.2 by taking $\lambda = 0$. Alternatively, take $F_0 = N(0,1)$, so that $G_0 = \chi_1^2$, $H_0 = \chi_1^2$ and apply Theorem 2.1.

In Theorem 2.1, we gave a characterization result based on the distribution of the quadratic statistics X_1^2, X_2^2 and $\frac{1}{2}(X_1+X_2)^2$. In the next theorem, we give a characterization based on the distribution of X^2 and $(X+a)^2$ for some constant $a \neq 0$.

THEOREM 2.2. *Let F_0 be a given distribution uniquely determined by its sequence of moments which all exist. Let $Y \sim F_0$. Let G_0 be the distribution of Y^2 and H_0 be the distribution of $(Y+a)^2$ for any constant $a \neq 0$. Let $X \sim F$, an unspecified distribution which admits moments of all order. Then $X^2 \sim G_0$ and $(X+a)^2 \sim H_0$ if and only if $F = F_0$.*

PROOF. The sufficiency follows directly from the definition of F_0, G_0 and H_0. The necessity follows along the same line of argument in the proof of Theorem 2.1, i.e., by induction, we can show that the moments of F coincide with the corresponding moments of F_0.

Like in Theorem 2.1, we immediately get the the following corollaries:

COROLLARY 2.4. *Let $H_0(\lambda)$ be the distribution of $(Y+a)^2$ where $Y \sim SN(\lambda)$ and $a \neq 0$ is a given constant. Let X be a random variable with a distribution that admits moments of all order. Then $X^2 \sim \chi_1^2$, $(X+a)^2 \sim H_0(\lambda)$ if and only if $X \sim SN(\lambda)$ for some λ.*

PROOF. Take $F_0 = SN(\lambda)$, $G_0 = \chi_1^2$ and $H_0 = H_0(\lambda)$. Then apply Theorem 2.2.

COROLLARY 2.5. *Let $a \neq 0$ be a given constant and let X be a random variable with a distribution that admits moments of all order. Then $X^2 \sim \chi_1^2$, $(X+a)^2 \sim \chi_{1,a^2}^2$ if and only if $X \sim N(0,1)$.*

PROOF. Take $F_0 = N(0,1)$, $G_0 = \chi_1^2$ and $H_0 = \chi_{1,a^2}^2$. Then apply Theorem 2.2.

3. Decomposition of the $SN3$ family

In Section 2, we presented characterization results for the skew-normal distribution based on quadratic statistics. In particular, the quadratic statistic $\frac{1}{2}(X_1 + X_2)^2$ was used in Corollary 2.2 for characterizing the skew-normal distribution. It is not difficult to see that when the quadratic statistic $\frac{1}{2}(X_1+X_2)^2$ is replaced by the quadratic statistic $(AX_1 + BX_2)^2$ for some non-zero constants A and B satisfying $A^2 + B^2 = 1$, then the result of Corollary 2.2 will still hold.

Lemma 1.4 shows that a $SN(\lambda)$ distributed random variable can be obtained by a linear combination of two independent random variables whose squares are distributed as χ_1^2. It is interesting to know whether this is true for any random variable whose square is χ_1^2 distributed. For lack of good notation, we will denote the distribution of such a variable by $SN3$. The notation is to reflect the fact that the $SN(\lambda)$ family is a subset of the skew-symmetric family we will denote by $SN2(\lambda)$ whose members have p.d.f. of the form $2F(\lambda z)\phi(z)$, where F is the c.d.f. of an absolutely continuous distribution whose p.d.f. is symmetric about the origin. The $SN2(\lambda)$, briefly discussed in Gupta *et al.* (2002*b*) is in turn a subset of the $SN3$ family. The study of these larger families might shed some light on the $SN(\lambda)$ family. To this end, we have the following result:

THEOREM 3.1. *Let X and Y be two independent random variables whose moments all exist and let A and B be non-zero constants such that $A^2 + B^2 = 1$. Let X^2, Y^2 and $(AX + BY)^2$ all be distributed as χ_1^2. Then*

(i) *at least one of X and Y is standard normal; and*

(ii) *if X and Y are identically distributed, then both X and Y are $N(0,1)$.*

PROOF. Let $W = AX + BY$. Denote by Ψ_Z the characteristic function of an arbitrary random variable Z. Since X^2, Y^2 and W^2 are all distributed as χ_1^2, then from Lemma 1.2, we have

$$\Psi_X(t) + \Psi_X(-t) = \Psi_Y(t) + \Psi_Y(-t) = \Psi_W(t) + \Psi_W(-t) = 2\exp(-t^2/2). \tag{3.1}$$

Also we have, $\Psi_W(t) = \Psi_X(At)\Psi_Y(Bt)$ and $\Psi_W(-t) = \Psi_X(-At)\Psi_Y(-Bt)$. Adding the last two equations and from (3.1) we get

$$\Psi_X(At)\Psi_Y(Bt) + \Psi_X(-At)\Psi_Y(-Bt) = 2\exp(-t^2/2). \tag{3.2}$$

From (3.1), we also get

$$(\Psi_X(At) + \Psi_X(-At))(\Psi_Y(Bt) + \Psi_Y(-Bt)) = 4\exp(-t^2/2). \tag{3.3}$$

Simplifying (3.3) and subtracting (3.2) gives

$$\Psi_X(At)\Psi_Y(-Bt) + \Psi_X(-At)\Psi_Y(Bt) = 2\exp(-t^2/2). \tag{3.4}$$

Equating (3.2) and (3.4) now gives

$$[\Psi_X(At) - \Psi_X(-At)][\Psi_Y(Bt) - \Psi_Y(-Bt)] = 0. \tag{3.5}$$

Let M_X and M_Y be the real part of Ψ_X and Ψ_Y, respectively, and let N_X and N_Y be the imaginary part of Ψ_X and Ψ_Y, respectively. Then, M_X and M_Y are even functions of t, N_X and N_Y are odd functions of t, N_X and N_Y are continuous, and $N_X(0) = N_Y(0) = 0$. Also,

$$\Psi_X(At) = M_X(At) + iN_X(At) \quad \text{and} \quad \Psi_Y(Bt) = M_Y(Bt) + iN_Y(Bt),$$
$$\Psi_X(At) - \Psi_X(-At) = 2iN_X(At)$$

and

$$\Psi_Y(Bt) - \Psi_Y(-Bt) = 2iN_Y(Bt).$$

So, (3.5) is equivalent to $N_X(At)N_Y(Bt) = 0$ which in turn is equivalent to

$$N_X(t)N_Y(Bt/A) = 0. \tag{3.6}$$

We note that this last equation is valid for all t.

Now, since all the odd moments μ_{2k+1}, $k = 1, 2, \ldots$ of X exist, there exists an open interval around 0 with length $\delta_1 > 0$ such that the Taylor series representation

$$N_X(t) = \sum_{k=0}^{\infty} \frac{N_X^{(2k+1)}(0)t^{2k+1}}{(2k+1)!}$$

is valid for all t in this interval. We then have either of the following two cases:

Case 1. If all odd moments of X are zero then X must have a distribution symmetric at the origin.

Case 2. If at least one odd moment of X is nonzero, let μ_{2m+1} be the first nonzero odd moment of X. It follows that the derivative $N_X^{(2m+1)}(0) \neq 0$. This would then imply that there exists an open interval around 0 with length $\delta_2 > 0$ such that $N_X^{(2m+1)}(t) \neq 0$ for all $t \in (-\delta_2, \delta_2)$. Since N_X is an odd function, it must be strictly monotone in this interval implying from (3.6) that $N_Y(Bt/A) = 0$ for all $t \in (-\delta_2, \delta_2)$. It follows that $N_Y(t) = 0$ in an open interval around 0. The last statement implies that all odd moments of Y are 0 and that Y must have a distribution symmetric at the origin.

Suppose without loss of generality that Case 2 holds. Then Y has a symmetric distribution with respect to the origin so that $\Psi_Y(t) = \Psi_Y(-t)$. It therefore follows from Lemma 1.2 that Y must have a standard normal distribution.

Remark 1. Part (ii) of the previous theorem reduces to the necessity part of Corollary 2.3 in the case where $A = B = 1/\sqrt{2}$. The sufficiency part of Corollary 2.3 in this case is well known. It is straight forward to see that the sufficiency part of Corollary 2.3 holds true in the more general case where the only restriction on A and B is the equation $A^2 + B^2 = 1$.

Remark 2. The second part of Theorem 3.1 is exactly what Roberts wanted to show in Roberts (1971). He suggested there that this result may not be true but that he was not able to give a counter-example.

Remark 3. One consequence of Theorem 3.1 is the result that not all random variables with distribution belonging to the $SN3$ family can be decomposed as a linear combination of two independent random variables whose squares are distributed as χ_1^2. To see this, we only need to consider the random variable $Y = |X|$ where $X^2 \sim \chi_1^2$. Clearly $Y^2 \sim \chi_1^2$ so the distribution of Y belongs to the $SN3$ family. If Y can be represented as a linear combination of two independent random variables whose squares are distributed as χ_1^2, then by Theorem 3.1, one of these random variables must be standard normal. This forces the support of Y to be the whole real line which cannot be since the support of Y must be a subset of the positive real line.

To study the decomposition of the $SN2(\lambda)$ family, it might be helpful to look first at the decomposition of the $SN(\lambda)$ distribution. We therefore give the following result:

THEOREM 3.2. *Let A and B be two non-zero constants such that $A^2 + B^2 = 1$ and let $X \sim N(0,1)$ and Y be independent. If $AX + BY \sim SN(\lambda)$ then $Y \sim SN(\operatorname{sign}(\lambda/B)|\lambda|/\sqrt{B^2 + \lambda^2(B^2-1)})$ provided $|\lambda/\sqrt{1+\lambda^2}| \leq |B|$.*

PROOF. Let $W = AX + BY$. Denote by Ψ_Z the characteristic function of an arbitrary random variable Z. Then,

$$\Psi_W(t) = \Psi_X(At)\Psi_Y(Bt). \tag{3.7}$$

From Lemma 1.5, $\Psi_W(t) = \exp(-t^2/2)(1 + i\tau(\delta t))$ where for $x \geq 0$, $\tau(x) = \int_0^x \sqrt{2/\pi}\exp(u^2/2)du$, $\tau(-x) = -\tau(x)$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. Also, $\Psi_X(At) = \exp(-A^2t^2/2)$. Hence, from (3.7) we have

$$\begin{aligned}\Psi_Y(Bt) &= \frac{\exp(-t^2/2)(1+i\tau(\delta t))}{\exp(-A^2t^2/2)} \\ &= \exp(-B^2t^2/2)(1+i\tau(\delta t)).\end{aligned}$$

Thus, replacing t by t/B, we get $\Psi_Y(t) = \exp(-t^2/2)(1 + i\tau(\delta t/B))$ which is the characteristic function of a $SN(\operatorname{sign}(\lambda/B)|\lambda|/\sqrt{B^2 + \lambda^2(B^2-1)})$ random variable provided $|\lambda/\sqrt{1+\lambda^2}| \leq |B|$.

Remark 4. If we take $B = \lambda/\sqrt{1+\lambda^2}$ in Theorem 3.2, we get the result that Y has a half-normal distribution which is suggested by Lemma 1.4.

We end this paper with the following conjecture:

CONJECTURE 3.1. Let A and B be two non-zero constants such that $A^2 + B^2 = 1$ and let $X \sim N(0,1)$ and Y be independent. Let $F(\lambda) \in SN2(\lambda)$. If $AX + BY \sim F(\lambda)$ then under possibly some inequality constraints on B and λ, $Y \sim F(\Delta)$ where Δ is a function of λ and B.

Acknowledgements

We would like to thank the two anonymous referees for their comments and suggestions which greatly improved the presentation of this paper.

References

Arnold, B. and Beaver, R. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting (with discussion), *Test*, **11**, 7–54.

Arnold, B., Beaver, R., Groeneveld, R. and Meeker, W. (1993). The nontruncated marginal of a truncated bivariate normal distribution, *Psychometrika*, **58**, 471–488.

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

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

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distributions, *Journal of the Royal Statistical Society, Series B*, **61**, 579–602.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, *Journal of the Royal Statistical Society, Series B*, **65**, 367–389.

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

Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, **79**, 99–113.

Chiogna, M. (1998). Some results on the scalar skew-normal distribution, *Journal of the Italian Statistical Society*, **1**, 1–13.

Genton, M., He, L. and Liu, X. (2001). Moments of skew-normal random vectors and their quadratic forms, *Statistics & Probability Letters*, **51**, 319–325.

Gupta, A. K. and Chen, T. (2001). Goodness-of-fit tests for the skew-normal distribution, *Communications in Statistics: Simulation and Computation*, **30**, 907–930.

Gupta, A. K. and Huang, W. J. (2002). Quadratic forms in skew-normal variates, *Journal of Mathematical Analysis and Applications*, **273**, 558–564.

Gupta, A. K., Chang, F. C. and Huang, W. J. (2002*a*). Some skew-symmetric models, *Random Operators and Stochastic Equations*, **10**, 133–140.

Gupta, A. K., Nguyen, T. T. and Sanqui, J. A. T. (2002*b*). Characterization of the skew-normal distribution. Tech. Report, No. 02-02, Department of Mathematics and Statistics, Bowling Green State University, Ohio.

Gupta, R. C. and Brown, N. (2001). Reliability studies of the skew-normal distribution and its application to a strength-stress model, *Communications in Statistics: Theory and Methods*, **30**, 2427–2445.

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

Liseo, B. (1990). The skew-normal class of densities: Inferential aspects from a bayesian viewpoint (in Italian), *Statistica*, **50**, 71–82.

Loperfido, N. (2001). Quadratic forms of skew-normal random vectors, *Statistics & Probability Letters*, **54**, 381–387.

Nguyen, T. T., Dinh, K. T. and Gupta, A. K. (2000). A location-scale family generated by a symmetric distribution and the sample variance, Tech. Report, No. 00-15, Department of Mathematics and Statistics, Bowling Green State University, Ohio.

Pewsey, A. (2000*a*). Problems of inference for Azzalini's skew-normal distribution, *Journal of Applied Statistics*, **27**, 859–870.

Pewsey, A. (2000*b*). The wrapped skew-normal distribution on the circle, *Communications in Statistics: Theory and Methods*, **29**, 2459–2472.

Roberts, C. (1971). On the distribution of random variables whose mth absolute power is gamma, *Sankhyā, Ser. A*, **33**, 229–232.

Roberts, C. and Geisser, S. (1966). A necessary and sufficient condition for the square of a random variable to be gamma, *Biometrika*, **53**, 275–277.

Salvan, A. (1986). Locally most powerful invariant tests of normality (in Italian), *Atti della XXXIII Riunione Scientifica della Società Italiana di Statistica*, **2**, 173–179.

Shohat, J. A. and Tamarkin, J. D. (1943). *The Problem of Moments*, American Mathematical Society, New York.