© Statistical Research and Training Center
مجله‌ی پژوهش‌های آماری ایران
J. Statist. Res. Iran **3** (2006): 47-61
دوره‌ی ۳، شماره‌ی ۱، بهار و تابستان ۱۳۸۵، صص ۴۷-۶۱
پژوهشکده‌ی آمار

A New Skew-normal Density

Maryam Sharafi† and Javad Behboodian‡,*

† Shiraz University
‡ Islamic Azad University, Shiraz Branch

Abstract. We present a new skew-normal distribution, denoted by NSN(λ). We first derive the density and moment generating function of NSN(λ). The properties of SN(λ), the known skew-normal distribution of Azzalini, and NSN(λ) are compared with each other. Finally, a numerical example for testing about the parameter λ in NSN(λ) is given.

Keywords. skew-normal distribution; a new skew-normal distribution; moment generating function; skewness; kurtosis; testing hypothesis.

1 Introduction

Azzalini (1985) introduced the following density function by the name of skew-normal density with parameter λ,

$$\varphi(z;\lambda) = 2\varphi(z)\Phi(\lambda z), \qquad z \in \mathbb{R}, \quad \lambda \in \mathbb{R}$$

where $\varphi(z)$ and $\Phi(z)$ are density and distribution functions of standard normal random variable, respectively. We denote a random variable Z_λ with the above density by $Z_\lambda \sim \mathrm{SN}(\lambda)$. This density and its generalization have been studied during the past years. For example, the distribution GSN(λ), given by Gupta and Gupta (2004), is a useful generalization of SN(λ).

* Corresponding author

In this paper we consider the density

$$f(x;\lambda) = c(\lambda)\varphi(x)\Phi^2(\lambda x), \tag{1}$$

where $c(\lambda)$ is given by

$$c(\lambda) = \frac{1}{E\{\Phi^2(\lambda U)\}}, \tag{2}$$

with $U \sim \mathrm{N}(0,1)$. We denote a random variable with this new density by $X_\lambda \sim \mathrm{NSN}(\lambda)$, which is in fact a special case of $\mathrm{GSN}(\lambda)$. Azzalini (1985) showed that the maximum value of skewness for $\mathrm{SN}(\lambda)$ is about 0.995 . The motivation of introducing this new density is the fact that it has a bigger skewness.

In section 2, we use orthant probability to compute $c(\lambda)$ easily without integration and illustrate a real data. Section 3 presents two representation theorems regarding the properties of $\mathrm{NSN}(\lambda)$. The moment generating function, some moments, skewness, and kurtosis of $\mathrm{NSN}(\lambda)$ are given in section 4. In section 5, we compare the properties of $\mathrm{SN}(\lambda)$ and $\mathrm{NSN}(\lambda)$. Finally, in section 6, we generate some data from $\mathrm{NSN}(\lambda)$ and we concentrate on a testing about the parameter λ.

2 Calculation of $\mathbf{c}(\boldsymbol{\lambda})$ by Orthant Probability

An orthant probability is the probability $P(V_1 > 0, V_2 > 0, \ldots, V_n > 0)$ where $\mathbf{V} = (V_1, V_2, \ldots, V_n)$ is a multivariate normal vector with mean 0 and covariance matrix $\boldsymbol{\Sigma} = (\rho_{ij})$, where $\rho_{ii} = 1$ and $\rho_{ij} = \mathrm{cov}(V_i, V_j)$, $i, j = 1, 2, \ldots, n$ (see Kotz et al., 2000).

Now, we write

$$\begin{aligned} E\{\Phi^2(\lambda U)\} &= \int_{-\infty}^{\infty} \Phi^2(\lambda u)\varphi(u)\ du \\ &= \int_{-\infty}^{\infty} P(U_1 \leqslant \lambda u,\ U_2 \leqslant \lambda u)\varphi(u)\ du, \end{aligned}$$

where U_1 and U_2 are i.i.d. $\mathrm{N}(0,1)$ and independent from $U \sim \mathrm{N}(0,1)$. Using the above integral we have

$$\begin{aligned} E\{\Phi^2(\lambda U)\} &= \int_{-\infty}^{\infty} P(U_1 \leqslant \lambda U,\ U_2 \leqslant \lambda U\ |U = u)\varphi(u)\ du \\ &= P(U_1 \leqslant \lambda U,\ U_2 \leqslant \lambda U) \\ &= P(V_1 \geqslant 0,\ V_2 \geqslant 0), \end{aligned}$$

where

$$V_1 = \frac{\lambda U - U_1}{\sqrt{1+\lambda^2}}, \qquad V_2 = \frac{\lambda U - U_2}{\sqrt{1+\lambda^2}},$$

with

$$(V_1, V_2) \sim \mathrm{N}_2(\mathbf{0}, \mathbf{\Sigma}), \qquad \mathbf{\Sigma} = \begin{pmatrix} 1 & \frac{\lambda^2}{1+\lambda^2} \\ \frac{\lambda^2}{1+\lambda^2} & 1 \end{pmatrix}.$$

Hence, we have

$$E\{\Phi^2(\lambda U)\} = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\left(\frac{\lambda^2}{1+\lambda^2}\right).$$

Using the simple trigonometric relation,

$$\frac{1}{4} + \frac{1}{2\pi}\sin^{-1} x = \frac{1}{\pi}\tan^{-1}\sqrt{\frac{1+x}{1-x}}, \qquad -1 < x < 1$$

from (2), we obtain

$$c(\lambda) = \frac{1}{E\{\Phi^2(\lambda U)\}} = \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}. \tag{3}$$

This coefficient is obtained in Arnold et al. (2002) by integration, but our method, by orthant probability, is much easier.

Therefore, density (1) becomes

$$f(x;\lambda) = \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\varphi(x)\Phi^2(\lambda x). \tag{4}$$

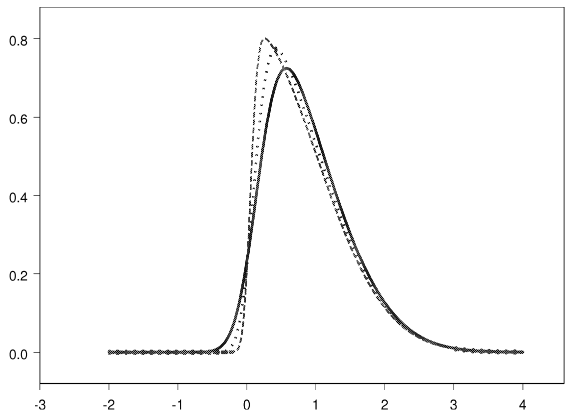

Figure 1. The density of NSN(λ) for $\lambda = 3$ (soiled line), $\lambda = 5$ (dotted line), and $\lambda = 10$ (dashed line).

Table 1. MLE's parameters for IQ score data

Distribution	$\mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$	$\mathbf{SN}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda})$	$\mathbf{NSN}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda})$
$\widehat{\mu}$	106.653	98.79	94.88
$\widehat{\sigma}$	8.23	11.38	13.07
$\widehat{\lambda}$	——	1.71	2.09
Log-likelihood	−183.387	−182.436	−182.206

Figure 1 shows the shape of $f(x;\lambda)$ for some values of λ.

The location-scale of NSN(λ) is defined as that of $Y = \mu + \sigma X_\lambda$, where $\mu \in \mathbb{R}$, $\sigma > 0$, and $X_\lambda \sim \mathrm{NSN}(\lambda)$. Its density is given by

$$f(y;\theta) = \left(\frac{\pi}{\sigma \tan^{-1}\sqrt{1+2\lambda^2}}\right) \varphi\left(\frac{y-\mu}{\sigma}\right) \Phi^2\left(\lambda \cdot \frac{y-\mu}{\sigma}\right),$$

where $\theta = (\mu, \sigma, \lambda)$. We denote this by $Y \sim \mathrm{NSN}(\mu, \sigma, \lambda)$.

We now fit this new distribution on the real IQ score data for 52 non-white males given by Gupta and Brown (2001). Table 1 and Figure 2 show that our distribution better fits the data comparing with the normal and Azzalini's distributions.

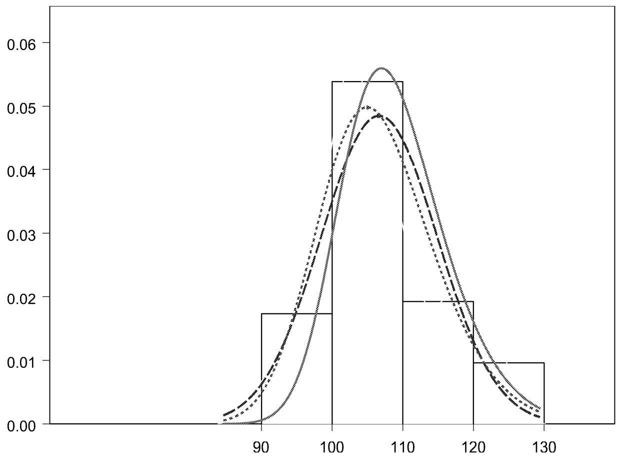

Figure 2. Histogram of 52 Otis IQ Scores. The lines represent distributions fitted using maximum likelihood estimation: $\mathrm{N}(\hat{\mu}, \hat{\sigma}^2)$ (dotted line), $\mathrm{SN}(\hat{\mu}, \hat{\sigma}, \hat{\lambda})$ (dashed line), and $\mathrm{NSN}(\hat{\mu}, \hat{\sigma}, \hat{\lambda})$ (soiled line).

3 Two Representation Theorems about NSN($\boldsymbol{\lambda}$)

Azzalini and Dalla-Valle (1996) define the density of the bivariate skew-normal vector $(Y_1, Y_2)^{\mathrm{T}} \sim \mathrm{SN}_2(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ by

$$\varphi_2(y_1, y_2; \rho, \boldsymbol{\alpha}) = 2\varphi(y_1, y_2; \rho)\Phi(\alpha_1 y_1 + \alpha_2 y_2),$$

where $\varphi(y_1, y_2; \rho)$ is the standard bivariate normal density with

$$\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \qquad |\rho| < 1,$$

and

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

By the following representation theorems, we relate X_λ with normal and bivariate skew-normal variables.

Theorem 1 (first representation theorem). *If* $(Y_1, Y_2)^{\mathrm{T}} \sim \mathrm{SN}_2(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ *with* $\alpha = \left(\rho/\sqrt{1-\rho^2}, 0\right)^{\mathrm{T}}$*, then*

$$(Y_1 | Y_2 > 0) \overset{\mathrm{D}}{=} X_\lambda \sim \mathrm{NSN}(\lambda), \qquad \lambda = \frac{\rho}{\sqrt{1-\rho^2}},$$

where $\overset{\mathrm{D}}{=}$ *denotes equality in distribution.*

Proof.

$$\begin{aligned} P(Y_1 \leqslant x | Y_2 > 0) &= \frac{P(Y_1 \leqslant x, Y_2 > 0)}{P(Y_2 > 0)} \\ &= \frac{2\int_{-\infty}^{x}\int_{0}^{\infty} \varphi(y_1, y_2; \rho)\Phi(\lambda y_1)\ dy_2 dy_1}{P(Y_2 > 0)} \\ &= \frac{2\int_{-\infty}^{x} \varphi(y_1)\Phi^2(\lambda y_1)\ dy_1}{P(Y_2 > 0)}. \end{aligned} \tag{5}$$

On the other hand, by Azzalini and Capitanio (1999),

$$Y_2 \sim \mathrm{SN}\left(\frac{\lambda^2}{\sqrt{1+2\lambda^2}}\right),$$

and by Gupta and Brown (2001),

$$P(Y_2 > 0) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\frac{\lambda^2}{\sqrt{1+2\lambda^2}}.$$

Now, using the following trigonometric relation for $x \in \mathbb{R}$,

$$\frac{1}{2} + \frac{1}{\pi} \tan^{-1} \frac{x^2}{\sqrt{1+2x^2}} = \frac{2}{\pi} \tan^{-1} \sqrt{1+2x^2},$$

we have

$$P(Y_2 > 0) = \frac{2}{\pi} \tan^{-1} \sqrt{1+2\lambda^2}. \tag{6}$$

From (5) and (6), we obtain

$$P(Y_1 \leqslant x | Y_2 > 0) = \int_{-\infty}^{x} \frac{\pi}{\tan^{-1} \sqrt{1+2\lambda^2}} \varphi(y_1) \Phi^2 (\lambda y_1) \; dy_1, \tag{7}$$

which is the distribution of X_λ by (4). This completes the proof.

Theorem 2 (second representation theorem). *If* $(U_1, U_2, U_3) \sim \mathrm{N}_3(\mathbf{0}, \mathbf{\Sigma})$ *with*

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho^2 \\ \rho & \rho^2 & 1 \end{pmatrix},$$

then

$$(U_1 | \min\{U_2, U_3\} > 0) \stackrel{\mathrm{D}}{=} X_\lambda \sim \mathrm{NSN}(\lambda), \qquad \lambda = \frac{\rho}{\sqrt{1-\rho^2}}.$$

Proof. Let $U = \min\{U_2, U_3\}$ and compute the joint density of U_1 and U. Using the fact that the pair (U_2, U_3) is exchangeable and $(U_1, U_2) \stackrel{\mathrm{D}}{=} (U_1, U_3)$, we have

$$\begin{aligned} P(U_1 \leqslant u_1,\ U \leqslant u_2) &= P(U_1 \leqslant u_1,\ U_2 \leqslant u_2 | U_2 \leqslant U_3) P(U_2 \leqslant U_3) \\ &\quad + P(U_1 \leqslant u_1,\ U_3 \leqslant u_2 | U_3 \leqslant U_2) P(U_3 \leqslant U_2) \\ &= P(U_1 \leqslant u_1,\ U_2 \leqslant u_2 | U_0 \geqslant 0), \end{aligned} \tag{8}$$

where

$$U_0 = \frac{U_3 - U_2}{\sqrt{2\,(1-\rho^2)}}.$$

On the other hand, $(U_0, U_1, U_2) \sim \mathrm{N}_3(\mathbf{0}, \mathbf{\Sigma}^*)$ with

$$\mathbf{\Sigma}^* = \begin{pmatrix} 1 & 0 & -\sqrt{\frac{1-\rho^2}{2}} \\ 0 & 1 & \rho \\ -\sqrt{\frac{1-\rho^2}{2}} & \rho & 1 \end{pmatrix}.$$

By Azzalini and Dalla-Valle (1996)

$$((U_1, U_2)|U_0 > 0) \sim \mathrm{SN}_2(\boldsymbol{\Omega}, \boldsymbol{\alpha}), \tag{9}$$

where

$$\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \frac{\rho}{\sqrt{1-\rho^2}} \\ \\ \frac{-1}{\sqrt{1-\rho^2}} \end{pmatrix}.$$

Now, from (8) and (9) we have

$$(U_1, U) \sim \mathrm{SN}_2(\Omega, \alpha).$$

Then, by the approach of Theorem 1, we conclude that

$$(U_1 | \min\{U_2, U_3\} > 0) \overset{\mathrm{D}}{=} X_\lambda.$$

4 Moment Generating Function of NSN($\boldsymbol{\lambda}$)

In this section, we find the moment generating function (m.g.f.) of X_λ which has density (4).

Theorem 3 *The m.g.f. of X_λ is*

$$M_{X_\lambda}(t) = \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}} \ \exp\left\{\frac{t^2}{2}\right\} \Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}; \frac{1}{\sqrt{1+2\lambda^2}}\right), \tag{10}$$

where $\Phi_1(z;\theta)$ is the distribution function of $Z_\theta \sim \mathrm{SN}(\theta)$, given by Azzalini (1985), as follows.

$$\Phi_1(z;\theta) = \Phi(z) - 2\int_z^\infty \int_0^{\theta w} \varphi(u)\varphi(w)\ du\ dw. \tag{11}$$

Proof. Using density (4) and the change of variable $x - t = u$, we have

$$\begin{aligned} M_{X_\lambda}(t) = E(\exp\{tX_\lambda\}) &= \int_{-\infty}^{\infty} \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}} \exp\{tx\}\varphi(x)\Phi^2(\lambda x)\ dx \\ &= \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}} \ \exp\left\{\frac{t^2}{2}\right\} E\{\Phi^2(\lambda U + \lambda t)\}, \end{aligned}$$

with $U \sim \mathrm{N}(0,1)$. We can show that

$$E\{\Phi^2(\lambda U + \lambda t)\} = \Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}; \frac{1}{\sqrt{1+2\lambda^2}}\right)$$

(see Appendix A1). Therefore, we have (10).

4.1 Moments of $\mathbf{X_\lambda}$

We can find expectation, variance, and third and forth moments of X_λ by taking derivative from $M_{X_\lambda}(t)$ with respect to t. The results are

$$\mu_\lambda = E(X_\lambda) = \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}} \cdot \frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}, \tag{12}$$

$$\sigma_\lambda^2 = \mathrm{var}(X_\lambda) = 1 + \frac{\lambda^2}{(1+\lambda^2)\sqrt{1+2\lambda^2}\tan^{-1}\sqrt{1+2\lambda^2}} - \mu_\lambda^2. \tag{13}$$

To find skewness and kurtosis of X_λ, we have

$$E(X_\lambda^3) = \left(\frac{3+2\lambda^2}{1+\lambda^2}\right)\mu_\lambda, \tag{14}$$

$$E(X_\lambda^4) = 3E(X_\lambda^2) + \frac{\lambda^2(5\lambda^2+3)}{(1+\lambda^2)^2(1+2\lambda^2)^{\frac{3}{2}}\tan^{-1}\sqrt{1+2\lambda^2}} \tag{15}$$

(see Appendix A2). Using (12)-(15) and some computation we obtain

$$3 < \text{kurtosis of } X_\lambda = \frac{E(X_\lambda - \mu_\lambda)^4}{\sigma_\lambda^4} < 3.8692$$

$$-5.6330 < \text{skewness of } X_\lambda = \frac{E(X_\lambda - \mu_\lambda)^3}{\sigma_\lambda^3} < 5.6330$$

We may consider the more general skew-normal density

$$f(x;\lambda) = c(\lambda)\varphi(x)\Phi^n(\lambda x), \qquad n = 1, 2, \ldots$$

by the name of skew-normal density of order n with distribution GSN(λ), which was introduced by Gupta and Gupta (2004). More Properties of this density are studied by Sharafi and Behboodian (2006). The result of Theorem 3 can also be obtained from this general density.

5 Comparison of SN($\boldsymbol{\lambda}$) and NSN($\boldsymbol{\lambda}$)

In this section, we discuss some properties of $Z_\lambda \sim \mathrm{SN}(\lambda)$ and $X_\lambda \sim \mathrm{NSN}(\lambda)$.

1. For $\lambda = 0$, $X_\lambda \overset{\mathrm{D}}{=} Z_\lambda \sim \mathrm{N}(0,1)$.
2. As $\lambda \to \pm\infty$, the densities of X_λ and Z_λ go to the half-normal density, i.e., the density of $|U|$ with $U \sim \mathrm{N}(0,1)$.

3. $-X_\lambda \sim \text{NSN}(-\lambda)$ and $-Z_\lambda \sim \text{SN}(-\lambda)$.

4. X_λ and Z_λ are both strongly unimodal (see Karlin, 1968).

5. Skewness of X_λ and Z_λ is positive for $\lambda > 0$ and negative for $\lambda < 0$.

6. $E(X_\lambda^k) \geqslant E(Z_\lambda^k) \geqslant 0 \quad \text{for} \quad \lambda \geqslant 0, \quad k = 1, 3$
   $E(X_\lambda^k) \leqslant E(Z_\lambda^k) \leqslant 0 \quad \text{for} \quad \lambda \leqslant 0, \quad k = 1, 3.$

7. $E(X_\lambda^2) \geqslant E(Z_\lambda^2) \quad \text{for} \quad \lambda \in \mathbb{R}.$

8. $\text{var}(X_\lambda) \leqslant \text{var}(Z_\lambda) \quad \text{for} \quad \lambda \in \mathbb{R}.$

9. skewness of $X_\lambda \geqslant$ skewness of $Z_\lambda \geqslant 0$ for $\lambda \geqslant 0$
   skewness of $X_\lambda \leqslant$ skewness of $Z_\lambda \leqslant 0$ for $\lambda \leqslant 0$.

10. kurtosis of $X_\lambda \geqslant$ kurtosis of Z_λ for $\lambda \in \mathbb{R}$.

11. $P(X_\lambda \geqslant 0) \geqslant P(Z_\lambda \geqslant 0) \quad \text{for} \quad \lambda \geqslant 0$
    $P(X_\lambda \geqslant 0) \leqslant P(Z_\lambda \geqslant 0) \quad \text{for} \quad \lambda \leqslant 0.$

(see Appendix A3).

6 Hypothesis Testing about $\boldsymbol{\lambda}$

Let $X_1, X_2, \ldots, X_n$ be i.i.d. from NSN(λ). We want to test $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda \neq \lambda_0$. It is easy to show that

$$T = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{\Phi(\lambda X_i)}$$

is a decreasing function of λ. Using

$$E(Z_\lambda) = \sqrt{\frac{2}{\pi}} \left(\frac{\lambda}{\sqrt{1+\lambda^2}} \right)$$

(see Azzalini, 1985), we have

$$\mu_T = E(T) = \sqrt{\frac{\pi}{2}} \left(\frac{\lambda}{\sqrt{1+\lambda^2}} \right) \left(\frac{1}{\tan^{-1} \sqrt{1+2\lambda^2}} \right),$$

$$\sigma_T^2 = \text{var}(T) = \frac{\pi}{n \tan^{-1} \sqrt{1+2\lambda^2}} \left(1 - \frac{\lambda^2}{2(1+\lambda^2) \tan^{-1} \sqrt{1+2\lambda^2}} \right).$$

(see Appendix A4). For large n, we have approximately (under H_0)

$$Z = \frac{T - \mu_T}{\sigma_T} \sim \mathrm{N}(0, 1).$$

Therefore, the critical region for the above test at the level α is $|Z| > z_{\alpha/2}$, where $P(Z > z_{\alpha/2}) = \alpha/2$.

We observe that this statistic is simpler than the similar statistic used by Gupta and Gupta (2004). Using the following generated data from NSN(1) with size 60, we want to test $H_0 : \lambda = 1$ versus $H_1 : \lambda \neq 1$. Under H_0 we obtain

$$t = 0.96331, \quad \mu_T = 0.846284, \quad \sigma_T = 0.195198,$$

$$z = 0.599830, \quad p\text{-value} = 2P(Z > 0.599830) = 0.54862$$

Therefore H_0 is not rejected, and the power of the test for $\lambda = 0$ is 0.9638 at level $\alpha = 0.05$.

1.01597	1.16863	0.29148	1.00454	0.29135	1.80183
0.44544	−0.00673	1.81410	−0.57200	1.47283	0.04330
2.46565	−0.03888	1.13157	0.44446	0.64613	1.13833
1.49379	0.00786	0.54528	2.12881	1.58779	1.32810
2.05071	0.56590	0.70947	1.60860	1.70980	1.36679
1.19152	−0.01076	−0.19823	0.58204	−0.02293	0.27317
1.96351	−0.09878	0.46880	1.60463	0.48174	1.48968
0.68240	1.46606	0.61545	1.46024	0.56457	1.53633
0.10162	0.14067	0.48807	0.59445	0.58263	1.01765
0.45428	0.89725	2.09898	0.22074	0.18980	1.43234

Acknowledgement

The authors sincerely thank the editor and referees for their valuable comments and suggestions.

References

Arellano, R.B.; Gomez, H.W.; Quintana, F.A. (2004). A new class of skew-normal distribution, *Comm. Statist. Theory Methods* **33**, 1465-1480.

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

Azzalini, A. (1985). A class of distributions with includes the normal ones. *Scand. J. Statist.* **12**, 171-178.

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

Azzalini, A.; Capitanio, A. (1999). Statitical applications of the multivarite skew-normal distributions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **61**, 579-602.

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

Branco, M.; Dey, D. (2001). A general class of multivariate elliptical distributions. *J. Multivariate Anal.* **79**, 99-113.

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

Gupta, R.C.; Gupta, R.D. (2004). Generalized skew normal density. *Test* **13**, 501-524.

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

Karlin, S. (1968). *Total Positivity,* vol. 1. Stanford University Press, Stanford, CA.

Kotz, S.; Balakrishnan, N.; Johnson, N.L. (2000). *Continuous Multivariate Distributions,* vol. 1, 2nd ed. Wiley, New York.

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

Sharafi, M.; Behboodian, J. (2006). The Balakrishnan skew-normal density. *Statist. Papers,* to appear.

Appendix A1

Let $\Phi_1(z;\theta)$ *be the distribution function of* $Z_\theta \sim \mathrm{SN}(\theta)$, *and* $U \sim \mathrm{N}(0,1)$. *Then*

$$E\{\Phi^2(\lambda U+\lambda t)\} = \Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right) \tag{16}$$

Proof. Consider U, U_1, U_2 i.i.d $\mathrm{N}(0,1)$,

$$\begin{aligned}
E\{\Phi^2(\lambda U+\lambda t)\} &= \int_{-\infty}^{\infty} \Phi^2(\lambda u+\lambda t)\varphi(u)\ du \\
&= \int_{-\infty}^{\infty} P(U_1 \leqslant \lambda U+\lambda t,\ U_2 \leqslant \lambda U+\lambda t|U=u)\varphi(u)\ du \\
&= P(U_1 \leqslant \lambda U+\lambda t,\ U_2 \leqslant \lambda U+\lambda t) \\
&= P\left(W_1 \leqslant \frac{\lambda t}{\sqrt{1+\lambda^2}},\ W_2 \leqslant \frac{\lambda t}{\sqrt{1+\lambda^2}}\right) \\
&= P\left\{\max(W_1,W_2) \leqslant \frac{\lambda t}{\sqrt{1+\lambda^2}}\right\},
\end{aligned}$$

where $W_i = (U_i - \lambda U)/\sqrt{1+\lambda^2}$, $i=1,2$, and $(W_1,W_2) \sim \mathrm{N}_2\left(0,0,1,1,\frac{\lambda^2}{1+\lambda^2}\right)$. By Loperfido (2001) we have

$$\max\{W_1,W_2\} \sim \mathrm{SN}\left(\frac{1}{\sqrt{1+2\lambda^2}}\right).$$

Thus, we have (16).

Appendix A2

Proof of the formulas (12) and (14): We know that

$$E(X_\lambda^k) = \frac{d^k}{dt^k} M_{X_\lambda}(t)|t=0.$$

Therefore

$$\begin{aligned}
E(X_\lambda) &= \frac{d}{dt} M_{X_\lambda}(t)\Big|_{t=0} \\
&= \frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\Bigg[t\exp\left\{\frac{t^2}{2}\right\}\Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right) \\
&\quad + \exp\left\{\frac{t^2}{2}\right\}\frac{d}{dt}\Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right)\Bigg]\Bigg|_{t=0}.
\end{aligned}$$

Because of

$$\frac{d}{dt}\Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right)=\frac{\lambda}{\sqrt{1+\lambda^2}}2\varphi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right)\Phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}\sqrt{1+2\lambda^2}}\right),$$

we have

$$\mu_\lambda=E(X_\lambda)=\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\right).$$

Therefore the formula (12) is proved.

For $E(X_\lambda^3)$, by taking derivative and some calculations we obtain

$$\begin{aligned}
E(X_\lambda^3)&=\frac{d^3}{dt^3}M_{X_\lambda}(t)\big|_{t=0}\\
&=3\frac{d}{dt}\Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right)+\frac{d^3}{dt^3}\Phi_1\left(\frac{\lambda t}{\sqrt{1+\lambda^2}};\frac{1}{\sqrt{1+2\lambda^2}}\right)\Bigg|_{t=0}\\
&=3\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\right)\\
&\quad-\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{\lambda^3}{\sqrt{2\pi}(1+\lambda^2)\sqrt{1+\lambda^2}}\right)\\
&=\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\right)\left(3-\frac{\lambda^2}{1+\lambda^2}\right)\\
&=\left(\frac{3+2\lambda^2}{1+\lambda^2}\right)\mu_\lambda.
\end{aligned}$$

Appendix A3

Proof of some properties (6)-(11):

Proof of (6). For $k=1$ (see Azzalini, 1985),

$$\begin{aligned}
E(X_\lambda)-E(Z_\lambda)&=\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\right)-\sqrt{\frac{2}{\pi}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)\\
&=\sqrt{\frac{2}{\pi}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)\left(\frac{\pi}{2\tan^{-1}\sqrt{1+2\lambda^2}}-1\right).
\end{aligned}$$

Because $\left\{\pi/\left(2\tan^{-1}\sqrt{1+2\lambda^2}\right)\right\}-1\geqslant 0$ for $\lambda\in\mathbb{R}$, we can see that

$$E(X_\lambda)\geqslant E(Z_\lambda)\geqslant 0\quad\text{for}\quad\lambda\geqslant 0,$$

and

$$E(X_\lambda)\leqslant E(Z_\lambda)\leqslant 0\quad\text{for}\quad\lambda\leqslant 0.$$

Now consider $k = 3$. By (14) and Azzalini (1985),

$$\begin{aligned} E(X_\lambda^3) - E(Z_\lambda^3) &= \left(\frac{3+2\lambda^2}{1+\lambda^2}\right) E(X_\lambda) - \left(\frac{3+2\lambda^2}{1+\lambda^2}\right) E(Z_\lambda) \\ &= \left(\frac{3+2\lambda^2}{1+\lambda^2}\right) \{E(X_\lambda) - E(Z_\lambda)\}. \end{aligned}$$

By the above result for $E(X_\lambda) - E(Z_\lambda)$, we obtain (6) for $k = 3$.

Proof of (7). By (13) and Azzalini (1985),

$$\begin{aligned} E(X_\lambda^2) - E(Z_\lambda^2) &= 1 + \left\{\frac{\lambda^2}{(1+\lambda^2)\sqrt{1+2\lambda^2}\tan^{-1}\sqrt{1+2\lambda^2}}\right\} - 1 \\ &= \frac{\lambda^2}{(1+\lambda^2)\sqrt{1+2\lambda^2}\tan^{-1}\sqrt{1+2\lambda^2}} \geqslant 0. \end{aligned}$$

Therefore $E(X_\lambda^2) \geqslant E(Z_\lambda^2)$ for $\lambda \in \mathbb{R}$.

Appendix A4

Proof of μ_T and σ_T^2:

$$\begin{aligned} \mu_T &= E(T) \\ &= E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{X_i}{\Phi(\lambda X_i)}\right) \\ &= E\left(\frac{X}{\Phi(\lambda X)}\right) \\ &= \int \left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{x}{\Phi(\lambda x)}\right)\varphi(x)\Phi^2(\lambda x)\ dx \\ &= \left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right) E(Z_\lambda) \\ &= \sqrt{\frac{\pi}{2}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)\left(\frac{1}{\tan^{-1}\sqrt{1+2\lambda^2}}\right). \end{aligned}$$

The variance of T is

$$
\begin{aligned}
\sigma_T^2 &= \mathrm{var}(T) \\
&= \frac{1}{n}\mathrm{var}\left(\frac{X}{\Phi(\lambda X)}\right) \\
&= \frac{1}{n}\left\{\int\left(\frac{\pi}{\tan^{-1}\sqrt{1+2\lambda^2}}\right)\left(\frac{x^2}{\Phi^2(\lambda x)}\right)\varphi(x)\Phi^2(\lambda x)dx - \mu_T^2\right\} \\
&= \frac{\pi}{n\tan^{-1}\sqrt{1+2\lambda^2}}(1-\mu_T^2) \\
&= \frac{\pi}{n\tan^{-1}\sqrt{1+2\lambda^2}}\left(1-\frac{\lambda^2}{2(1+\lambda^2)\tan^{-1}\sqrt{1+2\lambda^2}}\right).
\end{aligned}
$$

Maryam Sharafi
Department of Statistics,
Shiraz University,
Shiraz, Iran.
e-mail: *msharafi@susc.ac.ir*

Javad Behboodian
Department of Mathematics,
Islamic Azad University, Shiraz Branch
Shiraz, Iran.
e-mail: *behboodian@stat.susc.ac.ir*