METRON (2017) 75:87–107
DOI 10.1007/s40300-017-0106-2

A flexible extension of skew generalized normal distribution

Mahdi Rasekhi[1] · **G. G. Hamedani**[2] · **Rahim Chinipardaz**[3]

Received: 29 February 2016 / Accepted: 22 February 2017 / Published online: 6 March 2017
© Sapienza Università di Roma 2017

Abstract We introduce an extension of the skew generalized normal distribution called shape-skew generalized normal distribution. The proposed distribution has certain type of flexibility which is different from those given in other flexible skew normal distributions. It possesses properties such as uni/bimodality, skewness, wider range of the Pearson's excess kurtosis coefficient (γ_2) with respect to skew generalized normal distribution and preserving the most desirable features of the skew generalized normal distribution. Some basic distributional properties of the new extension including moments, moment generating function, characterization and relation to other distributions are derived. Also, the multivariate case of our proposed distribution is introduced and some of its properties are studied. The suitability of our model is demonstrated via comparisons with other generalized models.

Keywords Skew-symmetric distributions · Shape parameter · Skewing function · Moment · Stochastic representation · Skewness

✉ Mahdi Rasekhi
rasekhimahdi@gmail.com

G. G. Hamedani
g.hamedani@mu.edu

Rahim Chinipardaz
chinipardaz_r@scu.ac.ir

1 Department of Statistics, Faculty of Mathematical Sciences and Statistics, Malayer University, Malayer, Iran

2 Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, USA

3 Department of Statistics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Iran

1 Introduction

Various types of skew-symmetric distributions have been proposed by many researchers in the literature. In general, there are four methods of constructing a skew-symmetric distribution with a symmetric density function [17]. One of these methods is perturbation of a symmetric density via skewing function, i.e. a skew probability density function (pdf) is created by multiplying a symmetric pdf with a skewing function. A skewing function is a function with range [0,1]. In fact, the starting point of all these studies was the skew-normal (SN) distribution introduced by [5],

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x), \quad x \in \mathbb{R}, \tag{1}$$

where ϕ and Φ are the pdf and cumulative distribution function (cdf) of the standard normal, respectively. A random variable X with the above density is denoted by $X \sim SN(\lambda)$. [1] introduced a generalization of (1) with nice properties, called skew generalized normal distribution (SGN) with pdf of the form

$$f(x;\lambda_1,\lambda_2) = 2\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad x \in \mathbb{R}. \tag{2}$$

Skew-curved normal distribution (SCN) is a SGN distribution with parameter $\lambda_2 = \lambda_1^2$. A number of researchers proposed extension of this density such as [10,12,15,22]. Choudhury and Abdul Matin [10] added one parameter to SGN family and called it, extended skew generalized normal (ESGN) distribution with following density

$$f(x;\lambda_1,\lambda_2,\lambda_3) = 2\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{\lambda_2 x^2+\lambda_3 x^4}}\right). \tag{3}$$

They showed that ESGN distribution is more flexible since the range of Pearson's excess kurtosis coefficient of ESGN distribution is wider than those of SN and SGN distributions. Certain studies was done for creating flexibility (bimodality) in the skew-symmetric family of distributions by [2,3,9,11,14,16,18–20].

In this paper, an extension of SGN distribution is introduced by adding a shape parameter. The addition of this parameter make our proposed distribution to be uni-bimodal and have a wider range of Pearson's excess kurtosis coefficient. Also, the multivarite version of our distribution with multimodal shape is introduced.

The rest of the paper is organized as follows. In Sect. 2, we present the definition of our proposed distribution and various graphs of its pdf. We also derive some important results about this distribution and its relationship with other distributions. The main properties of our proposed model such as moments, stochastic representation and characterizations are also discussed in this section. Section 3 is devoted to maximum likelihood estimation. In Sect. 4, we introduce the multivarite case of our proposed distribution and study some of its properties. Finally, in Sect. 5, we use two real data sets to illustrate the usefulness of this family of distributions.

2 The shape-skew generalized normal distribution and its main properties

In this section, we introduce a flexible class of skew normal distributions generalizing (2).

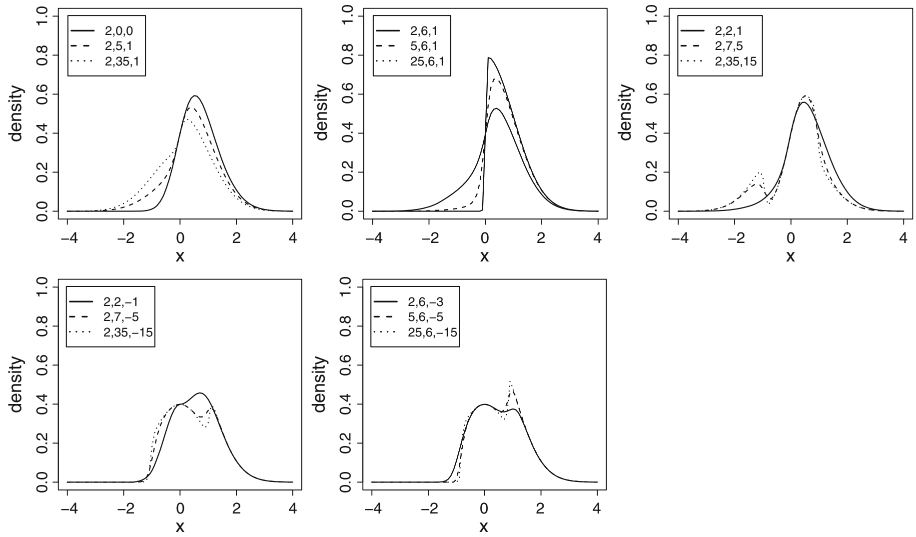

Fig. 1 Some possible shapes of $SSGN(\lambda_1, \lambda_2, \alpha)$ distribution by different parameters

Definition 1 The random variable X has shape skew generalized normal distribution if its density is given by

$$f(x; \lambda_1, \lambda_2, \alpha) = 2\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2|x|^{2\alpha}}}\right) \quad x \in \mathbb{R}, \tag{4}$$

where $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in [0, \infty)$ are skewing parameters, $\alpha \in \mathbb{R} - \{0\}$ is a shape parameter with the following conditions: if $\lambda_1 = 0$ then λ_2 and α must be zero and one, respectively and if $\lambda_2 = 0$ then $\alpha = 1$. We denote this by $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$. The resulting distribution for the special case $\lambda_2 = \lambda_1^2$ is called shape skew-curved normal (SSCN) and is denoted by $SSCN(\lambda_1, \alpha)$.

We like to point out that (4) is indeed a density, due to the fact that skewing function is constructed based on [17] conditions for skewing function (See page 2). This condition is presented as follows:

A skewing function is a mapping $\Pi : \mathbb{R}^k \times \mathbb{R}^k \to [0, 1]$ such that

$$\Pi(-\mathbf{z}, \delta) + \Pi(\mathbf{z}, \delta) = 1\,\forall \mathbf{z}, \delta \in \mathbb{R}^k,\, \Pi(\mathbf{z}, \delta^*) = \frac{1}{2}\,\forall \mathbf{z} \in \mathbb{R}^k, \tag{5}$$

where δ^* is special case of δ. The parameter δ is a skewness/asymmetry parameter and the normalizing constant equals 2. If α is zero formula (4) is not a density.

Figure 1 illustrates the various graphs of (4) under different choices of $\lambda_1, \lambda_2, \alpha$ which shows that SSGN density can change to uni/bimodality, high and low Pearsons excess kurtosis coefficient and heavy tail shape taking different parameters. To see the modality behavior of a SSGN distribution, we used some graphical methods and observed that the derivative of the density (4) changes sign at most once from positive to negative when $\alpha \in \{-1, 1\}$ and changes sign two more times when $\alpha \notin \{-1, 1\}$. Therefore, the distribution in question is either unimodal or bimodal. Figure 2 shows the effect of α on $SSCN(\lambda, \alpha)$ and it is compared with $SCN(\lambda)$, $(\alpha = 1)$ introduced by the [1]. We see that for the positive shape parameter,

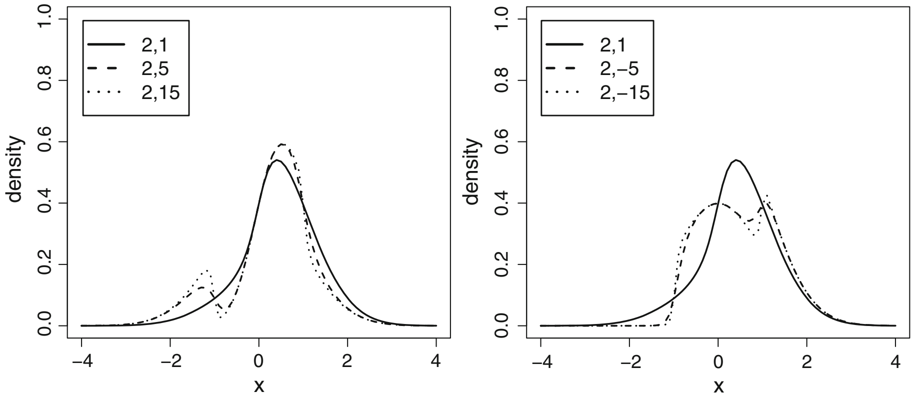

Fig. 2 Effect of positive or negative shape parameter on $SSCN(\lambda_1, \alpha)$ distribution

the density tends to be more heavy tail and bimodal. On the other hand for the negative shape parameter, mode of SCN density divides into two modes.

Basic properties of a $SSGN(\lambda_1, \lambda_2, \alpha)$:

Proposition 1 *If* $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$ *then we have:*

1. $SSGN(0, 0, 1) = N(0, 1)$
2. $SSGN(\lambda_1, 0, 1) = SN(\lambda_1)$*, for all* $\lambda_1 \in \mathbb{R}$
3. $SSGN(\lambda_1, \lambda_2, 1) = SGN(\lambda_1, \lambda_2)$*, for all* $\lambda_1 \in \mathbb{R}$*,* $\lambda_2 \geq 0$
4. $SSGN(\lambda_1, \lambda_2, 2) = ESGN(\lambda_1, 0, \lambda_2)$*, for all* $\lambda_1 \in \mathbb{R}$*,* $\lambda_2 \geq 0$.
5. $-X \sim SSGN(-\lambda_1, \lambda_2, \alpha)$
6. $f(x, \lambda_1, \lambda_2, \alpha) + f(-x, \lambda_1, \lambda_2, \alpha) = 2\phi(x)$*, for all* $x \in \mathbb{R}, \lambda_1 \in \mathbb{R}$*,* $\lambda_2 \geq 0$*,* $\alpha \in \mathbb{Z} - \{0\}$.
7. $\lim_{\lambda_1 \to \infty} f(x, \lambda_1, \lambda_2, \alpha) = 2\phi(x)I(x \geq 0)$ (*half normal distribution*)*, for all* $\lambda_2 \geq 0$*,* $\alpha \in \mathbb{Z} - \{0\}$.
8. $\lim_{\lambda_1 \to -\infty} f(x, \lambda_1, \lambda_2, \alpha) = 2\phi(x)I(x \leq 0)$ (*half normal distribution*)*, for all* $\lambda_2 \geq 0$*,* $\alpha \in \mathbb{Z} - \{0\}$.
9. *If* $Z \sim N(0, 1)$*, then for every even function* $h(\cdot)$*,* $(h(u) = h(-u))$*, we have* $h(Z) \stackrel{d}{=} h(X)$ *where* $\stackrel{d}{=}$ *means the equality in distribution.*
10. *If* $Y \sim SSGN(\lambda_1^*, \lambda_2^*, \alpha^*)$*, then for every even function* $h(\cdot)$*,* $(h(u) = h(-u))$*, we have* $h(Y) \stackrel{d}{=} h(X)$.

Proof The proof is straightforward. □

Proposition 2 *Let* $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$ *and* $F(x, \lambda_1, \lambda_2, \alpha)$ *be the cdf of X, then we have:*

$$F(x, \lambda_1, \lambda_2, \alpha) = \Phi(x) - 2H(x, \lambda_1, \lambda_2, \alpha), \tag{6}$$

where

$$H(x, \lambda_1, \lambda_2, \alpha) = \int_{-x}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu. \tag{7}$$

Proof We have

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt = \int_{-\infty}^{x} \int_{-\infty}^{\infty} \phi(t)\phi(u)dtdu$$

$$= \int_{-\infty}^{x} \left(\int_{-\infty}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)dt + \int_{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}}^{0} \phi(t)dt + \frac{1}{2} \right) \phi(u)du$$
$$= \frac{1}{2} F(x, \lambda_1, \lambda_2, \alpha) + \frac{1}{2}\Phi(x) + \int_{-\infty}^{x} \int_{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}}^{0} \phi(t)\phi(u)dtdu, \tag{8}$$

from which the following equality is obtained

$$\int_{-\infty}^{x} \int_{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}}^{0} \phi(t)\phi(u)dtdu = \int_{-x}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu = H(x, \lambda_1, \lambda_2, \alpha). \tag{9}$$

□

Properties of H function:

Proposition 3 *The H function in the cdf of SSGN distribution has two following properties:*

(1) $H(x, \lambda_1, \lambda_2, \alpha) = H(-x, \lambda_1, \lambda_2, \alpha)$, *for all* $x \in \mathbb{R}, \lambda_1 \in \mathbb{R}$, $\lambda_2 \geq 0$, $\alpha \in \mathbb{Z} - \{0\}$.
(2) $H(x, -\lambda_1, \lambda_2, \alpha) = -H(x, \lambda_1, \lambda_2, \alpha)$, *for all* $x \in \mathbb{R}, \lambda_1 \in \mathbb{R}$, $\lambda_2 \geq 0$, $\alpha \in \mathbb{Z} - \{0\}$.

Proof To prove (1), we start with definition of −H function

$$- H(x, \lambda_1, \lambda_2, \alpha) = \int_{-\infty}^{x} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu$$
$$= \int_{-\infty}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu - \int_{x}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu$$
$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu - \frac{1}{2} \right)$$
$$- \int_{x}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu$$
$$= 0 - \int_{x}^{\infty} \int_{0}^{\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^{2\alpha}}}} \phi(t)\phi(u)dtdu. \tag{10}$$

Equality (1) is obtained by multiplying both sides of the above equation by −1. The proof of part (2) is straightforward. □

Now, we obtain the moments of SSGN. Note that in view part (9) of Proposition 1, the even moments of SSGN and standard normal distribution are the same i.e.

$$E(X^{2K}) = 1 \times 3 \times 5 \times \cdots \times (2K - 1), \quad K = 1, 2, \ldots. \tag{11}$$

The odd moments of SSGN can be obtained using the following proposition.

Proposition 4 *Let* $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$. *Then for* $K = 0, 1, 2, \ldots$ *we have*

$$E(X^{2K+1}) = 2\left(b_K(\lambda_1, \lambda_2, \alpha) - b_K(0, \lambda_2, \alpha)\right), \tag{12}$$

when

$$b_K(\lambda_1, \lambda_2, \alpha) = \int_{0}^{\infty} \frac{u^k}{\sqrt{2\pi}} e^{-u/2} \Phi\left(\frac{\lambda_1 \sqrt{u}}{\sqrt{1 + \lambda_2 u^{\alpha}}} \right) dx, \tag{13}$$

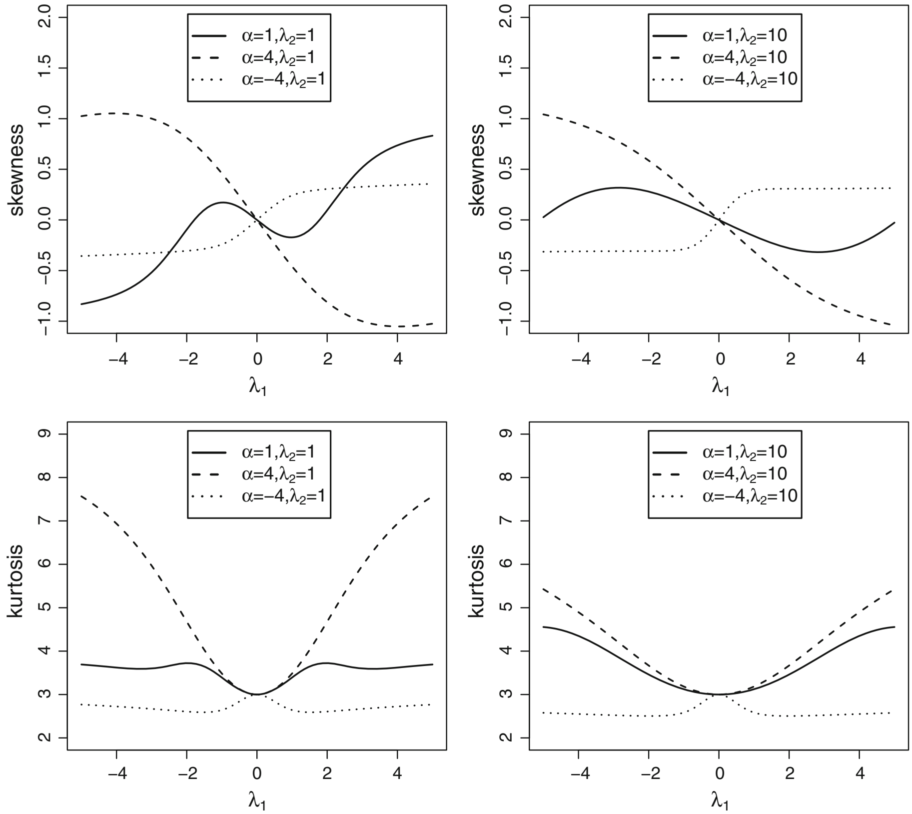

Fig. 3 Asymmetry and Kurtosis range variations of SSGN for some parameter values

and

$$b_K(0, \lambda_2, \alpha) = \frac{2^K \Gamma(K+1)}{\sqrt{2\pi}}. \tag{14}$$

Proof

$$\begin{aligned} E(X^{2K+1}) &= 2\int_0^\infty x^{2K+1}\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^{2\alpha}}}\right)dx - 2\int_0^\infty x^{2K+1}\phi(x)dx \\ &= 2b_K(\lambda_1, \lambda_2, \alpha) - \frac{2^{\mathrm{K}+1}\Gamma(K+1)}{\sqrt{2\pi}} \end{aligned} \tag{15}$$

□

Using the above formulas, we can obtain the skewness and Pearson's excess kurtosis coefficients of $SSGN(\lambda_1, \lambda_2, \alpha)$ for selective values of α. These coefficients are obtained via

$$\gamma_1 = \frac{E(X^3) - 3\mu\sigma^2 - \mu^3}{(\sigma^2)^{\frac{3}{2}}} \quad and \quad \gamma_2 = \frac{E(X^4) - 4E(X^3)\mu + 6E(X^2)\mu^2 - 3\mu^4}{(\sigma^2)^2}.$$

Figure 3, shows the variability of these coefficients for various values of the parameters.

Moment generating function is given by

$$M_X(t) = 2e^{t^2/2} E(\Phi\left(\frac{\lambda_1(Z+t)}{\sqrt{1+\lambda_2(Z+t)^{2\alpha}}}\right) \tag{16}$$

where Z has a standard normal distribution.

Certain relations of SSGN distribution with well-known distributions are mentioned in the following proposition:

Proposition 5 • *If* $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$ *and* $Y \sim \chi^2_{(k)}$ *then for* $G = \frac{X}{\sqrt{\frac{Y}{k}}}$, *we have*

$$\begin{aligned} f_G(g) &\to 2 f_{T_{(k)}}(g) I(g \geq 0) \quad as \quad \lambda_1 \to \infty \\ f_G(g) &\to 2 f_{T_{(k)}}(g) I(g < 0) \quad as \quad \lambda_1 \to -\infty \end{aligned} \tag{17}$$

where $f_{T_{(k)}}$ *is density of student t distribution with k degrees of freedom.*

- *If* $X_1, X_2 \overset{iid}{\sim} SSGN(\lambda_1, \lambda_2, \alpha)$ *and* $D = \frac{X_1}{|X_2|}$, *then*

$$\begin{aligned} f_D(d) &\to 2 f_U(d) I(d \geq 0) \quad as \quad \lambda_1 \to \infty \\ f_D(d) &\to 2 f_U(d) I(d < 0) \quad as \quad \lambda_1 \to -\infty \end{aligned} \tag{18}$$

where f_U *is density of standard Cauchy distribution* ($C(0, 1)$) *and iid stands for independent and identically distributed.*

- *If* $X_1, X_2, \ldots, X_n \overset{iid}{\sim} SSGN(\lambda_1, \lambda_2, \alpha)$ *then* $D = \sum_{i=1}^{n} X_i^2 \sim \chi^2_n$.
- *If* $X|Y = y \sim SSGN(\frac{\lambda_1}{y}, \frac{\lambda_2}{y^{2\alpha}}, \alpha)$, $Y \sim SN(\theta)$ *and* $V = \frac{X}{Y}$, *then*

$$f_V(v) = 2g(v)\Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right) \tag{19}$$

where $g(v)$ *is the standard Cauchy density* ($C(0, 1)$).

Note the connection between Normal and SSGN distribution: If $X \sim N(0, 1)$, the resulting random variable G has $T_{(k)}$ distribution. If $X_1, X_2 \sim N(0, 1)$ the resulting random variable D has $C(0, 1)$ distribution and finally if $X_1, X_2, \ldots, X_n \sim N(0, 1)$ the resulting random variable D has χ^2_n distribution. So, we only prove (19).

Proof Let $f_V(v)$ denote the pdf of V. Then

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} 2\phi(vy)\Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right) 2\phi(y)\Phi(\theta y)\,|y|\,dy \\ &= 2\Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right) \int_{-\infty}^{\infty} 2\frac{|y|}{\sqrt{2\pi}}\phi(y\sqrt{1+v^2})\Phi(\theta y)dy \\ &= \frac{2}{\sqrt{2\pi}\,(1+v^2)}\Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right) \int_{-\infty}^{\infty} 2\,|y|\,\phi(y)\Phi(\frac{\theta y}{\sqrt{1+d^2}})dy. \end{aligned} \tag{20}$$

Since $|W| \stackrel{d}{=} |Z|$, where $Z \sim N(0,1)$ and $W \sim SN\left(\frac{\theta}{\sqrt{1+d^2}}\right)$, we have

$$f_V(v) = \frac{2}{\sqrt{2\pi}(1+v^2)} \Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right) \int_{-\infty}^{\infty} |y| \phi(y) dy = 2g(v) \Phi\left(\frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^{2\alpha}}}\right), \tag{21}$$

which completes the proof. □

The family of skew-Cauchy distribution was introduced by considering the distribution of $\frac{X_\lambda}{X}$, where $X_\lambda \sim SN(\lambda)$ and $X \sim N(0,1)$ are independent random variables [7]. Another family of two parameters skew-Cauchy distribution which includes the skew-Cauchy distribution as a special case was proposed by [13]. Nekokhou et al. [19] introduced three parameters skew-Cauchy distribution based on relationships between SN and flexible skew generalized normal (FSGN) distributions. The pdf (19) presents another three parameters skew-Cauchy distribution based on SN and SSGN distributions without the assumption of independence of the random variables.

Method of generating data from SSGN distribution is presented by the stochastic representation. The first stochastic representation of SSGN distribution will be introduced in Proposition 6 which is constructed on random variables with the standard normal distributions.

Proposition 6 *Let Y and Z be iid random variables with $N(0,1)$ distribution, then*

$$X = \begin{cases} Y & if \quad Z \leq \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}} \\ -Y & if \quad Z > \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}} \end{cases} \tag{22}$$

has $SSGN(\lambda_1, \lambda_2, \alpha)$ distribution.

Proof Observe that

$$\begin{aligned} F_X(x) &= P\left(X \leq x, Z \leq \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) + P\left(X \leq x, Z > \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) \\ &= P\left(Y \leq x, Z \leq \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) + P\left(-Y \leq x, Z > \frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) \\ &= \int_{-\infty}^{x} \int_{-\infty}^{\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}} \phi(z)\phi(y) dz dy + \int_{-x}^{\infty} \int_{\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}}^{\infty} \phi(z)\phi(y) dz dy \\ &= \int_{-\infty}^{x} \phi(y) \Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy + \int_{-\infty}^{x} \phi(y) \Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy \\ &= 2\int_{-\infty}^{x} \phi(y) \Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy. \end{aligned} \tag{23}$$

Now, differentiating $F_X(x)$ with respect to x, we have:

$$f_X(x) = 2\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^{2\alpha}}}\right). \tag{24}$$

□

The second approach for generating data from SSGN distribution via standard normal and uniform distributions is presented in the next proposition.

Proposition 7 *Let Y and U be independent random variables with distributions* $N(0, 1)$ *and the uniform distribution on the interval [0,1]* $(U(0, 1))$*, respectively. The random variable*

$$X = \begin{cases} Y & if \quad U \leq \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) \\ -Y & if \quad U > \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right) \end{cases} \tag{25}$$

has $SSGN(\lambda_1, \lambda_2, \alpha)$ *distribution.*

Proof The proof is similar to that of Proposition 6. □

The next proposition establishes the stochastic representation of SSGN based on Normal, Uniform and Bernoulli distributions.

Proposition 8 *Let* Y, U *and* V *be independent random variables with* $N(0, 1)$, $U(0, 1)$ *and the Bernoulli* $(B(1, p))$ *distributions, respectively. Define* $X_1 = Y|U \leq \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right)$ *and* $X_2 = -Y|U > \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right)$*. Then*

$$X_1 \stackrel{d}{=} X_2 \sim SSGN(\lambda_1, \lambda_2, \alpha), \tag{26}$$

and

$$H = VX_1 + (1 - V)X_2 \tag{27}$$

has $SSGN(\lambda_1, \lambda_2, \alpha)$ *distribution.*

Proof Note that

$$P(X_1 \leq y) = P\left(Y \leq y|U \leq \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right)\right) = \frac{P\left(Y \leq y,\ U \leq \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right)\right)}{P\left(U \leq \Phi\left(\frac{\lambda_1 Y}{\sqrt{1+\lambda_2 Y^{2\alpha}}}\right)\right)}$$

$$= \frac{\int_{-\infty}^{y} \phi(y)\Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy}{\int_{-\infty}^{\infty} \phi(y)\Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy} = \int_{-\infty}^{y} 2\phi(y)\Phi\left(\frac{\lambda_1 y}{\sqrt{1+\lambda_2 y^{2\alpha}}}\right) dy. \tag{28}$$

Similarly X_2 has the same SSGN distribution. Now, we show that H has SSGN distribution:

$$\begin{aligned} F_H(h) &= P(H \leq h) = P(H \leq h|V = 1)P(V = 1) + P(H \leq h|V = 0)p(V = 0) \\ &= P(X_1 \leq h)p + P(X_2 \leq h)(1 - p) = P(X_1 \leq h). \end{aligned} \tag{29}$$

The last equality is based on the fact that X_1, X_2 and H are identically distributed. □

3 Maximum likelihood estimation

We consider the location-scale version of SSGN distribution for fitting a model to a real data set. For this purpose, we define $Y = \mu + \sigma X$ where $X \sim SSGN(\lambda_1, \lambda_2, \alpha)$ and $(\mu \in \mathbb{R}, \sigma > 0)$. Then

$$f_Y(y\,|\theta) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\frac{\lambda_1\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{1+\lambda_2\left(\frac{y-\mu}{\sigma}\right)^{2\alpha}}}\right) \tag{30}$$

where $\theta = (\mu, \sigma, \lambda_1, \lambda_2, \alpha)$ i.e. $Y \sim SSGN\,(\mu, \sigma, \lambda_1, \lambda_2, \alpha)$.

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a population with pdf (30). Then the likelihood function of the sample is

$$\ell(\mu, \sigma, \lambda_1, \lambda_2, \alpha) = n\ln\left(\frac{2}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n} z_i^2 + \sum_{i=1}^{n}\log\left(\Phi\left(\frac{\lambda_1 z_i}{\sqrt{1+\lambda_2 z_i^{2\alpha}}}\right)\right), \tag{31}$$

where $z_i = \frac{x_i-\mu}{\sigma}$.

Since space of α is discrete, ML estimation is performed by the following algorithm based on profile likelihood: suppose $\alpha \in \{A = -N, \ldots, -1, 1, \ldots, N\}$ for each $N \in \mathbb{Z}$.

For $i = 1, \ldots, 2N$

- Set $\alpha = A(i)$.
- With numerical calculation based on following score function equal to zero, we find MLE of $\mu, \sigma, \lambda_1, \lambda_2$ and show them by $(\hat{\mu}, \hat{\sigma}, \hat{\lambda}_1, \hat{\lambda}_2)$

$$\begin{aligned}
\frac{\partial\ell}{\partial\mu} &= -\sum_{i=1}^{n}\frac{z_i}{\sigma} - \sum_{i=1}^{n}\frac{\lambda_1(1+\lambda_2(z_i^{2\alpha}+\alpha z_i^{2\alpha-1}))}{\sigma(1+\lambda_2 z_i^{2\alpha})^{\frac{3}{2}}}w_i^*\\
\frac{\partial\ell}{\partial\sigma} &= -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{z_i^2}{\sigma} + \sum_{i=1}^{n}\frac{\lambda_1(2\alpha\lambda_2 z_i^{2\alpha+1} - z_i\sqrt{1+\lambda_2 z_i^{2\alpha}})}{\sigma(1+\lambda_2 z_i^{2\alpha})}w_i^*\\
\frac{\partial\ell}{\partial\lambda_1} &= \sum_{i=1}^{n}\frac{z_i}{\sqrt{1+\lambda_2 z_i^{2\alpha}}}w_i^*\\
\frac{\partial\ell}{\partial\lambda_2} &= \sum_{i=1}^{n}\frac{z_i^{2\alpha}}{(1+\lambda_2 z_i^{2\alpha})^{\frac{3}{2}}}w_i^*
\end{aligned} \tag{32}$$

where $w_i^* = \dfrac{\phi\left(\frac{\lambda_1 z_i}{\sqrt{1+\lambda_2 z_i^{2\alpha}}}\right)}{\Phi\left(\frac{\lambda_1 z_i}{\sqrt{1+\lambda_2 z_i^{2\alpha}}}\right)}$.

- now calculate $l(i) = \ell(\hat{\mu}, \hat{\sigma}, \hat{\lambda}_1, \hat{\lambda}_2, \alpha)$.

Return to the first step with i replaced by $i + 1$ and after final step, we find $\hat{\alpha}$ based on maximum of log likelihood l function and corresponding estimation of other parameters of $SSGN(\mu, \sigma, \lambda_1, \lambda_2, \alpha)$.

4 The multivariate shape skew generalized normal distribution

In this section, we introduce certain interesting results when $X = (X_1, \ldots, X_n)$ has a multivariate shape skew generalized normal density. This multivariate version can be used in graphical models since we show that its conditional distribution belongs to this family as well.

Definition 2 An n-dimensional random variable $X = (X_1, \ldots, X_n)$ has the multivariate shape skew generalized normal, $MSSGN(\lambda_1, \lambda_2, \alpha)$ distribution with the following conditions: if $\lambda_1 = 0$, then λ_2 and α must be zero and one, respectively, and if $\lambda_2 = 0$, then $\alpha = 1$ with the following density:

$$f(x_1, \ldots, x_n; \lambda_1, \lambda_2, \alpha) = c(\lambda_1, \lambda_2, \alpha). \prod_{i=1}^{n} \phi(x_i).\Phi\left(\frac{\lambda_1 \prod_{i=1}^{n} x_i}{\sqrt{1 + \lambda_2 \left(\prod_{i=1}^{n} x_i\right)^{2\alpha}}}\right), \quad (33)$$

where $\phi(x_1)$,…,$\phi(x_n)$ are standard normal densities and

$$\begin{aligned} c(\lambda_1, \lambda_2, \alpha) &= \frac{1}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n} \phi(x_i)\Phi\left(\frac{\lambda_1 \prod_{i=1}^{n} x_i}{\sqrt{1+\lambda_2(\prod_{i=1}^{n} x_i)^{2\alpha}}}\right) dx_1 \ldots dx_n} \\ &= \frac{1}{E\left(\Phi\left(\frac{\lambda_1(\prod_{i=1}^{n} U_i)}{\sqrt{1+\lambda_2(\prod_{i=1}^{n} U_i)^{2\alpha}}}\right)\right)}, \end{aligned} \quad (34)$$

where $U_1, \ldots, U_n \overset{iid}{\sim} N(0, 1)$, with the following property:
$c(-\lambda_1, \lambda_2, \alpha) = c(\lambda_1, \lambda_2, \alpha)$.

Proposition 10 presents the relation between MSSGN and SSGN distributions and Proposition 11 presents an stochastic representation of MSSGN distribution.

Proposition 9 *Let* $(X_1, \ldots, X_n) \sim MSSGN(\lambda_1, \lambda_2, \alpha)$. *The following properties hold:*

(1) *The conditional distribution of each random variable given the other random variables has a shape skew normal distribution, i.e.,*

$$\begin{aligned} &X_i \,|(X_1, \ldots, X_{i-1}, X_{i+1} \ldots, X_n) = (x_1, \ldots, x_{i-1}, x_{i+1} \ldots, x_n) \\ &\sim SSGN\left(\lambda_1. \prod_{\substack{j=1 \\ j\neq i}}^{n} x_i, \lambda_2. \prod_{\substack{j=1 \\ j\neq i}}^{n} x_i^{2\alpha}, \alpha\right), \end{aligned} \quad (35)$$

for $i = 1, \ldots, n$.

(2) *The conditional distribution of each random vector given the other random variables has a multivariate shape skew normal distribution, i.e.,*

$$\begin{aligned}(X_1,\ldots,X_r)\,|X_{r+1},\ldots,X_n &= (x_{r+1},\ldots,x_n)\\ &\sim MSSGN\left(\lambda_1.\prod_{j=r+1}^{n}x_i,\lambda_2.\prod_{j=r+1}^{n}x_i^{2\alpha},\alpha\right)\end{aligned} \tag{36}$$

for $1<r<n$.

Proof of part (1) Let $Y=X_i\,|(X_1,\ldots,X_{i-1},X_{i+1}\ldots,X_n)=(x_1,\ldots,x_{i-1},x_{i+1}\ldots,x_n)$. From the definition of conditional pdf, we have

$$\begin{aligned}f_Y(y) &= \frac{f_{X_1,\cdots X_n}(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)}{f_{X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n}(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)}\\ &= \frac{c(\lambda_1,\lambda_2,\alpha).\prod_{\substack{j=1\\ j\neq i}}^{n}\phi(x_j).\phi(y).\Phi\left(\frac{\left(\lambda_1\prod_{\substack{j=1\\ j\neq i}}^{n}x_j\right)y}{\sqrt{1+\left(\lambda_2\prod_{\substack{j=1\\ j\neq i}}^{n}x_j^{2\alpha}\right)y^{2\alpha}}}\right)}{\int_{-\infty}^{\infty}c(\lambda_1,\lambda_2,\alpha).\prod_{\substack{j=1\\ j\neq i}}^{n}\phi(x_j).\phi(y).\Phi\left(\frac{\left(\lambda_1\prod_{\substack{j=1\\ j\neq i}}^{n}x_j\right)y}{\sqrt{1+\left(\lambda_2\prod_{\substack{j=1\\ j\neq i}}^{n}x_j^{2\alpha}\right)y^{2\alpha}}}\right)dy}\\ &\propto d\phi(y)\Phi\left(\frac{\left(\lambda_1\prod_{\substack{j=1\\ j\neq i}}^{n}x_j\right)y}{\sqrt{1+\left(\lambda_2\prod_{\substack{j=1\\ j\neq i}}^{n}x_j^{2\alpha}\right)y^{2\alpha}}}\right)\end{aligned} \tag{37}$$

where d is the normalizing constant. Therefore, Y has $SSGN\left(\lambda_1.\prod_{\substack{j=1\\ j\neq i}}^{n}x_j,\lambda_2.\prod_{\substack{j=1\\ j\neq i}}^{n}x_j^{2\alpha},\alpha\right)$ distribution.

Proof of part (2) Let $\mathbf{Y}=X_1,\ldots,X_r\,|X_{r+1},\ldots,X_n=(x_{r+1},\ldots,x_n)$. From the definition of conditional pdf, we have

$$\begin{aligned}f_{\mathbf{Y}}(y_1,\ldots,y_r) &= \frac{f_{X_1,\cdots X_n}(y_1,\ldots,y_r,x_{r+1},\ldots,x_n)}{f_{X_{r+1},\ldots,X_n}(x_{r+1},\ldots,x_n)}\\ &= \frac{c(\lambda_1,\lambda_2,\alpha).\prod_{j=r+1}^{n}\phi(x_j).\prod_{i=1}^{r}\phi(y_i).\Phi\left(\frac{\left(\lambda_1\prod_{j=r+1}^{n}x_j\right)\prod_{i=1}^{r}y_i}{\sqrt{1+\left(\lambda_2\prod_{j=r+1}^{n}x_j^{2\alpha}\right)\prod_{i=1}^{r}y_i^{2\alpha}}}\right)}{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}c(\lambda_1,\lambda_2,\alpha).\prod_{j=r+1}^{n}\phi(x_i).\prod_{i=1}^{r}\phi(y_i).\Phi\left(\frac{\left(\lambda_1\prod_{j=r+1}^{n}x_j\right)\prod_{i=1}^{r}y_i}{\sqrt{1+\left(\lambda_2\prod_{j=r+1}^{n}x_j^{2\alpha}\right)\prod_{i=1}^{r}y_i^{2\alpha}}}\right)dy_1,\ldots,dy_r}\\ &= d\prod_{i=1}^{r}\phi(y_i).\Phi\left(\frac{\left(\lambda_1\prod_{j=r+1}^{n}x_j\right)\prod_{i=1}^{r}y_i}{\sqrt{1+\left(\lambda_2\prod_{j=r+1}^{n}x_j^{2\alpha}\right)\prod_{i=1}^{r}y_i^{2\alpha}}}\right)\end{aligned} \tag{38}$$

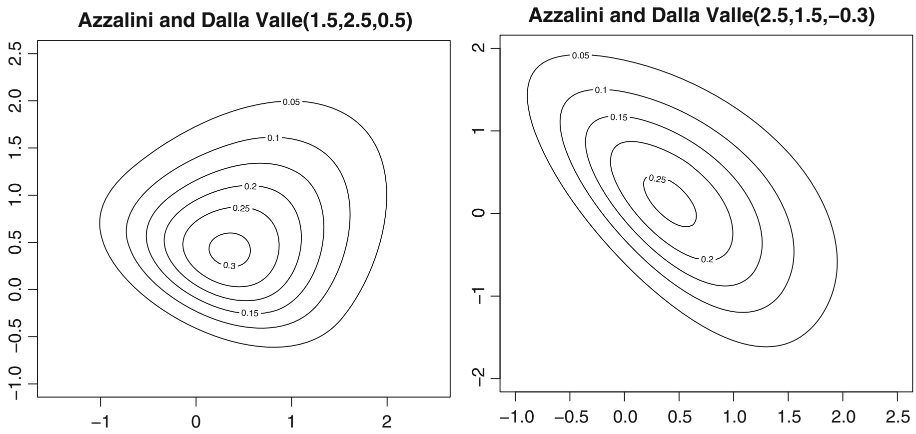

Fig. 4 Some possible contours of bivariate [6] for several values of λ_1, λ_2 and ρ

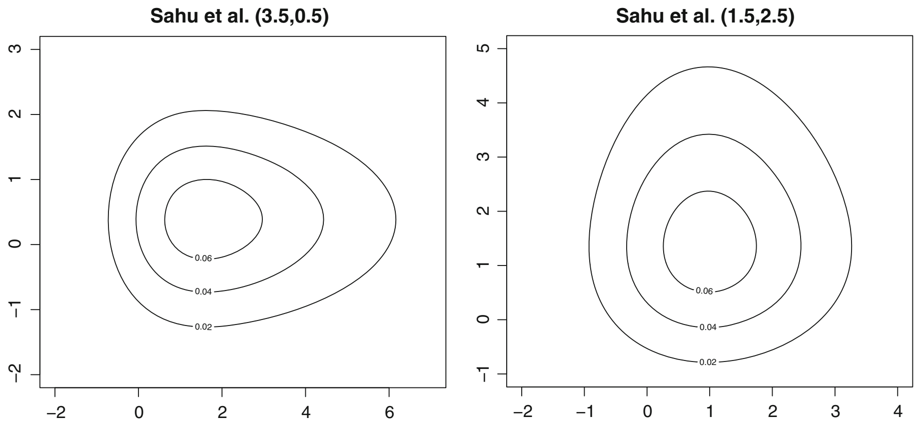

Fig. 5 Some possible contours of bivariate [21] for several values of δ_1 and δ_2

where d is the normalizing constant. Thus, $\mathbf{Y}$ has $MSSGN\left(\lambda_1 \prod_{j=r+1}^{n} x_i, \lambda_2 \prod_{j=r+1}^{n} x_i^{2\alpha}, \alpha\right)$ distribution. □

Proposition 10 *Let $X_1, \ldots, X_n$ and Z be i.i.d. random variables with $N(0, 1)$ distribution. Then,*

$$(X_1, \ldots, X_n) \left| \left\{ Z \leq \frac{\lambda_1 \prod_{i=1}^{n} X_i}{\sqrt{1 + \lambda_2 \left(\prod_{i=1}^{n} X_i^{2\alpha}\right)}} \right\} \right. \sim MSSGN(\lambda_1, \lambda_2, \alpha). \tag{39}$$

Proof Let $B = \left\{ Z \leq \frac{\lambda_1 \prod_{i=1}^{n} X_i}{\sqrt{1+\lambda_2 \left(\prod_{i=1}^{n} X_i\right)^{2\alpha}}} \right\}$. Then, we have

$$f_{(X_1,\ldots,X_n)|B}\,(x_1, \ldots, x_n\,|B) = \frac{p(B\,|(X_1, \ldots, X_n) = (x_1, \ldots, x_n)).f_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n)}{p\left(Z \leq \frac{\lambda_1 \prod_{i=1}^{n} X_i}{\sqrt{1+\lambda_2 \left(\prod_{i=1}^{n} X_i^{2\alpha}\right)}}\right)}$$

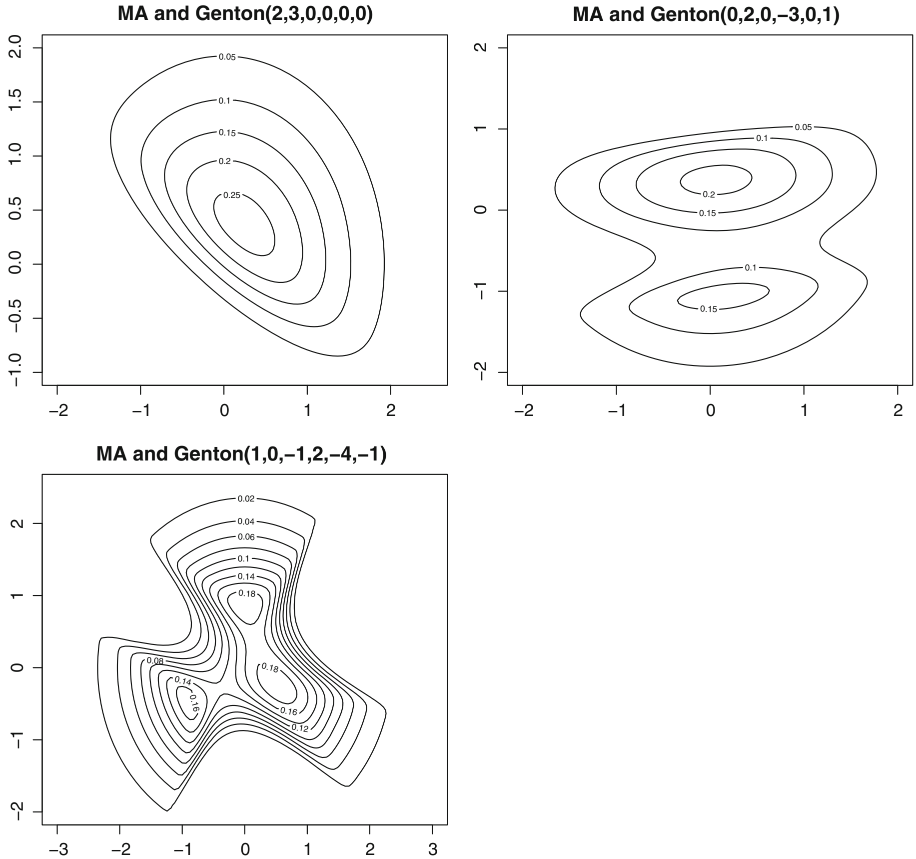

Fig. 6 Some possible contours of bivariate [18] for several values of α_1,α_2, β_1, β_2, β_3 and β_4

$$= \frac{p\left(Z \leq \frac{\lambda_1 \prod_{i=1}^n x_i}{\sqrt{1+\lambda_2(\prod_{i=1}^n x_i^{2\alpha})}}\right) \prod_{i=1}^n \phi(x_i)}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n \phi(x_i) \Phi\left(\frac{\lambda_1 \prod_{i=1}^n x_i}{\sqrt{1+\lambda_2(\prod_{i=1}^n x_i)^{2\alpha}}}\right) dx_1, \ldots, dx_n}$$

$$= c(\lambda_1, \lambda_2, \alpha). \prod_{i=1}^n \phi(x_i).\Phi\left(\frac{\lambda_1 \prod_{i=1}^n x_i}{\sqrt{1+\lambda_2\left(\prod_{i=1}^n x_i\right)^{2\alpha}}}\right). \tag{40}$$

Thus $(X_1, \ldots, X_n) \left| \left\{ Z \leq \frac{\lambda_1 \prod_{i=1}^n X_i}{\sqrt{1+\lambda_2(\prod_{i=1}^n X_i^{2\alpha})}} \right\} \right.$ has $MSSGN(\lambda_1, \lambda_2, \alpha)$ distribution. □

The MSSGN distribution reduces to multivariate normal distribution, $MN_n(\mathbf{0}, \mathbf{I})$ if $\lambda_1 = 0$. In Fig. 7, some possible contours of bivariate $MSSGN(\lambda_1, \lambda_2, \alpha)$ are shown for several values of α, λ_1 and λ_2. This figure shows that our proposed class is more flexible than classical multivariate skew normal such as [6] (Fig. 4), [21] (Fig. 5) and the density of MSSGN distribution has different shape than the pdf of [18] (Fig. 6) distribution.

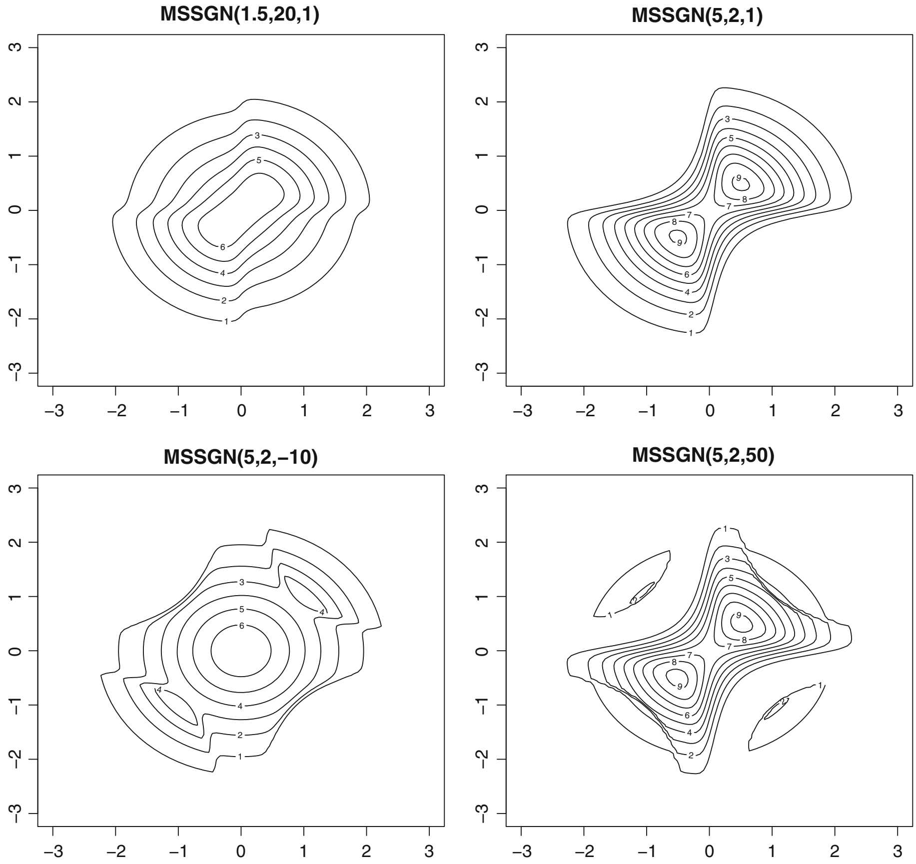

Fig. 7 Some possible contours of bivariate $MSSGN(\lambda_1, \lambda_2, \alpha)$ for several values of α, λ_1 and λ_2

Table 1 Descriptive statistics for the first and second examples

Example	n	$\bar{x}$	s	γ_1	γ_2
1	240	8.109	2.434	−0.209	2.534
2	49	8805.694	4553.915	0.093	2.085

5 Data analysis

We consider the variable E-Shiny (first example) available in the database creaminess of cream cheese which can be found at http://www.models.kvl.dk/Cream and the Kevlar data represent the failure times when the pressure is at 70 percent stress level that is presented by [4]. Table 1 shows the summary statistics (length, mean, standard deviation, skewness ($\gamma_1 = \frac{m_3}{s^3}$) and kurtosis ($\gamma_2 = \frac{m_4}{s^4}$))) for these two examples. (m_r is the rth central sample moments about mean). In Tables 2 and 5, two distributions are fitted to the data of the first and second examples, respectively. They are SGN [1] and $FSGN$ [19]. Also, we compare our proposed distribution with two component mixture normal $(\mu_1, \sigma_1, \mu_2, \sigma_2, p)$ distributions. In all

Table 2 Estimated parameters and log-likelihood for the models SGN, FSGN, FGSN, SSGN and Mixture Normal for the first example (the numbers in brackets are the estimates are standard errors)

Model	μ	σ	λ_1	λ_2	λ_3	α	Log-L	Information
SGN	7.840 (0.156)	2.443 (0.111)	0.187 (0.104)	0.313 (0.658)	–	–	−553.329	$\mu_X = 8.112$ $\sigma_X = 2.427$ $\gamma_{1X} = -0.041$ $\gamma_{2X} = 3.019$
FSGN	9.626 (0.135)	2.863 (0.108)	0.262 (0.307)	5.500 (1.977)	−2.478 (0.354)	–	−548.109	$\mu_X = 8.130$ $\sigma_X = 2.441$ $\gamma_{1X} = -0.271$ $\gamma_{2X} = 2.615$
SSGN	9.693 (0.094)	2.876 (0.092)	−1.230 (0.169)	0.010 (0.015)	–	−13 (1.932)	−547.571	$\mu_X = 8.196$ $\sigma_X = 2.456$ $\gamma_{1X} = -0.271$ $\gamma_{2X} = 2.614$
	μ_1	σ_1	μ_2	σ_2	p			
Mix-normal	7.167 (0.190)	2.101 (0.134)	10.569 (0.199)	1.182 (0.146)	0.723 (0.041)		−548.544	$\mu_X = 8.107$ $\sigma_X = 2.428$ $\gamma_{1X} = -0.189$ $\gamma_{2X} = 2.465$

Table 3 Formal goodness of fit statistics for first example

Model	Goodness of fit criteria	
	AIC	$CAIC$
SGN	1114.658	1114.827
FSGN	1106.218	1106.472
SSGN	1105.142	1105.396
Mix-Normal	1107.088	1107.342

Table 4 Formal goodness of fit statistics for second example

Model	Goodness of fit criteria	
	AIC	$CAIC$
SGN	971.592	972.113
FSGN	968.926	969.814
SSGN	965.484	966.372
Mix-normal	970.096	970.984

cases, the models are augmented by the inclusion of location (μ) and scale (σ) parameters. In the second example, $FSGN$ reduces to $FGSN$ [18], since $\hat{\lambda}_2 = 0$.

In all of the cases, the parameters are estimated by the method of maximum likelihood using the stats package, optim function, of R software. If data set has unimodal histogram, then the parameter α can have values -1 , 1 and if it has bimodal histogram, we must search for α in $\mathbb{Z} - \{-1, 0, 1\}$. In the following examples, we are faced with two bimodal data sets. Thus, in view of Sect. 3, we must define a loop on the parameter α in $\mathbb{Z} - \{-1, 0, 1\}$. For simplicity, however, we choose an N in $\mathbb{Z}$ and define a loop on $\{-N, \ldots, N\} - \{-1, 0, 1\}$. Then at each step of the loop, by optim function in R program, the MLE and the corresponding log-likelihood values are obtained. After completing all the steps in the loop, the MLEs of all the parameters are obtained by maximizing the log-likelihood function. The standard errors of all the parameters except α are calculated using observed Fisher Information Matrix based on Hessian Matrix. The Hessian Matrix is obtained via "$Hessian = T$" code in optim function and finally just for the parameter α of SSGN distribution, standard error of the MLE is calculated using parametric bootstrap with the same sample size.

Akaike information criterion (AIC) and Corrected Akaike information criterion (CAIC) [8] statistics are used for goodness of fit test criterion fitting mentioned distributions applied to two data sets. The lower value of these statistics show better fit considering the number parameters of models. Tables 3 and 4 present the values of these statistics for two real data sets.

To compare the SSGN distribution with the SGN model for both data sets, consider testing the null hypothesis of an SGN distribution against a SSGN distribution using the likelihood ratio statistics based on the ratio $\Lambda = L_{SGN}(\hat{\mu}, \hat{\sigma}, \hat{\lambda}_1, \hat{\lambda}_2)/L_{SSGN}(\hat{\mu}, \hat{\sigma}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\alpha})$. Substituting the estimated values, we obtain $-2log\Lambda$ for the first and the second example as 11.515 and 8.108, respectively. When compared with the 95 percent critical value of the $\chi^2_{(1)} = 3.841$, we conclude that the null hypotheses are clearly rejected and there is a strong indication that the SSGN distribution presents a much better fit than the SGN distribution to the data sets under consideration.

Table 5 Estimated parameters and log-likelihood for the models SGN, FSGN, SSGN and mixture normal for the second example (the numbers in brackets are the estimates are standard errors)

Model	μ	σ	λ_1	λ_2	λ_3	α	Log-L	Information
SGN	5381.566 (645.277)	5667.984 (537.025)	1.1419 (0.305)	0.001 (0.583)	–	–	–481.796	$\mu_X = 8782.818$
								$\sigma_X = 4534.041$
								$\gamma_{1_X} = 0.180$
								$\gamma_{2_X} = 3.090$
FSGN	5475.085 (440.895)	5627.097 (416.833)	−0.624 (0.424)	0.000 (1.312)	2.325 (0.663)	–	−479.463	$\mu_X = 8692.201$
								$\sigma_X = 4616.658$
								$\gamma_{1_X} = 0.324$
								$\gamma_{2_X} = 2.567$
SSGN	5878.861 (181.590)	5373.976 (191.144)	93.669 (121.180)	65.269 (168.243)	–	−33 (4.099)	−477.742	$\mu_X = 8703.135$
								$\sigma_X = 4572.036$
								$\gamma_{1_X} = 0.329$
								$\gamma_{2_X} = 2.475$
	μ_1	σ_1	μ_2	σ_2	p			
Mix-normal	10,856.025 (716.68)	3569.749 (517.26)	3868.825 (669.48)	2012.860 (474.15)	0.706 (0.08)		−480.048	$\mu_X = 8801.788$
								$\sigma_X = 4507.913$
								$\gamma_{1_X} = 0.094$
								$\gamma_{2_X} = 2.292$

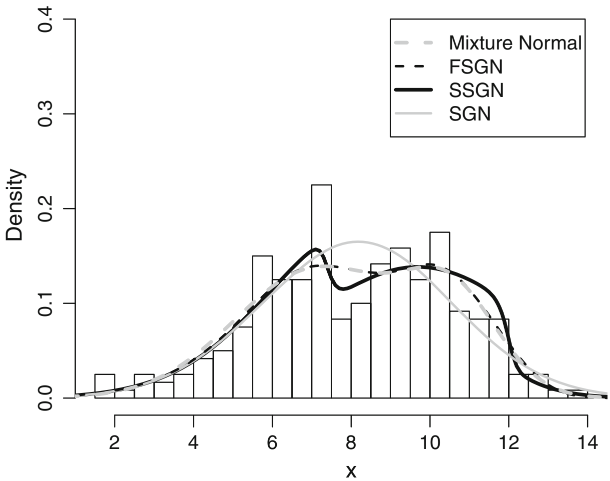

Fig. 8 Histogram for the E-Shiny variable. The *curves* represent densities fitted by maximum likelihood

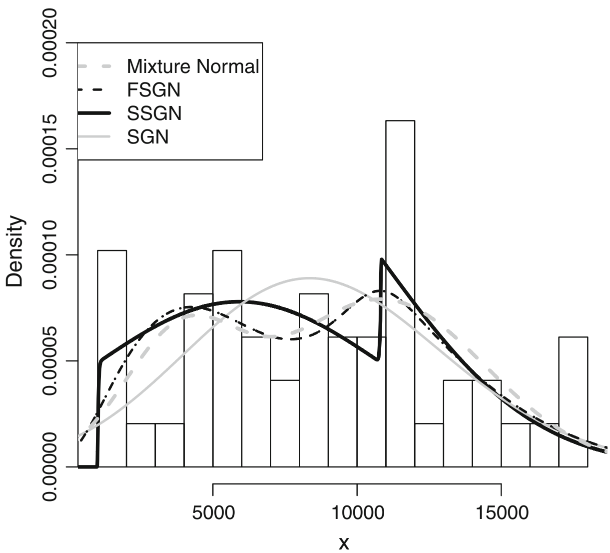

Fig. 9 Histogram for the failure times variable. The *curves* represent densities fitted by maximum likelihood

In both examples theoretical mean, standard deviation, skewness and kurtosis coefficients (γ_1 and γ_2) of distributions are presented in Tables 2 and 5. By considering scale of data sets, all theoretical and empirical statistics (Table 1) are approximately equal.

6 Conclusion

This paper introduces a flexible generalization of skew generalized normal distribution by adding a shape parameter to define shape skew generalized normal (SSGN) distribution, which also includes the Azzalini skew normal, Arellano-Valle et al. [1] skew generalized normal and special case of extended skew generalized normal distribution [10]. Inferential properties and three generation procedures are mentioned for our model. This model includes popular structure such as uni/bimodality, skewness, heavy tail and wider range for Pearson's excess kurtosis coefficient than SN and SGN distributions. Therefore, the proposed distribution is appropriate for other aspects of statistical analysis. The bivariate version of the distribution can model data sets with at most four modes, and its multivariate version can be used in graphical models such as directed acyclic graphs (DAG) (Figs. 8, 9).

Acknowledgements The authors are grateful to the Editor and the Reviewers for their comments and suggestion which greatly enhanced the content of this work.

References

1. Arellano-Valle, R.B., Gomez, H.W., Quintana, F.A.: A new class of skew-normal distributions. Commun Stat Theory Meth **33**(7), 1465–1480 (2005)
2. Arellano-Valle, R.B., Cortes, M.A., Gomez, H.W.: An extension of the epsilon-skew-normal distribution. Commun. Stat. Theory Meth. **39**(5), 912–922 (2010)
3. Arnold, B.C., Gomez, H.W., Salinas, H.S.: A doubly skewed normal distribution. Statistics **49**(4), 842–858 (2015)
4. Al-Aqtash, R., Lee, C., Famoye, F.: Gumbel–Weibull distribution: properties and applications. J. Modern Appl. Stat. Meth. **13**(2), 201–225 (2014)
5. Azzalini, A.: A class of distributions which include the normal. Scand. J. Stat. **12**(2), 171–178 (1985)
6. Azzalini, A., Dalla-Valle, A.: The multivariate skew-normal distribution. Biometrika **83**(4), 715–726 (1996)
7. Behboodian, J., Jamalizadeh, A., Balakrishnan, N.: A new class of skew-Cauchy distributions. Stat. Probab. Lett. **76**(14), 1488–1493 (2006)
8. Burnham, K.P., Anderson, D.R.: Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer, Berlin (2002)
9. Chakraborty, S., Hazarika, P.J., Masoom Ali, M.: A multimodal skewed extension of normal distribution: its properties and applications. Statistics **49**(4), 859–877 (2015)
10. Choudhury, K., Abdul Matin, M.: Extended skew generalized normal distribution. METRON Int. J. Stat. **69**(3), 265–278 (2011)
11. Gomez, H.W., Elal-Olivero, D., Salinas, H.S., Bolfarine, H.: Bimodal extension based on the skew-normal distribution with application to pollen data. Environmetrics **22**(1), 50–62 (2009)
12. Hasanalipour, P., Sharafi, M.: A new generalized Balakrishnan skew-normal distribution. Stat. Pap. **53**(1), 219–228 (2012)
13. Huang, W.J., Chen, Y.H.: Generalized skew-Cauchy distribution. Stat. Probab. Lett. **77**(11), 1137–1147 (2007)
14. Kim, H.J.: On a class of two-piece skew-normal distributions. Statistics **39**(6), 537–553 (2005)
15. Kumar, C.S., Anusree, M.R.: On an extended version of skew generalized normal distribution and some of its properties. Commun. Stat. Theory Meth. **44**(3), 573–586 (2015)
16. Kumar, C.S., Anusree, M.R.: A generalized two-piece skew normal distribution and some of its properties. Statistics **47**(6), 1370–1380 (2013)
17. Ley, C.: Skew Distributions. Encyclopedia of Environmetrics. Wiley, New York (2013)
18. Ma, Y., Genton, M.G.: Flexible class of skew-symmetric distributions. Scan. J. Stat. **31**(3), 459–468 (2004)
19. Nekoukhou, V., Alamatsaz, M.H., Aghajani, A.H.: A flexible skew-generalized normal distribution. Commun. Stat. Theory Meth. **42**(13), 2324–2334 (2013)
20. Rasekhi, M., Chinipardaz, R., Alavi, S.M.R.: A flexible generalization of the skew normal distribution based on a weighted normal distribution. Stat. Meth. Appl. **25**(3), 375–394 (2016)

21. Sahu, S.K., Branco, M.D., Dey, D.K.: A new class of multivariate skew distributions with applications to Bayesian regression models. Can. J. Stat. **31**(2), 129–150 (2003)
22. Venegas, O., Sanhueza, A.I., Gomez, H.W.: An extension of the skew-generalized normal distribution and its derivation. Proyecciones J. Math. **30**(3), 401–413 (2011)