Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

A Flexible Skew-Generalized Normal Distribution

V. Nekoukhou [a] , M. H. Alamatsaz [a] & A. H. Aghajani [a]

[a] Department of Statistics , University of Isfahan , Isfahan , Iran

To cite this article: V. Nekoukhou , M. H. Alamatsaz & A. H. Aghajani (2013): A Flexible Skew-Generalized Normal Distribution, Communications in Statistics - Theory and Methods, 42:13, 2324-2334

To link to this article: http://dx.doi.org/10.1080/03610926.2011.599003

Communications in Statistics—Theory and Methods, 42: 2324–2334, 2013

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610926.2011.599003

Taylor & Francis
Taylor & Francis Group

A Flexible Skew-Generalized Normal Distribution

V. NEKOUKHOU, M. H. ALAMATSAZ,
AND A. H. AGHAJANI

Department of Statistics, University of Isfahan, Isfahan, Iran

Skew-symmetric distributions of various types have been the center of attraction by many researchers in the literature. In this article, we will introduce a uni/bimodal generalization of the Azzalini's skew-normal distribution which is indeed an extension of the skew-generalized normal distribution obtained by Arellano-Valle et al. (2004). Our new distribution contains more parameters and thus it is more flexible in data modeling. Indeed, certain univariate case of the so called flexible skew-symmetric distribution of Ma and Genton (2004) is also a particular case of our proposed model. We will first study some basic distributional properties of the new extension, such as its distribution function, limiting behavior and moments. Then, we will investigate some useful results regarding its relation with other known distributions, such as student's t and skew-Cauchy distributions. In addition, we will present certain methods to generate the new distribution and, finally, we shall apply the model to a real data set to illustrate its behavior comparing to some rival models.

Keywords Fitting; Flexible skew-normal distribution; Skew-normal distribution; Stochastic representation.

Mathematics Subject Classification 60E; 62E.

1. Introduction

Recently, much of interest has been shown on a family of distributions called *skew-symmetric* having probability density function (pdf)

$$2f(x)\kappa(x), \tag{1}$$

where f is a symmetric pdf (about zero) and κ is a lebesgue measurable function such that:

(a) $0 \leq \kappa(x) \leq 1$, $x \in \mathfrak{R}$,
(b) $\kappa(x) + \kappa(-x) = 1$, *a.e. on* $\mathfrak{R}$.

Received January 16, 2011; Accepted June 17, 2011
Address correspondence to M. H. Alamatsaz, Department of Statistics, University of Isfahan, Isfahan, Iran; E-mail: alamatho@sci.ui.ac.ir

Particularly, κ can be an absolutely continuous symmetric distribution function. As a special case, κ can be the cumulative distribution function (cdf) of f; see, e.g., Arnold and Lin (2004).

Nadarajah and Kotz (2003) introduced the family of skew-symmetric-normal distributions as

$$2\phi(x)F(\lambda x), \tag{2}$$

where ϕ is the pdf of the standard normal distribution and F is an absolutely continuous cdf with a symmetric density and λ is a real constant. They produced various skew-symmetric distributions by choosing F as normal, student's t, Laplace, logistic, and uniform. Recently, Nekoukhou and Alamatsaz (2012) introduced the family of skew-symmetric-Laplace distributions. They obtained certain properties of this family of distributions and discussed some stochastic ordering properties of the family of skew-symmetric distributions. In fact, the starting point of all these studies was the skew-normal distribution introduced by Azzalini (1985),

$$2\phi(x)\Phi(\lambda x) \tag{3}$$

in which ϕ and Φ are the standard normal density and distribution functions, respectively. A random variable Z_λ with the above density is denoted by $Z_\lambda \sim SN(\lambda)$ in the literature.

Arellano-Valle et al. (2004) introduced a generalization of $SN(\lambda)$, called skew-generalized normal distribution, having pdf of the form

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

where $\lambda_1 \in \Re$ and $\lambda_2 \geq 0$ are real constants. The last authors used the notation $SGN(\lambda_1, \lambda_2)$ to represent their distribution. They also investigated some other properties of their skew distribution. As it appears, $SGN(\lambda_1, \lambda_2)$ distributions are unimodal. In addition, Ma and Genton (2004) introduced a flexible class of skew-symmetric distributions. In fact, they introduced the multivariate family

$$2\phi_p(\mathbf{x})\Phi(P_K(\mathbf{x})), \tag{5}$$

where ϕ_p denotes a p-dimensional multivariate standard normal pdf and P_K is an odd polynomial of order K. They revealed some properties of this family of distributions.

In this article, we will introduce a flexible skew-generalized normal distribution which involves three parameters and, depending on its parameter values, may be unimodal or bimodal.

This article is organized as follows: Sec. 2 introduces the flexible skew-generalized normal distribution and discusses some of its important features and properties such as its limiting behavior, distribution function, and moments. We will also reveal some interesting relations between the new distribution and other known distributions, such as the student's t and skew-Cauchy distributions. In Sec. 3, we will show that how one can generate a flexible skew-generalized normal distribution in different manners. Finally, Sec. 4 examines the proposed model with a real data set and illustrates its better fit comparing to some competitive models.

2. Definition and Basic Properties

In this section, we will introduce a flexible skew-generalized normal density function which is an extension of the skew-generalized normal distribution obtained by Arellano-Valle et al. (2004) and also a certain generalization of the flexible skew-normal distribution discussed in Ma and Genton (2004) in the univariate case. Indeed, the purpose of the present section is to introduce a uni/bimodal extension of the skew-normal density function whose mode(s) can be controlled by a suitable choice of values of its parameters. Such an extension is potentially relevant for practical applications, because there are far fewer distributions available for dealing with bimodal data than in the unimodal case.

Definition 2.1. A random variable X is said to be flexible skew-generalized normal distributed, denoted by $FSGN(\lambda_1, \lambda_2, \lambda_3)$, if its pdf has the form

$$f(x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x)\Phi\left(\frac{\lambda_1 x + \lambda_3 x^3}{\sqrt{1+\lambda_2 x^2}}\right), \quad x \in \mathfrak{R}, \tag{6}$$

where $\lambda_1, \lambda_3 \in \mathfrak{R}$ and $\lambda_2 \geq 0$ are constants.

It is obvious that (6) is indeed a probability density. This is seen by the fact that

$$2\Phi\left(\frac{\lambda_1 x + \lambda_3 x^3}{\sqrt{1+\lambda_2 x^2}}\right) - 1$$

is an odd function in $x \in \mathfrak{R}$ and therefore

$$\int_{-\infty}^{\infty} 2\phi(x)\Phi\left(\frac{\lambda_1 x + \lambda_3 x^3}{\sqrt{1+\lambda_2 x^2}}\right) dx - 1 = E\left[2\Phi\left(\frac{\lambda_1 X + \lambda_3 X^3}{\sqrt{1+\lambda_2 X^2}}\right) - 1\right] = 0.$$

We should note that the univariate flexible skew-normal distribution, $FSN(\alpha, \beta)$, of Ma and Genton (2004), for $K = 3$ in (5), reduces to

$$g(x) = 2\phi(x)\Phi(\alpha x + \beta x^3), \tag{7}$$

which is clearly a particular case of (6), i.e., it is a $FSGN(\alpha, 0, \beta)$ distribution. These authors revealed that a general statement on the relation between the number of modes of their distribution and the order of the polynomial involved seems unavailable in the general case (5). However, they verified that the pdf g in (7) has at most two modes.

To see the modality behavior of a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distribution, we used some graphical methods and observed that the derivative of the density (6) changes sign at most once from positive to negative when $\lambda_1\lambda_3 > 0$ and changes sign two more times when $\lambda_1\lambda_3 < 0$. Therefore, the distribution in question is either unimodal or bimodal. This is also illustrated in Fig. 1.

The following basic properties of a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distribution can be obtained directly from (6).

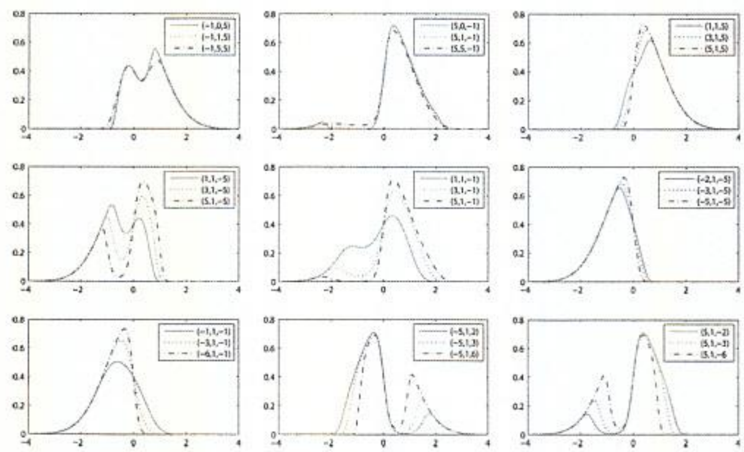

Figure 1. Illustrations of the pdf of a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distribution for several values of $(\lambda_1, \lambda_2, \lambda_3)$. (color figure available online.)

Proposition 2.1. *If $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$, then we have*:

(a) $FSGN(0, \lambda, 0) = N(0, 1)$, *for all* $\lambda > 0$;
(b) $FSGN(\lambda, 0, 0) = SN(\lambda)$, *for all* $\lambda \in \mathcal{R}$;
(c) $FSGN(\lambda_1, \lambda_2, 0) = SGN(\lambda_1, \lambda_2)$, *for all* $\lambda_1 \in \mathcal{R}$ *and* $\lambda_2 \geq 0$;
(d) $FSGN(\lambda_1, 0, \lambda_3) = FSN(\lambda_1, \lambda_3)$, *for all* $\lambda_1, \lambda_3 \in \mathcal{R}$;
(e) $-X \sim FSGN(-\lambda_1, \lambda_2, -\lambda_3)$;
(f) $f(x; \lambda_1, \lambda_2, \lambda_3) + f(-x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x)$; *for all* $x \in \mathcal{R}$, $\lambda_1, \lambda_3 \in \mathcal{R}$ *and* $\lambda_2 \geq 0$;
(g) $\lim_{\lambda_1 \to \infty} f(x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x) I_{\{x \geq 0\}}$, *for all* $\lambda_2, \lambda_3 \geq 0$;
(h) $\lim_{\lambda_3 \to \infty} f(x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x) I_{\{x \geq 0\}}$, *for all* $\lambda_1, \lambda_2 \geq 0$;
(i) $\lim_{\lambda_1 \to -\infty} f(x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x) I_{\{x \leq 0\}}$, *for all* $\lambda_2 \geq 0$ *and* $\lambda_3 \leq 0$;
(j) $\lim_{\lambda_3 \to -\infty} f(x; \lambda_1, \lambda_2, \lambda_3) = 2\phi(x) I_{\{x \leq 0\}}$, *for all* $\lambda_1 \leq 0$ *and* $\lambda_2 \geq 0$;
(k) *If* $Z \sim N(0, 1)$, *then for every even function* $h(\cdot)$ *we have* $h(Z) \stackrel{d}{=} h(X)$ *in which* $\stackrel{d}{=}$ *means the equality in distribution*;
(l) *If* $Y \sim FSGN(\lambda_1^*, \lambda_2^*, \lambda_3^*)$, *then for every even function* $h(\cdot)$ *we have* $h(X) \stackrel{d}{=} h(Y)$.

Proposition 2.2. *Let* $Z \sim N(0, 1)$, $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$ *and* $F(x; \lambda_1, \lambda_2, \lambda_3)$ *be the cdf of X. Then, we have*

$$F(x; \lambda_1, \lambda_2, \lambda_3) = \Phi(x) - 2T(x; \lambda_1, \lambda_2, \lambda_3), \quad x \in \mathcal{R}, \tag{8}$$

where

$$T(x; \lambda_1, \lambda_2, \lambda_3) = \int_{-x}^{\infty} \int_0^{\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}} \phi(u)\phi(t)\,dt\,du.$$

Proof. It is easy to see that

$$\begin{aligned}\Phi(x) &= \int_{-\infty}^{x}\int_{-\infty}^{\infty}\phi(t)\phi(u)dt\,du \\ &= \int_{-\infty}^{x}\left[\int_{-\infty}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)dt + \int_{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}^{0}\phi(t)dt + \frac{1}{2}\right]\phi(u)du \\ &= \frac{1}{2}F(x;\lambda_1,\lambda_2,\lambda_3) + \int_{-\infty}^{x}\int_{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}^{0}\phi(t)\phi(u)dt\,du + \frac{1}{2}\Phi(x). \qquad (9)\end{aligned}$$

Thus, the result follows since

$$\int_{-\infty}^{x}\int_{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}^{0}\phi(t)\phi(u)dt\,du = \int_{-x}^{\infty}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du.$$

Proposition 2.3. *The function $T(x;\lambda_1,\lambda_2,\lambda_3)$ has the following properties*:

(a) $T(x;\lambda_1,\lambda_2,\lambda_3) = T(-x;\lambda_1,\lambda_2,\lambda_3)$ *for all* $x\in\Re$, $\lambda_1,\lambda_3\in\Re$ *and* $\lambda_2\geq 0$;
(b) $T(x;-\lambda_1,\lambda_2,-\lambda_3) = -T(x;\lambda_1,\lambda_2,\lambda_3)$ *for all* $x\in\Re$, $\lambda_1,\lambda_3\in\Re$ *and* $\lambda_2\geq 0$.

Proof. (a) We have:

$$\begin{aligned}-T(x;\lambda_1,\lambda_2,\lambda_3) &= \int_{-\infty}^{x}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du \\ &= \int_{-\infty}^{\infty}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du - \int_{x}^{\infty}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)dt - \frac{1}{2}\right)\phi(u)du \\ &\quad - \int_{x}^{\infty}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du \\ &= 0 - \int_{x}^{\infty}\int_{0}^{\frac{\lambda_1 u+\lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}}\phi(t)\phi(u)dt\,du,\end{aligned}$$

as required. The proof of part (b) is straightforward.

Now, we discuss the moments of a $FSGN(\lambda_1,\lambda_2,\lambda_3)$ distribution. Using part (k) in Proposition 2.1, we conclude that the even moments of $FSGN(\lambda_1,\lambda_2,\lambda_3)$ are the same as those of $N(0,1)$ distribution. Hence, we only need to obtain an expression for its odd moments. Let

$$c_k(\lambda_1,\lambda_2,\lambda_3) = \int_0^{\infty}\frac{u^k}{\sqrt{2\pi}}e^{-u/2}\Phi\left(\frac{\lambda_1\sqrt{u}+\lambda_3 u^{3/2}}{\sqrt{1+\lambda_2 u}}\right)du.$$

Then, if $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$, it is easy to see that for $k = 0, 1, 2, \ldots$ we have

$$E(X^{2k+1}) = 2c_k(\lambda_1, \lambda_2, \lambda_3) - \frac{2^{k+1}k!}{\sqrt{2\pi}}. \tag{10}$$

The moment generating function (mgf) of $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$ can also be easily shown to be of the form

$$M_X(t) = 2e^{t^2/2}E\left[\Phi\left(\frac{\lambda_1(Z+t) + \lambda_3(Z+t)^3}{\sqrt{1+\lambda_2(Z+t)^2}}\right)\right], \tag{11}$$

where $Z \sim N(0, 1)$.

In the following, we reveal certain relations between $FSGN(\lambda_1, \lambda_2, \lambda_3)$ and other known distributions. The proof of these results are straightforward and thus omitted.

Proposition 2.4. *Let $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$ be independent of $Y \sim \chi^2_{(r)}$, i.e., a chi-square random variable with r degrees of freedom, and define $W = \frac{X}{\sqrt{\frac{Y}{r}}}$. Then, we have*

$$f_W(w) \to 2f_{T_{(r)}}(w)I_{\{w\geq 0\}} \quad as\ \lambda_1, \lambda_3 \to \infty,$$
$$f_W(w) \to 2f_{T_{(r)}}(w)I_{\{w< 0\}} \quad as\ \lambda_1, \lambda_3 \to -\infty,$$

where $T_{(r)}$ is a student's t random variable with r degrees of freedom.

Proposition 2.5. *If X_1 and X_2 are independently and identically $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distributed random variables and $U = \frac{X_1}{|X_2|}$, then*

$$f_U(u) \to 2g(u)I_{\{u\geq 0\}} \quad as\ \lambda_1, \lambda_3 \to \infty,$$
$$f_U(u) \to 2g(u)I_{\{u< 0\}} \quad as\ \lambda_1, \lambda_3 \to -\infty,$$

where $g(u)$ is a standard Cauchy density.

It is known that if Y and Z are independent and identical $N(0, 1)$ random variables, then $\frac{Y}{Z}$ and $\frac{Y}{|Z|}$ are distributed as standard Cauchy. Many statisticians have been interested in generalizing the Cauchy distribution to a skewed one. Behboodian et al. (2006) introduced a family of skew-Cauchy distributions in the form $\frac{Z_\lambda}{|Z|}$, where $Z_\lambda \sim SN(\lambda)$ and $Z \sim N(0, 1)$ are independent random variables. Huang and Chen (2007) introduced another family of two parameters skew-Cauchy distributions which includes the skew-Cauchy density obtained by Behboodian et al. (2006) as a special case. The independent property of the random variables involved is the key assumption in constructing the mentioned skew-Cauchy distributions. Our next proposition, however, reveals a useful result which generates a skew-Cauchy distribution wherein the independence assumption is relaxed.

Proposition 2.6. *If $X \mid Y = y \sim FSGN(\frac{\lambda_1}{y}, \frac{\lambda_2}{y^2}, \frac{\lambda_3}{y^3})$ and $Y \sim SN(\delta)$, then $U = \frac{X}{Y}$ has the following pdf*

$$f_U(u) = 2g(u)\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right),$$

where $g(u)$ is a standard Cauchy density.

Proof. Let $f_U(u)$ denote the pdf of U. Hence, we have

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} 2\phi(uy)\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right) 2\phi(y)\Phi(\delta y)|y|dy \\ &= 2\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right)\int_{-\infty}^{\infty} 2\frac{|y|}{\sqrt{2\pi}}\phi(y\sqrt{1+u^2})\Phi(\delta y)dy \\ &= \frac{2}{\sqrt{2\pi}(1+u^2)}\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right)\int_{-\infty}^{\infty} 2|y|\phi(y)\Phi\left(\frac{\delta y}{\sqrt{1+u^2}}\right)dy. \end{aligned}$$

Since $|W| \stackrel{d}{=} |Z|$, where $Z \sim N(0,1)$ and $W \sim SN(\frac{\delta}{\sqrt{1+u^2}})$ we can write

$$\begin{aligned} f_U(u) &= \frac{2}{(1+u^2)\sqrt{2\pi}}\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right)\int_{-\infty}^{\infty} |y|\phi(y)dy \\ &= 2g(u)\Phi\left(\frac{\lambda_1 u + \lambda_3 u^3}{\sqrt{1+\lambda_2 u^2}}\right), \end{aligned}$$

as required.

3. Stochastic Representation

In this section, we will present some methods for generating a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ random variable.

Proposition 3.1. *Let Y and Z be independent $N(0,1)$ random variables. Then, we have*

$$X = Y \mid Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}} \sim FSGN(\lambda_1, \lambda_2, \lambda_3).$$

Proof. Let $F_X(\cdot)$ and $f_X(\cdot)$ denote the cdf and pdf of X, respectively. Hence, we have

$$\begin{aligned} F_X(y) &= P\left(Y \leq y \mid Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right) \\ &= \frac{P\left(Y \leq y, Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right)}{P\left(Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right)} \\ &= \frac{\int_{-\infty}^{y}\int_{-\infty}^{\frac{\lambda_1 t + \lambda_3 t^3}{\sqrt{1+\lambda_2 t^2}}} \phi(z)\phi(t)dzdt}{\int_{-\infty}^{\infty} P\left(Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}} \mid Y = y\right)\phi(y)dy} \end{aligned}$$

$$= \frac{\int_{-\infty}^{y} \phi(t)\Phi\left(\frac{\lambda_1 t+\lambda_3 t^3}{\sqrt{1+\lambda_2 t^2}}\right) dt}{\int_{-\infty}^{\infty} P\left(Z \leq \frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}\right)\phi(y)dy}$$

$$= 2\int_{-\infty}^{y} \phi(t)\Phi\left(\frac{\lambda_1 t+\lambda_3 t^3}{\sqrt{1+\lambda_2 t^2}}\right) dt.$$

Therefore,

$$f_X(y) = \frac{dF_X(y)}{dy} = 2\phi(y)\Phi\left(\frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}\right),$$

as required.

Remark 3.1. Under conditions of Proposition 3.1, one can easily show that $-Y \mid Z > \frac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}$ also has a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distribution.

Proposition 3.2. *Let Y and Z be independent $N(0, 1)$ random variables. Then,*

$$X = \begin{cases} Y & \text{if } Z \leq \dfrac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}} \\ -Y & \text{if } Z > \dfrac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}, \end{cases}$$

is distributed as $FSGN(\lambda_1, \lambda_2, \lambda_3)$.

Proof. We have

$$\begin{aligned} F_X(x) &= P\left(X \leq x, Z \leq \frac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right) + P\left(X \leq x, Z > \frac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right) \\ &= P\left(Y \leq x, Z \leq \frac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right) + P\left(-Y \leq x, Z > \frac{\lambda_1 Y+\lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}\right) \\ &= \int_{-\infty}^{x}\int_{-\infty}^{\frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}} \phi(z)\phi(y)dzdy + \int_{-x}^{\infty}\int_{\frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}}^{\infty} \phi(z)\phi(y)dzdy \\ &= \int_{-\infty}^{x} \phi(y)\Phi\left(\frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}\right) dy + \int_{-\infty}^{x} \phi(y)\Phi\left(\frac{\lambda_1 y+\lambda_3 y^3}{\sqrt{1+\lambda_2 y^2}}\right) dy. \end{aligned}$$

Therefore,

$$f_X(x) = \frac{dF_X(x)}{dx} = 2\phi(x)\Phi\left(\frac{\lambda_1 x+\lambda_3 x^3}{\sqrt{1+\lambda_2 x^2}}\right),$$

as required.

Proposition 3.3. *Let* $Z \sim N(0, 1)$ *be independent of* $U \sim U(0, 1)$, *i.e. a uniform random variable. Then, we have*

$$Z \mid U \leq \Phi\left(\frac{\lambda_1 Z + \lambda_3 Z^3}{\sqrt{1+\lambda_2 Z^2}}\right) \sim FSGN(\lambda_1, \lambda_2, \lambda_3).$$

Proof. The proof is similar to that of Proposition 3.1 and, thus, it is omitted.

Proposition 3.4. *Let* $Z \sim N(0, 1)$ *be independent of* $U \sim U(0, 1)$. *Then, the random variable*

$$X = \begin{cases} Z & \text{if } U \leq \Phi\left(\frac{\lambda_1 Z + \lambda_3 Z^3}{\sqrt{1+\lambda_2 Z^2}}\right) \\ -Z & \text{if } U > \Phi\left(\frac{\lambda_1 Z + \lambda_3 Z^3}{\sqrt{1+\lambda_2 Z^2}}\right), \end{cases}$$

is distributed as $FSGN(\lambda_1, \lambda_2, \lambda_3)$.

Proof. The proof is similar to that of Proposition 3.2 and, thus, it is omitted.

Proposition 3.5. *Let independent* $N(0, 1)$ *random variables* Y *and* Z *be also independent of the random variable* $V \sim Ber(p)$, *i.e. a Bernoulli random variable. If we define random variables* $X_1 = Y \mid Z \leq \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}$ *and* $X_2 = -Y \mid Z > \frac{\lambda_1 Y + \lambda_3 Y^3}{\sqrt{1+\lambda_2 Y^2}}$, *then we have*

$$X = VX_1 + (1 - V)X_2 \sim FSGN(\lambda_1, \lambda_2, \lambda_3).$$

Proof. Let $F_X(\cdot)$ and $f_X(\cdot)$ denote the cdf and pdf of the random variable X, respectively. Then, we have

$$\begin{aligned} F_X(x) &= P(X \leq x \mid V = 1)P(V = 1) + P(X \leq x \mid V = 0)P(V = 0) \\ &= P(X_1 \leq x)p + P(X_2 \leq x)(1 - p) \\ &= F_{X_1}(x)p + F_{X_2}(x)(1 - p). \end{aligned}$$

Hence,

$$f_X(x) = \frac{dF_X(x)}{dx} = f_{X_1}(x)p + f_{X_2}(x)(1 - p).$$

Now, since by Proposition 3.1 and Remark 3.1 we have $X_1 \stackrel{d}{=} X_2 \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$, we can conclude that $X \sim FSGN(\lambda_1, \lambda_2, \lambda_3)$.

4. Application

The location-scale *FSGN* distribution is defined as the distribution of $Y = \mu + \sigma X$, where X has a $FSGN(\lambda_1, \lambda_2, \lambda_3)$ distribution. Here, $\mu \in \Re$ and $\sigma > 0$ are the location

Table 1
MLEs of parameters on different models for the waiting time of 299 consecutive eruptions

Distribution	N	SN	SGN	FSN	$FSGN$
$\hat{\mu}$	72.3144	50.3299	48.1640	49.7777	59.2332
$\hat{\sigma}$	13.8671	25.9925	27.8484	26.4612	19.0633
$\hat{\lambda}_1$	–	5.4679	14.9644	5.9709	−72.6709
$\hat{\lambda}_2$	–	–	19.4177	–	9663.1772
$\hat{\lambda}_3$	–	–	–	10.3431	297.2672
AIC	2424.9767	2470.1886	2470.3007	2471.9035	2385.5577
BIC	2432.3775	2481.2899	2485.1024	2486.7053	2404.0599

and scale parameters, respectively. Hence, the pdf of Y is given by

$$f(y;\theta) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\frac{\lambda_1(y-\mu)+\frac{\lambda_3}{\sigma^2}(y-\mu)^3}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right), \quad (12)$$

where $\theta = (\mu, \sigma, \lambda_1, \lambda_2, \lambda_3)$. We use the notation $FSGN(\mu, \sigma, \lambda_1, \lambda_2, \lambda_3)$ to represent this distribution. It is obvious that when $\lambda_1 = \lambda_3 = 0$, the density (12) reduces to a $N(\mu, \sigma^2)$ distribution. For $\lambda_2 = \lambda_3 = 0$, (12) coincides with the $SN(\mu, \sigma, \lambda_1)$ density, that is, the distribution of a random variable $W = \mu + \sigma Z_{\lambda_1}$ where $Z_{\lambda_1} \sim SN(\lambda_1)$. In what follows, we shall use a data set concerning the waiting time (in minutes) of 299 consecutive eruptions of the Old Faithful geyser in Yellowstone National Park

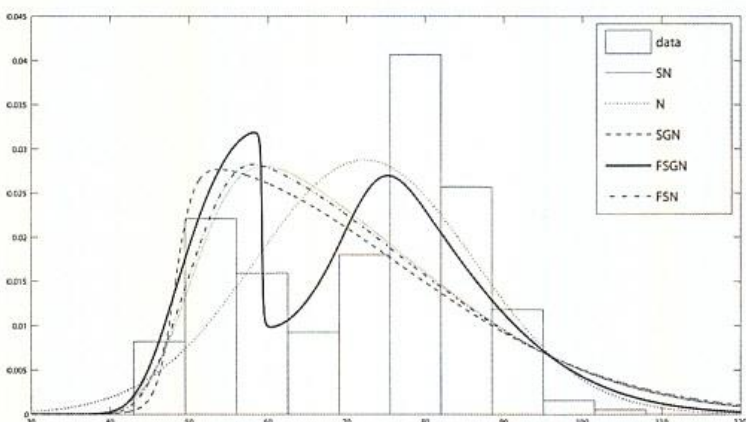

Figure 2. Histogram of the waiting time (in minutes) of 299 consecutive eruptions of the Old Faithful geyser in Yellowstone National Park and fitted densities. (color figure available online.)

of Azzalini and Bowman (1990) to examine our distribution. In order to compare *FSGN* with other competitors, we have obtained maximum likelihood estimation of θ using some numerical computations. The results of comparison of *FSGN* with four rival models are summarized in Table 1. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are used to compare the estimated models. As we observe, the preferred model is our *FSGN* with smallest values of AIC and BIC (see also Fig. 2). In addition, the test of hypotheses

$$\text{(a)}\quad H_0 : \lambda_2 = 0(FSN) \quad vs. \quad H_1 : \lambda_2 \neq 0(FSGN)$$

and

$$\text{(b)}\quad H_0 : \lambda_3 = 0(SGN) \quad vs. \quad H_1 : \lambda_3 \neq 0(FSGN),$$

can confirm our claim. The values of the likelihood ratio (LR) test statistics of (a) and (b) are 88.35 and 86.74, respectively, whereas the corresponding p-values are almost zero. Therefore, the null hypotheses are rejected in favor of the alternative hypotheses.

Acknowledgments

The authors sincerely thank two anonymous referees for their valuable comments that led to improvement of this article. The authors are also grateful to the Graduate Office of the University of Isfahan for their support.

References

Arellano-Valle, R. B., Gómez, H. W., Quintana, F. A. (2004). A new class of Skew-normal distributions. *Commun. Statist.* 33:1465–1480.

Arnold, B. C., Lin, G. D. (2004). Characterizations of the skew-normal and generalized Chi distributions. *Sankhyā* 66:593–606.

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

Azzalini, A., Bowman, A. W. (1990). A look at some data on the Old Faithful geyser. *Appl. Statist.* 39:357–365.

Behboodian, J., Jamalizadeh, A., Balakrishnan, N. (2006). A new class of skew-Cauchy distributions. *Statist. Probab. Lett.* 76:1488–1493.

Huang, W. J., Chen, Y. H. (2007). Generalized skew-Cauchy distribution. *Statist. Probab. Lett.* 77:1137–1147.

Ma, Y., Genton, M. G. (2004). Flexible class of skew-symmetric distributions. *Scan. J. Statist.* 31:459–468.

Nadarajah, S., Kotz, S. (2003). Skewed distributions generated by the normal kernel. *Statist. Probab. Lett.* 65:269–277.

Nekoukhou, V., Alamatsaz, M. H. (2012). A family of skew-symmetric-Laplace distributions. *Statist. Pap.* 53:685–696.