STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. 10, September 2022, pp 1126–1142.
Published online in International Academic Press (www.IAPress.org)

A New Weighted Skew Normal Model

Forough Naghibi , Sayed Mohammad Reza Alavi *, Rahim Chinipardaz

Deparment of Statistics, Faculty of Mathematics and Computer Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran

Abstract Weighted distribution is a valuable method for constructing flexible models and analyzing data sets. In this paper, a new weighted distribution of skew-normal is introduced with four parameters. The proposed model is a generalized version of several distributions, such as normal, bimodal normal, skew-normal, and skewed bimodal normal-normal. This weighted model is form-invariant under the proposed weight function. The basic characteristics of the model are expressed. A method has been used to generate data from the model. The maximum likelihood estimations of parameters are given and evaluated using a simulation study. The model is fitted to the three real data sets. The advantage of the proposed model has been shown on the rival distributions using appropriate criteria.

Keywords Skew normal distribution, Absolute-power weight function, simulation, maximum likelihood estimation, geyser data, pollen data, egg size data

AMS 2010 subject classifications 94A08, 68U10

DOI:10.19139/soic-2310-5070-1108

1. Introduction

Many data have a bimodal nature, symmetrical or asymmetrical, that the normal distribution is not suitable for fitting them. Azzalini [7] introduced the skew-normal distribution to model unimodal asymmetry data. This distribution has a skewness parameter, λ, and denoted by $SN(\lambda)$. Its probability density function, pdf, is

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

where $\phi(.)$ and $\Phi(.)$ are the pdf and cumulative distribution function (cdf) of standard normal distribution, respectively. By replacing a symmetric pdf instead of $\Phi(x)$ in (1), other skew normal distributions were introduced as skew-symmetric distributions based on Azzalini lemma (1985) [7].

Lemma 1
(Azzalini, [7]) Let f be a symmetric pdf about zero, and H is cdf of a symmetric distribution about zero. Then

$$f(x;\lambda) = 2f(x)H(\lambda x), \quad x \in R \tag{2}$$

is a pdf for any $\lambda \in R$.

Using this lemma, Gomez et al. [13] and Nekoukhou and Alamatsaz [30] introduced the skew t-normal distribution and skew symmetric-Laplace distribution, respectively. The generalizations of the skew-normal distribution of Azzalini [7] have been discussed by many authors. Nadarajah and Kotz

*Correspondence to: Sayed Mohammad Reza Alavi (Email: alavi_m@scu.ac.ir). Deparment of Statistics, Faculty of Mathematics and Computer Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran

ISSN 2310-5070 (online) ISSN 2311-004X (print)

[29] introduced the skew-normal-t distribution, skew-normal-Cauchy, skew normal-logistic, skew normal-uniform distributions. Yadegarietal et al. [38] and Sharafi and Behboodian [36] expressed generalizations of the Balakrishnan skew-normal distributions. Mameli and Musio [26] introduced the Beta skew-normal distributions. Ma and Genton [24] and Rasekhi et al. [34] expressed flexible classes of skew-symmetric distributions. Maleki and Nematollahi [25], and Arellano-valle et al. [4] examined the Bayesian estimates of the skew-Normal distributions. O'Hagan and Leonard [31] and Mukhopadhyay and Vidakovic [28] have used skew-symmetric models as priors in studying robustness. Some other studies in the field of skew models are Azzalini [6, 8], Azzalini and Regoli[10], Azzalini and Bowman [9], Gupta and Gupta [15], Arellano-Valle et al.[3], Henze [?] Arnold et al.[5], Martinez et al. [27], Jamalizadeh et al.[19], Gomez et al.[13], and Kumar and Anusree [22]. Weighted distribution introduced by Rao [33] is a useful method for constructing flexible models and analyzing data sets.

The weighted distributions have been used when the sampling mechanism records observations according to a certain chance. Suppose that $f(x;\theta)$ is pdf of the random variable X and the probability of the recording x of X is proportional to a non-negative weight function $w(x,\beta)$. The recorded x is an observation of X_w (weighted version of X) having the pdf

$$f^w(x;\theta,\beta) = \frac{w(x,\beta)f(x;\theta)}{(E[w(X,\beta)]}$$

where weight parameter $\beta$is a known or unknown. When $w(x,\beta) = x^\beta$, the weighted distribution called size-biased order β. The pdf of general weighted skew normal model is

$$f(x,\lambda,\beta) = Cw(x,\beta)\Phi(x)\Phi(\lambda x) \tag{3}$$

where $w(x,\beta)$ denotes weight function with weight parameter β and C represents the normalization constant given by

$$C^{-1} = \int w(x,\beta)\phi(x)\Phi(\lambda x)dx$$

In this paper the weight function $w(x\beta) = |x|^\beta$ is considered and called absolute-power order β weight function. The corresponding weighted model under this weight function is called weighted absolute-power skew normal order β. The symmetric version $(\lambda = 0)$ of model is called the weighted absolute-power normal order β (or bimodal normal order β). This symmetric distribution denoted by $BN(\beta)$, has been introduced by Alavi [1]. The pdf of $BN(\beta)$ is given by

$$g(x,\beta) = \frac{|x|^\beta}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}x^2} \quad x \in R \tag{4}$$

where $\beta \geq 0$ and $\Gamma(.)$ are mode parameter and gamma function, respectively. If $\beta = 0$, distribution is unimodal. The cdf of $BN(\beta)$ denoted by $G(x,\beta)$ is

$$G(x,\beta) = \begin{cases} \frac{1}{2} - \frac{1}{2}F_T(x^2), & x < 0 \\ \frac{1}{2} + \frac{1}{2}F_T(x^2), & x \geq 0 \end{cases} \tag{5}$$

where $F_T(.)$ is the cdf of the random variable T distributed as the gamma $\Gamma(\frac{\beta+1}{2}, \frac{1}{2})$. Alavi[1] described some of the properties of this distribution in the following theorem:

Theorem 1
Suppose $X \sim BN(\beta)$, then

1. The mean and variance of X are ?? and $(\beta+1)$, respectively.
2. $X^2 \sim \Gamma(\frac{\beta+1}{2}, 1/2)$

3. $Y = |X| \sim RHBN(\beta)$, where $RHBN(\beta)$ is right half bimodal normal distribution with the following pdf
$$f(y;\beta) = \frac{2y^\beta}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}} e^{\frac{1}{2}y^2}, \quad y \geq 0 \tag{6}$$
4. $Y = -|X| \sim LHBN(\beta)$, where $LHBN(\beta)$ is left half bimodal normal distribution with the following pdf
$$f(y;\beta) = \frac{2y^\beta}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}y^2}, \quad y \leq 0 \tag{7}$$

Proof See[1].
Alavi and Tarhani[2], based on Azzalini lemma [7] and the bimodal normal distribution order 2, introduced the skew bimodal normal distribution with three parameters (SBN). In this paper, a generalized version of [2] is introduced. In Section 2, the new weighted absolute-power skew-normal order β is introduced, and some special cases of the model are presented. The cdf, moments and one method for generating data from this model are presented in Section 3. In Section 4, the location-scale extension of the proposed model is given, and its parameters are estimated by the maximum likelihood method. We used the maximum likelihood methods, because of its asymptotically properties, such as efficiency and normality convergence. In Section 5, using the simulation study, the maximum likelihood estimation of parameters is evaluated. Section 6 is dedicated to the application of the proposed model to fit the actual three data sets. Section 7 presents the summarized results and suggestions.

2. Weighted absolute-power skew normal order $\boldsymbol{\beta}$

In this section, using Lemma 1 and pdf of $BN(\beta)$, the weighted absolute-power skew normal order β is considered as an element of skew-symmetric family, so it can be called skew bimodal normal-normal distribution and denoted by $SBNN(\lambda, \beta)$. The pdf of $SBNN(\lambda, \beta)$ is
$$f(x;\lambda,\beta) = 2g(x;\beta)\Phi(\lambda x), \quad x \in R, \ \lambda \in R \tag{8}$$
where λ is the skewness parameter and $g(.)$ is pdf of $BN(\beta)$. Figure ?? shows pdf of $SBNN(\lambda, \beta)$ for some values of λ and β. This distribution can model unimodal, bimodal, skew and heavy-tail datasets. The parameter λ controls skewness and kurtosis and the mode parameter, β, affects on shape of the distribution. The following distributions are special cases:

1. $SBNN(0, 0)$ is the standard normal distribution.
2. $SBNN(\lambda, 0)$ is $SN(\lambda)$.
3. $SBNN(0, \beta)$ is $BN(\lambda)$.
4. If $X \sim SBNN(\lambda, \beta)$, then $X \sim RHBN(\beta)$ as $\lambda \to \infty$.
5. If $X \sim SBNN(\lambda, \beta)$, then $X \sim LHBN(\beta)$ as $\lambda \to -\infty$.

An important property of weighted distributions is that the pdf of original and weighted distribution have the same form under weight function except possibly for a change in the parameters. This is known as the form-invariant property of weighted distributions. (see [1, 20]). Based on the theory of weighted distributions, the following results are easily obtained:

1. For even β, the size-biased order β of $SN(\lambda)$ is $SBNN(\lambda, \beta)$.
2. The weighted normal distribution under weight function $|x|^\beta\Phi(\lambda x)$ is $SBNN(\lambda, \beta)$.
3. The weighted $BN(\beta)$ under weight function $\Phi(\lambda x)$ is $SBNN(\lambda, \beta)$.
4. $SBNN(\lambda, \beta)$ under weight function $|x|^p$ is form-invarint. It means that the weighted $SBNN(\lambda, \beta)$ under weight function $|x|^p$ is $SBNN(\lambda, \beta + p)$

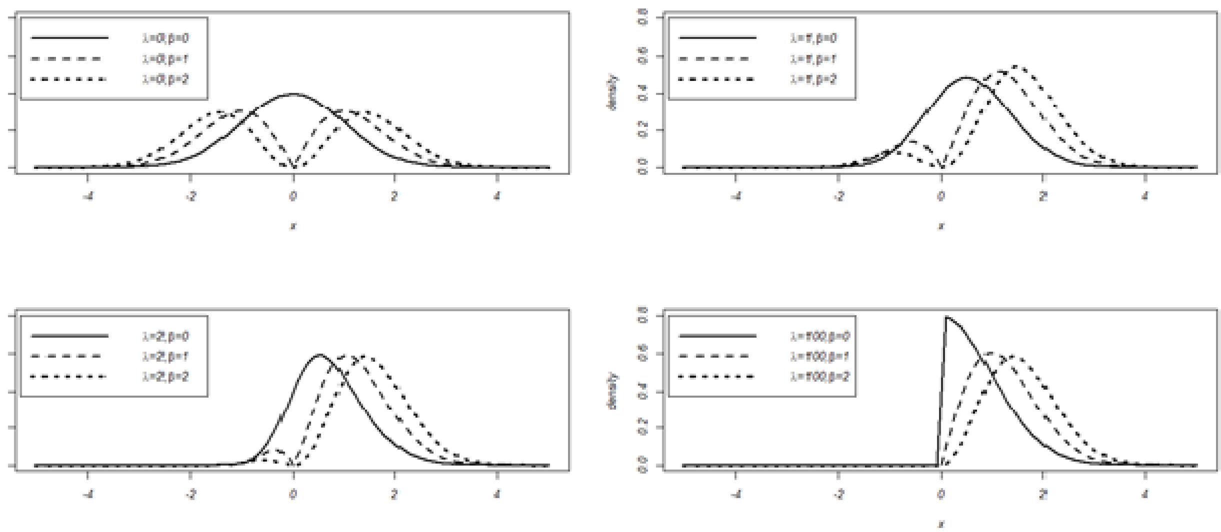

Figure 1. pdf of $SBNN(\lambda, \beta)$ for some values of λ and β.

3. Some Properties

Some properties of $SBNN(\lambda, \beta)$ are expressed in this section.

Theorem 2
If $G(x;\beta)$ and $F(x;\lambda,\beta)$ are the cdf of $BN(\beta)$ and $SBNN(\lambda,\beta)$, respectively, then for each real x and λ

$$F(x;\lambda,\beta) = G(x;\beta) - H(x;\lambda,\beta) \tag{9}$$

where

$$H(x;\lambda,\beta) = \frac{1}{\sqrt{\pi}\Gamma(\frac{\beta+1}{2})}\sum_{i=0}^{\infty}\frac{(-1)^i\lambda^{1+2i}\Gamma(\frac{\beta+2}{2}+i, \frac{1}{2}x^2)}{i!(1+2i)}. \tag{10}$$

Proof Without loss of generality, we assume that $\lambda \geq 0$ and $x \geq 0$ in (8). We have

$$\begin{aligned} F(x;\lambda,\beta) &= \int_{-\infty}^{x} 2g(t;\beta)\Phi(\lambda t)dt \\ &= \int_{-\infty}^{x}\int_{-\infty}^{\lambda t} 2g(t,\beta)\phi(u)dudt \end{aligned}$$

and

$$\begin{aligned} 2G(x;\beta) &= \int_{-\infty}^{x} 2g(t;\beta)dt = \int_{-\infty}^{x} 2g(t;\beta)\Big[\int_{-\infty}^{\lambda t}\Phi(u)du + \int_{\lambda t}^{+\infty}\phi(u)du\Big]dt \\ &= F(x;\lambda,\beta) + \int_{-\infty}^{x}\int_{\lambda t}^{0} 2g(t;\beta)\phi(u)dudt + \int_{-\infty}^{x}\int_{0}^{\infty} 2g(t;\beta)\phi(u)dudt \\ &= F(x;\lambda,\beta) + \int_{-\infty}^{x}\int_{\lambda t}^{0} 2g(t;\beta)\phi(u)dudt + G(x;\beta). \end{aligned}$$

Then

$$\begin{aligned}F(x;\lambda,\beta) &= G(x;\beta)-\int_{-\infty}^{x}\int_{\lambda t}^{0}2g(t;\beta)\phi(u)dudt = G(x;\beta)+\int_{-\infty}^{x}\int_{0}^{\lambda t}2g(t;\beta)\phi(u)dudt\\ &= G(x;\beta)+\Big[\int_{-\infty}^{\infty}\int_{0}^{\lambda}t2g(t;\beta)\phi(u)dudt-\int_{x}^{\infty}\int_{0}^{\lambda}t2g(t;\beta)\phi(u)dudt\Big]\\ &= G(x;\beta)+\int_{-\infty}^{\infty}2g(t;\beta)[\Phi(\lambda t)-0.5]dt-H(x;\lambda,\beta)=G(x;\beta)-H(x;\lambda,\beta)\end{aligned}$$

where

$$\begin{aligned}H(x;\lambda,\beta) &= \int_{x}^{\infty}\int_{0}^{\lambda t}2g(t)\phi(u)dudt=\int_{x}^{\infty}\frac{|t|^{\beta}e^{-\frac{1}{2}t^2}\gamma(\frac{1}{2},\frac{1}{2}\lambda^2t^2)}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}\sqrt{\pi}}dt\\ &= \int_{x}^{\infty}t^{\beta}e^{-\frac{1}{2}t^2}\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}\sqrt{\pi}\sum_{i=0}^{\infty}\frac{(-1)^i\big(\frac{1}{2}\lambda^2t^2\big)^{\frac{1}{2}+i}}{i!(\frac{1}{2}+i)}dt.\end{aligned}$$

If the transformation $y=\frac{1}{2}t^2$ is used, then (10) is obtained using the formulas 3.381(1) and 8.354(1) in [?] and $\Gamma(\alpha,x)=\int_x^{\infty}e^{-t}t^{\alpha-1}dt$. The following properties are easily obtained for function $H(x;\lambda,\beta)$

$$H(-x;\lambda,\beta) = H(x;\lambda,\beta) \quad (11)$$

$$H(x;-\lambda,\beta) = -H(x;\lambda,\beta). \quad (12)$$

For the negative values of x and λ using (11) and (12) and $(\lambda x)=1-\Phi(-\lambda x)$, (9) is also confirmed. So the proof is complete.

Theorem 3
Suppose that $X\sim SBNN(\lambda,\beta)$, then

1. $-X\sim SBNN(-\lambda,\beta)$
2. $F(x,-\lambda,\beta)=1-F(-x,\lambda,\beta)$
3. $X^2\sim\Gamma(\frac{\beta+1}{2},\frac{1}{2})$
4. $|X|\sim RHBN(\lambda,\beta)$
5. $-|X|\sim LHBN(\lambda,\beta)$

Proof (1) is straightforward using (8). To prove (2), we have

$$\begin{aligned}F(x;-\lambda,\beta) &= \int_{-\infty}^{x}2g(t;\beta)\Phi(-\lambda t)dt\\ &= \int_{-\infty}^{x}2g(t;\beta)[1-\Phi(\lambda t)]dt=2G(x;\beta)-F(x;\lambda,\beta)\end{aligned}$$

using (8), we have

$$F(x;-\lambda,\beta)=G(x;\beta)+H(x;\lambda,\beta)$$

and

$$1-F(-x;\lambda,\beta)=1-[G(-x;\beta)-H(-x;\lambda,\beta)]=G(x;\beta)+H(-x;\lambda,\beta)$$

according to equation (10) the proof (2) is completed. To prove (3), using transformation $Y=X^2$, the pdf of Y is given by

$$\begin{aligned}f_Y(y;\lambda) &= \frac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})]=\frac{1}{2\sqrt{y}}[2g(\sqrt{y})\Phi(\lambda\sqrt{y})+2g(-\sqrt{y})\Phi(-\lambda\sqrt{y})]\\ &= \frac{g(\sqrt{y})}{\sqrt{y}}=\frac{y^{\frac{\beta+1}{2}-1}}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}}e^{-\frac{1}{2}y},\quad y>0\end{aligned}$$

that is pdf of the gamma distribution $\Gamma(\frac{\beta+1}{2}, \frac{1}{2})$. Proof (4) and (5) are similar to (3). ¡

To obtain the moments of $SBNN(\lambda, \beta)$, the following Lemma and Theorem are necessary.

Lemma 2
Suppose $f(y)$ is pdf of a random variable Y having a symmetric distribution about zero and H is cdf of a symmetric distribution about zero. If pdf of random variable X is given by

$$f(x;\lambda) = 2f(x)H(\lambda x), \quad x \in R,$$

then

1. The even moments of X are independent of λ and the same as Y.
2. X^2 and Y^2 have the same distribution.
3. $|X|$ and $|Y|$ have the same distribution.

Proof See Gupta et al. (2002)[16].

Theorem 4
Under the assumptions of Lemma 2, let $U = |Y|$ and

$$T = \begin{cases} 1, & \text{with probability } H(\lambda u) \\ -1, & \text{with probability } 1 - H(\lambda u) \end{cases}$$

Then $X = TU$ has pdf (2).

Proof See Azzalini (1986)[6].

Theorem 5
Suppose that $X \sim SBNN(\lambda, \beta),\ V = |X|, Y \sim BN(\beta)$ and $U = |Y|$, then

1. For k even

$$E(X^k) = E(V^k) = E(U^k) = E(Y^k) = \frac{\Gamma\left(\frac{\beta+k+1}{2}\right)2^{\frac{k}{2}}}{\Gamma\left(\frac{\beta+1}{2}\right)},$$

2. For k odd

$$\begin{aligned} E(X^k) &= 2E\{\Phi(\lambda U)U^k\} - E(U^k) \\ &= \frac{2^{\frac{k}{2}+1-c}\lambda}{\Gamma\left(\frac{\beta+1}{2}\right)(1+\lambda^2)^{c+\frac{1}{2}}} \sum_{(n=0}^{\infty} \frac{(2c+2n-1)!!}{(2n+1)!!}\left[\frac{\lambda^2}{1+\lambda^2}\right]^n \end{aligned}$$

where $c = (k+\beta+1)/(2)$ and the double factorial operator for odd and even numbers is equal to $(2n-1)!! = \prod_{i=1}^{n}(2i-1)$ and $(2n)!! = \prod_{i=1})^n(2i) = 2^n n!$.

Proof. Without loss of generality we assume that $\lambda \geq 0$ in (7). I. According to Lemma 2

$$E(X^k) = E(V^k) = E(U^k) = E(Y^k)$$

So

$$\begin{aligned} E(Y^k) &= \int_{-\infty}^{\infty} \frac{y^k|y|^\beta}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}} \\ &= \int_{-\infty}^{0} \frac{y^k(-y)^\beta}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}y^2}dy + \int_{0}^{\infty} \frac{y^k y^\beta}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}y^2}dy. \end{aligned}$$

By transformation $u=-y$ in the first integral, we have:

$$E(Y^k)=\frac{2}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}}\int_0^\infty u^{\beta+k}e^{-\frac{1}{2}u^2}du.$$

and using $t=u^2$

$$E(Y^k)=\frac{1}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}}\int_0^\infty t^{\frac{\beta+k+1}{2}-1}e^{-\frac{1}{2}t}dt=\frac{\Gamma\left(\frac{\beta+k+1}{2}\right)2^{\frac{k}{2}}}{\Gamma\left(\frac{\beta+1}{2}\right)} \tag{13}$$

II. According to Theorem 4, we have:

$$\begin{aligned} E(X^k) &= E(TU^k)=E\{E(TU)U^k\} \\ &= E\{(H(\lambda U)-H(-\lambda U))U^k\}=2E\{(H(\lambda U)-1/2)U^k\}. \end{aligned}$$

Thus

$$E(X^k)=2E\{H(\lambda U)U^k\}-E(U^k). \tag{14}$$

For $SBNN(\lambda,\beta),\ H(.)=\Phi(.)$, so

$$E(X^k)=2E\{\Phi(\lambda U)U^k\}-E(U^k). \tag{15}$$

On the other hand

$$\begin{aligned} E\{\Phi(\lambda U)U^k\} &= \int_0^\infty \frac{u^k\Phi(\lambda u)2u^\beta}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}}e^{-\frac{1}{2}u^2}du \\ &= \frac{2}{\Gamma\left(\frac{\beta+1}{2}\right)2^{\frac{\beta+1}{2}}}\int_0^\infty u^{k+\beta}e^{-\frac{1}{2}u^2}\Phi(\lambda u)du. \end{aligned}$$

Using transformation $t=\frac{u^2}{2}$

$$E\{\Phi(\lambda U)U^k\}=\frac{2^{\frac{k}{2}}}{\Gamma\left(\frac{\beta+1}{2}\right)}\int_0^\infty t^{\frac{k+\beta+1}{2}-1}e^{-t}\Phi(\lambda\sqrt{2t})dt.$$

Assuming $c=(k+\beta+1)/(2)$ and using the standard extension

$$\Phi(x)=\frac{1}{2}+\Phi(x)\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!!}$$

we have

$$E\{\Phi(\lambda U)U^k\}=\frac{2^{\frac{k}{2}}}{\Gamma\left(\frac{\beta+1}{2}\right)}\left(\frac{1}{2}\Gamma(c)+\sum_{n=0}^{\infty}\frac{2^n\lambda^{2n+1}\Gamma\left(c+n+\frac{1}{2}\right)}{(2n+1)!!\sqrt{\pi}(1+\lambda^2)^{c+n+\frac{1}{2}}}\right).$$

So for $\lambda\geq 0$

$$E(X^k)=\frac{2^{\frac{k}{2}+1}\lambda}{\Gamma\left(\frac{\beta+1}{2}\right)\sqrt{\pi}(1+\lambda^2)^{c+\frac{1}{2}}}\sum_{n=0}^{\infty}\frac{\Gamma\left(c+n+\frac{1}{2}\right)}{(2n+1)!!}\left[\frac{2\lambda^2}{1+\lambda^2}\right]^n. \tag{16}$$

When $c=(k+\beta+1)/2$ is an integer, $\Gamma\left(n+\frac{1}{2}\right)=((2n-1)!!\sqrt{\pi})/2^n$ and the expression (16) simplifies to

$$E(X^k)=\frac{2^{\frac{k}{2}+1-c}\lambda}{\Gamma\left(\frac{\beta+1}{2}\right)(1+\lambda^2)^{c+\frac{1}{2}}}\sum_{n=0}^{\infty}\frac{(2c+2n-1)!!}{(2n+1)!!}\left[\frac{\lambda^2}{1+\lambda^2}\right]^n.$$

For $\lambda < 0$ the proof can be obtained using $\Phi(\lambda x) = 1 - \Phi(-\lambda x)$.
As an example, if $\beta = 2$ and $k = 1$, then $c = 2$ and $E(X) = \sqrt{2/\pi}(3\delta - \delta^3)$ where $\delta = \lambda/(\sqrt{1+\lambda^2})$. According to the idea described by Azzalini[8], the following theorem is presented for generating data from $SBNN(\lambda, \beta)$.

Theorem 6
Suppose that $Z \sim N(0,1)$ and $Y \sim BN(\beta)$ are two independent random variables. If the random variable X is defined as

$$X = \begin{cases} Y, & Z \leq \lambda Y \\ -Y, & Z > \lambda Y, \end{cases}$$

then $X \sim SBNN(\lambda, \beta)$.

Proof

$$\begin{aligned} P(X \leq x) &= \int_{-\infty}^{x} P(Z < \lambda Y Y = y) g(y;\beta) dy + \int_{-\infty}^{x} P(Z < \lambda Y - Y = y) g(-y;\beta) dy \\ &= \int_{-\infty}^{x} \Phi(\lambda y) g(y;\beta) dy + \int_{-\infty}^{x} [1 - \Phi(-\lambda y)] g(y;\beta) dy = \int_{-\infty}^{x} 2 g(y;\beta) \Phi(\lambda y) dy \end{aligned}$$

that is cdf of $SBNN(\lambda, \beta)$. Thus $X \sim SBNN(\lambda, \beta)$ and the proof is completed. Using the Theorem 6, a method for generating data from $SBNN(\lambda, \beta)$ is introduced by implementing the following steps:

1. Generate a random value Z, from standard normal distribution.

2. Generate a random value Y, from $BN(\beta)$ distribution.

3. If $\lambda Y > Z$, then $X = Y$; otherwise, $X = -Y$.
   When these three steps are repeated n times, a random sample of size n can be generated from $SBNN(\lambda, \beta)$.
   The exact and simulated density of $SBNN(1, 1)$, are shown in Figure 2.

4. Location-Scale model

Suppose $Y \sim SBNN(\lambda, \beta)$. By definition $X = \mu + \sigma Y$, the random variable X has Location-Scale distribution of weighted absolute-power skew normal order β. This distribution is denoted by $SBNN(\mu, \sigma, \lambda, \beta)$, where $\mu \in R$ and $\sigma > 0$. The pdf of X is

$$f_X(x;\beta,\mu,\sigma) = \frac{2|x-\mu|^\beta}{\Gamma\big(\frac{\beta+1}{2}\big) 2^{\frac{\beta+1}{2}} \sigma^{\beta+1}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \Phi\big(\lambda \frac{x-\mu}{\sigma}\big) \tag{17}$$

Note that $SBNN(0, 1, \lambda, \beta) = SBNN(\lambda, \beta)$.

Using Theorem 5 and Newton's binomial expansion, the k^{th} non-central moment of $SBNN(\mu, \sigma, \lambda, \beta)$ is equal to

$$E(X^k) = \sum_{j=0}^{k} \binom{k}{j} \mu^{k-j} \sigma^j E(Y^j).$$

This distribution is unimodal for $\beta = 0$ or large values of λ and bimodal for small values of λ. The parameters make $SBNN(\mu, \sigma, \lambda, \beta)$?exible distribution for uni/bimodal data. For the special case $\lambda = 0$,

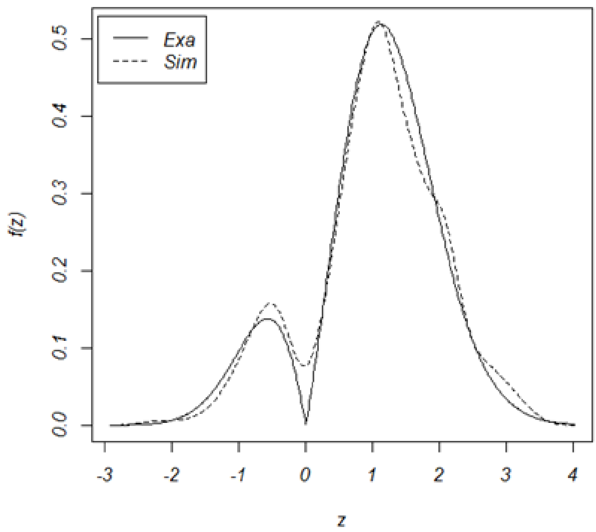

Figure 2. The exact and simulated density of $SBNN(1,1)$

the Bimodal Normal distribution order β with mean μ and variance $(\beta+1)\sigma^2$ denoted by $BN(\mu,\sigma,\beta)$, is resulted[1]. If the major and minor modes are shown with m_1 and m_2, respectively, For $\lambda > 0\,(\lambda < 0)$ inequality $m_1 > \mu > m_2\,(m_1 < \mu < m_2)$ exists between the modes and the location parameter. When $|\lambda|$ increases, the $SBNN(\mu,\sigma,\lambda,\beta)$ tends to a unimodal distribution, such that for $\lambda \to \infty(\lambda \to -\infty)$ the Right Half Bimodal Normal error β denoted by $RRHBN(\mu,\sigma,\beta)$ (Left Half Bimodal error β Normal denoted by $LHBN(\mu,\sigma,\beta)$ with the following pdf is obtained:

$$f_X(x;\beta,\mu,\sigma,\lambda) = \frac{2|x-\mu|^\beta}{\Gamma\big(\frac{\beta+1}{2}\big)2^{\frac{\beta+1}{2}}\sigma^{\beta+1}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\Phi\big(\lambda\frac{x-\mu}{\sigma}\big)$$

where $x \geq \mu(x \leq \mu)$.

4.1. The maximum likelihood estimation of parameters

Suppose $x_1, x_2, \ldots, x_n$ is a observed random sample of from $SBNN(\mu,\sigma,\lambda,\beta)$, then the log-likelihood function is

$$\begin{aligned} \ell(\mu,\sigma,\lambda,\beta) &= -n\big(\frac{\beta+1}{2}\big)\log 2 - n\log\Gamma\big(\frac{\beta+1}{2}\big) - n(\beta+1)\log\sigma \\ &+ \sum_{i=1}^{n}\log?(|x_i-\mu|^\beta) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 + \sum_{i=1}^{n}\log\Phi\big(\lambda\frac{x_i-\mu}{\sigma}\big). \end{aligned} \quad (18)$$

The maximum likelihood estimate (MLE) of μ, σ, λ and β are obtained by solving simultaneously the following equations using numerical methods such as Newton-Raphson iteration.

$$-\beta\sum_{i=1}^{n}\frac{1}{(x_i-\mu)} + \sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2} - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{\phi\big(\lambda\frac{x_i-\mu}{\sigma}\big)}{\Phi\big(\lambda\frac{x_i-\mu}{\sigma}\big)} = 0 \quad (19)$$

Table 1. Simulated means and standard errors for the MLEs of μ, σ, λ and $\beta(n = 200)$

μ	σ	λ	β	E[$\hat{\mu}$] (SD[$\hat{\mu}$])	E[$\hat{\sigma}$] (SD[$\hat{\sigma}$])	E[$\hat{\lambda}$] (SD[$\hat{\lambda}$])	E[$\hat{\beta}$] (SD[$\hat{\beta}$])
1	1	0	0.6	1.000 (0.041)	0.993(0.057)	-0.003(0.075)	0.137(0.632)
0	1	0.5	4	0.001 (0.061)	0.992 (0.053)	0.501 (0.061)	4.108 (0.484)
0	1	1	3	-0.003 (0.078)	0.991 (0.053)	1.008 (0.126)	3.112 (0.435)
0	1	1	1	0.003 (0.041)	0.993 (0.059)	1.007 (0.136)	1.044 (0.190)

Table 2. Simulated means and standard errors for the MLEs of μ, σ, λ and $\beta(n = 500)$

μ	σ	λ	β	$E[\hat{\mu}](SD[\hat{\mu}])$	$E[\hat{\sigma}](SD[\hat{\sigma}])$	$E[\hat{\lambda}](SD[\hat{\lambda}])$	$E[\hat{\beta}](SD[\hat{\beta}])$
1	1	0	0.6	0.999 (0.023)	0.999 (0.036)	0.000- (0.047)	0.087 (0.609)
0	1	0.5	4	-0.002 (0.039)	1.000 (0.034)	0.498 (0.038)	4.034 (0.304)
0	1	1	3	-0.002 (0.045)	0.997 (0.034)	1.003 (0.077)	3.038 (0.252)
0	1	1	1	-0.001 (0.025)	0.997 (0.035)	1.005 (0.082)	1.017 (0.111)

$$-\frac{n(\beta+1)}{\sigma} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^3} - \frac{\lambda}{\sigma^2} \sum_{i=1}^{n} \frac{(x_i-\mu)\phi\left(\lambda \frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda \frac{x_i-\mu}{\sigma}\right)} = 0 \tag{20}$$

$$\frac{1}{\sigma} \sum_{i=1}^{n} \frac{(x_i-\mu)\phi\left(\lambda \frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda \frac{x_i-\mu}{\sigma}\right)} = 0 \tag{21}$$

$$-\frac{1}{2}\Big[n\Psi_0\Big(\frac{\beta+1}{2}\Big) - 2\sum_{i=1}^{n} \log?(|x_i-\mu|) + (2\log\sigma + \log 2)n\Big] = 0 \tag{22}$$

where $\Psi_0(z) = \frac{d}{dz}\log\Gamma(z)$. The MLE of parameters can be calculated using the optim or nlminb commands in R package. The existence of MLE for λ parameter in SN distribution is discussed in [23] and [32]. They have explained if at least two elements of sample have different sign, then the MLE of parameter λ exists. The same result is given in[27] for each class of distributions with the following pdf

$$f(x) = 2h(x)\Phi(\lambda x) \tag{23}$$

where $h(.)$ is a symmetric pdf. Therefore, this condition is necessary for MLE of λ because the $SBNN$ is a member of this class. Asymptotic distribution of the MLE of parameters is multivariate normal distribution with mean vector $(\mu, \sigma, \lambda, \beta)'$ and covariance matrix of inversed Fisher information. The score vector and Hessian matrix are given in the ¡°Appendix¡±. Because the closed forms are not available for MLEs, they are evaluated using a simulation study.

5. Simulation study

In this section, the expected value and standard error (in parentheses) of $MLEs$ are obtained using the following simulation steps: First, a sample of size n (200 and 500) is generated from the $SBNN(\mu, \sigma, \lambda, \beta)$ for known parameters. In the second step, for each sample in step 1, the MLEs are computed using the optim command by the $L-BFGS-B$ method in the R package. Steps 1 and 2 are repeated 1000 times, then for each parameter, the mean and standard deviation of these 1000 repetitions are calculated as the simulated mean and standard error of the MLE. The results are presented in Tables 1 and 2. The tables show that the estimators are unbiased asymptotically, and their efficiency increase when sample size increases.

Figure 3 shows the densities of the simulated samples from $SBNN(1,1,1,1)$ for $n = 50, 200, 500, 1000$.

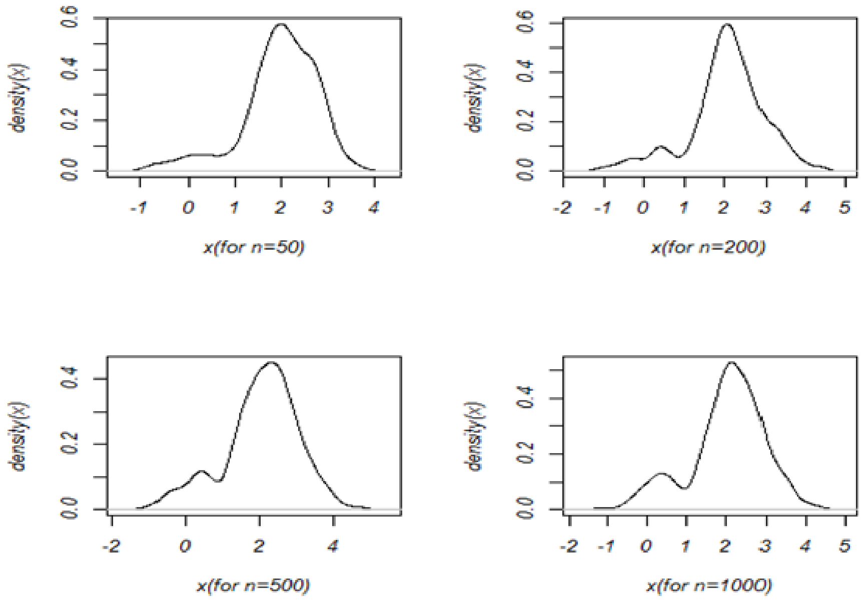

Figure 3. Densities of simulated samples from SBNN(1,1,1,1)

Table 3. Descriptive statistics of datasets

data	n	$\bar{x}$	s	β_1	β_2
Geyser data (duration variable)	299	3.46	1.14	-0.45	-1.43
Egg size data (log of the egg diameter)	88	6.07	0.98	0.13	-1.21
Pollen data (nub variable)	481	-0.04	5.19	0.23	-0.41

6. Applications

In this section, the $SBNN$ is fitted to the three datasets. Table 3 shows descriptive statistics for these datasets, where β_1 and β_2 represent the coefficients of skewness and kurtosis, respectively. The first dataset is used by Azzalini and Bowman [9]. This dataset is in "$MASS$" in R package called "geyser". The data consists of 299 pairs of measurements, referring to the time interval between the starts of successive eruptions (waiting variable) and the duration of the subsequent eruption (duration variable) of the Old Faithful geyser in Yellow stone National Park, Wyoming, USA. The duration variable is considered in this article. The second dataset, called "Egg size data", is studied by Sewell and Young [35] and Famoye et al. [11]. The data represent the logarithm of the egg diameters of 88 asteroid species. The histogram of the data is roughly symmetric bimodal, as shown in Figure 6. The third dataset is called "Pollen data". The nub variable of pollen data is available at http://lib.stat.cmu.edu/datasets/pollen.data. This data are resulted from measuring geometric characteristics of a certain type of pollen. Figure 3 displays that the distribution of the nub variable is bimodal.

Comparison of the fitted Models to the first dataset The pdf of SBN ([2]), $SSCN$([34]), BEP ([17]), $Mixed-N2$ (two-component mixture normal with five parameters $\mu_1, \sigma_1, \mu_2, \sigma_2$ and ρ) and SBNN (proposed model) are fitted to duration variable of the Old Faithful geyser data. The MLE of parameters

Table 4. The MLE and standard error (in parenthesis) of parameters of the models for the duration data

Model					
SBNN	$\hat{\mu} = 3.116(0.018)$	$\hat{\sigma} = 0.546(0.023)$	$\hat{\lambda} = 0.170(0.036)$	$\hat{\beta} = 3.802(0.368)$	
SBN	$\hat{\mu} = 3.117(0.025)$	$\hat{\sigma} = 0.691((0.017)$	$\hat{\lambda} = 0.215(0.046)$		
SSCN	$\hat{\mu} = 3.117(-)$	$\hat{\sigma} = 0.691(-)$	$\hat{\lambda} = 0.215(-)$	$\hat{\alpha} = 3196.176(-)$	
BEP	$\hat{\mu} = 3.149(0.020)$	$\hat{\sigma} = 1.306(0.091)$	$\hat{\delta} = 2.330(0.333)$	$\hat{\alpha} = 3.746(0.489)$	
Mixed-N2	$\hat{\mu}_1 = 4.174(0.040)$	$\hat{\sigma}_1 = 0.559(0.031)$	$\hat{\mu}_2 = 1.927(0.016)$	$\hat{\sigma}_2 = 0.144(0.015)$	$\hat{\rho} = 0.685(0.027)$

Table 5. Various goodness of fit criteria and Kolmogorov-Smirnov test statistic for the candidate distributions (duration data)

Model	AIC	AICC	BIC	KS-Test
SBNN	651.21	643.35	666.02	0.105
SBN	682.15	676.23	693.25	0.140
SSCN	684.16	676.29	698.96	0.140
BEP	657.26	649.39	672.06	0.186
Mixed-N2	623.03	613.23	641.53	0.147

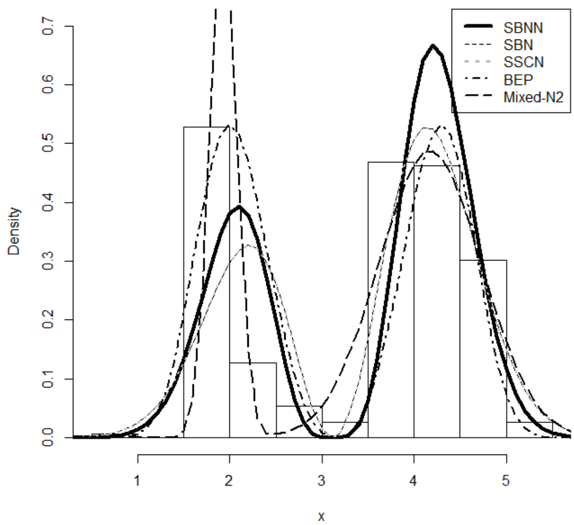

Figure 4. Histogram and pdf of the fitted models for duration data

for these models is calculated in Table 4. The Akaike criterion (AIC), modified Akaike criterion ($AICC$), Bayesian criterion (BIC), and Komogorov-Sminorv test (KS-Test) for the goodness of fit are given in Table 5. Based on KS- Test, in Table 5 the $SBNN$ is the best model among the rival models. But Based on AIC, $AICC$ and BIC criteria in Table 5 the $SBNN$ is the best model among the rival models except for Mixed-$N2$. Note that the $SBNN$ distribution doesn¡¯t need identifiability condition (switch case) for fitting data. The histogram of data and the pdf of fitted models are shown in Figure 4. Figure 4 confirms the superiority of $SBNN$.

Table 6. The MLE and standard error (in parenthesis) of parameters of the models for the Egg size data

Model					
SBNN	$\hat{\mu} = 5.826(0.042)$	$\hat{\sigma} = 0.704(0.060)$	$\hat{\lambda} = 0.221(0.108)$	$\hat{\beta} = 1.069(0.275)$	
SBN	$\hat{\mu} = 5.826(0.031)$	$\hat{\sigma} = 0.585(0.026)$	$\hat{\lambda} = 0.183(0.086)$		
SSCN	$\hat{\mu} = 5.904(0.092)$	$\hat{\sigma} = 0.621(0.034)$	$\hat{\lambda} = 0.133(0.113)$	$\hat{\alpha} = 3.762(2.193)$	
BEP	$\hat{\mu} = 5.865(0.048)$	$\hat{\sigma} = 1.075(0.315)$	$\hat{\delta} = 0.949(0.448)$	$\hat{\alpha} = 2.177(0.708)$	
Mixed-N2	$\hat{\mu}_1 = 6.746(0.091)$	$\hat{\sigma}_1 = 0.606(0.071)$	$\hat{\mu}_2 = 5.001(0.040)$	$\hat{\sigma}_2 = 0.223(0.030)$	$\hat{\rho} = 0.612(0.054)$

Table 7. Various goodness of fit criteria and Kolmogorov-Smirnov test statistic for the candidate distributions (Egg size data)

Model	AIC	AICC	BIC	KS-Test
SBNN	229.84	222.32	239.74	0.093
SBN	236.51	230.80	243.94	0.131
SSCN	230.61	223.09	240.52	0.094
BEP	234.35	226.84	244.26	0.140
Mixed-N2	212.61	203.34	225.00	0.068

Table 8. The MLE and standard error (in parenthesis) of parameters of the modelsfor the Pollen data

Model					
SBNN	$\hat{\mu} = 2.009(0.085)$	$\hat{\sigma} = 4.935(0.196)$	$\hat{\lambda} = -0.447(0.061)$	$\hat{\beta} = 0.280(0.069)$	
SBN	$\hat{\mu} = 4.419(0.018)$	$\hat{\sigma} = 3.954(0.073)$	$\hat{\lambda} = -0.706(0.057)$		
SSCN	$\hat{\mu} = 1.659(0.599)$	$\hat{\sigma} = 4.017(0.201)$	$\hat{\lambda} = -0.306(0.119)$	$\hat{\alpha} = 0.738(0.190)$	
BEP	$\hat{\mu} = 0.20(-)$	$\hat{\sigma} = 8.252(0.366)$	$\hat{\delta} = 0.000(-)$	$\hat{\alpha} = 2.690(0.283)$	
Mixed-N2	$\hat{\mu}_1 = 3.845(1.555)$	$\hat{\sigma}_1 = 3.954(0.676)$	$\hat{\mu}_2 = -3.591(0.895)$	$\hat{\sigma}_2 = 3.298(0.365)$	$\hat{\rho} = 0.457(0.159)$

Comparison of fitted Model to the second dataset

Similar to the first dataset, the pdf of SBN, $SSCN$, BEP, $Mixed-N2$ and $SBNN$ are fitted to the Egg size data. The MLEs of parameters for these models are calculated in Table 6. The AIC, $AICC$, BIC and KS-Test for the goodness of fit are given in Table 7. Table 7 shows that the $SBNN$ is the best model among the rival models except $Mixed-N2$. But the $SBNN$ distribution does not need a identifiability condition (switch case) for fitting data. Histogram of data and the pdf of fitted models are shown in Figure 5.

Comparison of fitted Model to the third dataset

Similar to the two previous datasets, the pdf of SBN, SSCN, BEP, Mixed-N2 and SBNN are fitted to the Pollen data. The MLE of parameters for these models is calculated in Table 8. The AIC, AICC, BIC and KS-Test for the goodness of fit are given in Table 9. Table 9 shows that the SBNN is the best model among the rival models. Histogram of data and pdf of fitted models are shown in Figure 6. Figure 6 confirms the superiority of SBNN.

7. Conclusion

In this paper, a new weighted model with four parameters of skew-normal distributions called weighted absolute-power skew normal of order β, was introduced. The normal distribution, the skew-normal

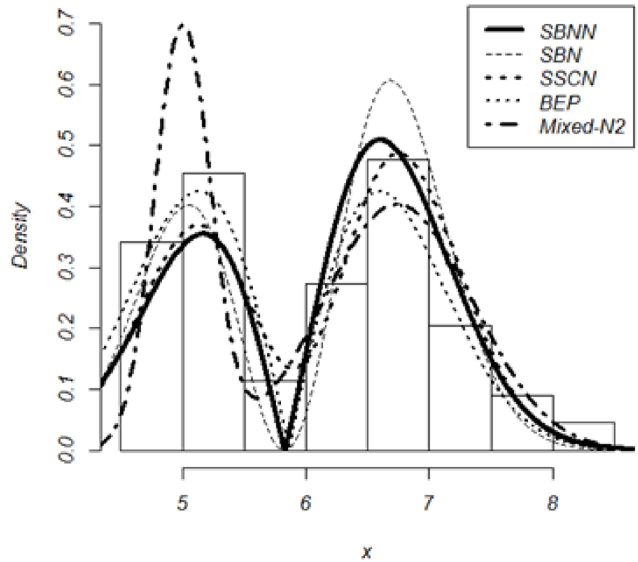

Figure 5. Histogram and pdf of the fitted models for Egg size data

Table 9. Various goodness of fit criteria and Kolmogorov-Smirnov test statistic for the candidate distributions (Pollen data)

Model	AIC	AICC	BIC	KS-Test
SBNN	2939.11	2931.19	2955.81	0.02
SBN	3359.07	3353.12	3371.60	0.16
SSCN	2940.94	2933.03	2957.65	0.02
BEP	2949.60	2941.68	2966.30	0.05
Mixed-N2	2942.16	2932.29	2963.47	0.02

distribution, the bimodal normal distribution and the skew bimodal normal distribution are special cases of this model. A method for generating data from this model was presented. The maximum likelihood estimates of parameters were obtained by numerical methods and evaluated using a simulation study. This model was fitted to the duration of the eruption of the famous Old Faithful geyser data, Egg size data and Pollen data. The superiority of the model was shown by some goodness of fit criteria on the rival distributions. Although the estimates proposed in this paper have been shown to work better than other methods, comparing these estimates with Bayesian estimates may be the subject of further research.

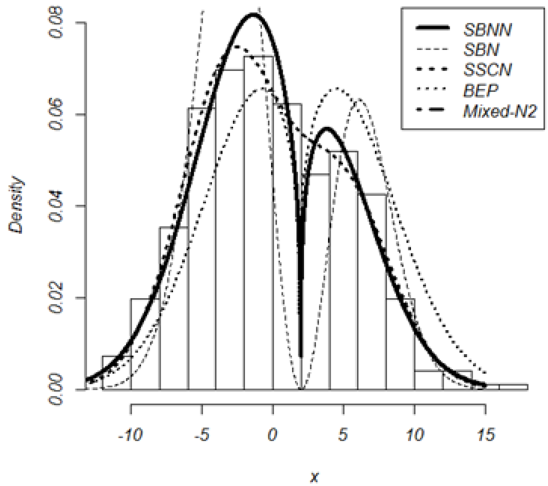

Figure 6. Histogram and pdf of the fitted models for Pollen datat

Acknowledgement

Appendix

Score vector and hessian matrix

Suppose $x_1, x_2, \ldots, x_n$ is an observed random sample of size n from $SBNN(\mu,\sigma,\lambda,\beta)$ with the log-likelihood function (18). The following elements of the score vector are obtained by deriving from (18) relative to the parameters

$$\begin{aligned}
\ell_\mu &= -\beta\sum_{i=1}^{n}\frac{1}{(x_i-\mu)}+\sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2}-\frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_\sigma &= -\frac{n(\beta+1)}{\sigma}+\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^3}-\frac{\lambda}{\sigma^2}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_\lambda &= \frac{1}{\sigma}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_\beta &= -\frac{n\Psi_0\left(\frac{\beta+1}{2}\right)-2\sum_{i=1}^{n}\log ?(|x_i-\mu|)+(2\log\sigma+\log 2)n}{2}
\end{aligned}$$

where $\Psi_0(z)=\frac{d}{dz}\log\Gamma(z)$ and $\Psi_n(z)=\frac{d^n}{dz^n}\log\Gamma(z)$.
The hessian matrix is given by

$$\begin{bmatrix}
\ell_{\mu,\mu} & \ell_{\mu,\sigma} & \ell_{\mu,\lambda} & \ell_{\mu,\beta}\\
\ell_{\sigma,\mu} & \ell_{\sigma,\sigma} & \ell_{\sigma,\lambda} & \ell_{\sigma,\beta}\\
\ell_{\lambda,\mu} & \ell_{\lambda,\sigma} & \ell_{\lambda,\lambda} & \ell_{\lambda\beta}\\
\ell_{\beta,\mu} & \ell_{\beta,\sigma} & \ell_{\beta,\lambda} & \ell_{\beta,\beta}
\end{bmatrix}$$

where

$$\begin{aligned}
\ell_{\mu,\mu} &= -\frac{n}{\sigma^2} - \beta\sum_{i=1}^{n}\frac{1}{(x_i-\mu)^2} - \frac{\lambda^3}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{\lambda^2}{\sigma^2}\sum_{i=1}^{n}\frac{\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\sigma,\sigma} &= \frac{n(\beta+1)}{\sigma^2} - 3\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^4} + \frac{2\lambda}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
&- \frac{\lambda^3}{\sigma^5}\sum_{i=1}^{n}\frac{(x_i-\mu)^3\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{\lambda^2}{\sigma^4}\sum_{i=1}^{n}\frac{(x_i-\mu)^2\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\lambda,\lambda} &= -\frac{\lambda}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)^3\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\beta,\beta} &= -\frac{n\Psi_1\left(\frac{\beta+1}{2}\right)}{4}
\end{aligned}$$

and

$$\begin{aligned}
\ell_{\mu,\sigma} &= \ell_{\sigma,\mu} = -2\sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^3} + \frac{\lambda}{\sigma^2}\sum_{i=1}^{n}\frac{\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
&- \frac{\lambda^3}{\sigma^4}\sum_{i=1}^{n}\frac{(x_i-\mu)^2\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{\lambda^2}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\mu,\lambda} &= \ell_{\lambda,\mu} = \frac{\lambda^2}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)^2\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{1}{\sigma}\sum_{i=1}^{n}\frac{\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} + \frac{\lambda}{\sigma^2}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\mu,\beta} &= \ell_{\beta,\mu} = -\sum_{i=1}^{n}\frac{1}{(x_i-\mu)}\\
\ell_{\sigma,\lambda} &= \ell_{\lambda,\sigma} = \frac{\lambda^2}{\sigma^4}\sum_{i=1}^{n}\frac{(x_i-\mu)^3\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)} - \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
&+ \frac{\lambda}{\sigma^3}\sum_{i=1}^{n}\frac{(x_i-\mu)^2\phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}{\Phi^2\left(\lambda\frac{x_i-\mu}{\sigma}\right)}\\
\ell_{\sigma,\beta} &= \ell_{\beta,\sigma} = -\frac{n}{\sigma}, \qquad \ell_{\lambda,\beta} = \ell_{\beta,\lambda} = 0.
\end{aligned}$$

REFERENCES

1. Alavi, S. M. R., (2011). On a New Bimodal Normal Family, Journal of Statistical Research of Iran, 8(2), 163-175.
2. Alavi, S. M. R. and Tarhani, M., (2016). On a Skew Bimodal Normal-Normal distribution fitted to the Old-Faithful geyser data, Communications in Statistics - Theory and Methods, 46(15), 7301-7312.
3. Arellano-Valle, R. B., Gomez H. W. and Quintana, F. A., (2004). A New Class of Skew-Normal Distributions, Communications in Statistics - Theory and Methods, 33(7), 1465-1480.
4. Arellano-Valle, R. B., Cortes, M. A. and Gomez, H.W., (2010). An Extension of the Epsilon-Skew Normal Distribution, Communications in Statistics - Theory and Methods, 39(5), 912-922.
5. Arnold, B. C., Gomez, H. W. and Salinas, H. S., (2015). A Doubly Skewed Normal Distribution, Statistics, 49(4), 842-858.
6. Azzalini, A., (1986). Further Results on a Class of Distributions which Includes the Normal Ones. Statistica, 46, 199-208.
7. Azzalini, A., (1985). A Class of Distributions which Includes the Normal Ones. Scandinavian Journal of Statistics, 12, 171-178.
8. Azzalini, A., (2005). The Skew-Normal Distribution and Relative Multivariate Families. Scandinavian Journal of Statistics, 32, 159-188.

9. Azzalini, A., Bowman, A. W., (1990). A Look at Some Data on the Old Faithful Geyser, Journal of Applied Statistics. 39, 357-365.
10. Azzalini, A. and Regoli, G., (2012). Some Properties of Skew-Symmetric Distributions, Annals of the Institute of Statistical Mathematics, 64, 857-879.
11. Famoye, F., Lee, C., and Eugene, N., (2004). Beta-Normal Distribution: Bimodality Properties and Application, Journal of Modern Applied Statistical Methods, 3(1), 85-103.
12. Gomez, H. W., Elal-Olivero, D., Salinas, H. S. and Bolfarine, H., (2011). Bimodal Extension Based on the Skew-Normal Distribution with Application to Pollen Data, Environmetrics, 22(1), 50-62.
13. G¨®mez, H. W., Venegas, O. and Bolfarine, H., (2007). Skew-Symmetric Distributions Generated by the Distribution Function of the Normal Distribution, Environmetrics,18, 395-407.
14. Gradshteyn, I. S. and Ryzhik, I. M., (1965). Tables of Integrals, Series and Products. Academic Press, New York.
15. Gupta, R. C. and Gupta, R. D., (2004). Generalized Skew Normal Model, Test, 13, 501-524.
16. Gupta, A.K., Chang, F. C. and Huang, W.J., (2002). Some Skew-Symmetric Models, Random Operators Stochastic Equations, 10, 133-140.
17. Hassan, M. Y. and Hijazi, R. H., (2010). A Bimodal Exponential Power Distribution. Pakistan Journal of Statistics, 26(2), 379-396.
18. Henze, N., (1986). A Probabilistic Representation of the Skew-Normal Distribution. Scandinavian Journal of Statistics, 13(4), 271-275.
19. Jamalizadeh, A., Behboodian, J. and Balakrishnan, N., (2008). A Two-Parameter Generalized Skew-Normal Distribution, Statistical and Probability Letters, 78, 1722-1728.
20. Karimi, M. and Alavi, S. M. R., (2014). The E?ect of Weight Function on Hypothesis Testing in Weighted Sampling. Journal of Applied Statistics, 41(11), 2493-2503.
21. Kazemi, M. R., Haghbin, H. and Behboodian, J., (2011). Another Generalization of the Skew Normal Distribution, World Applied Sciences Journal, 12, 1034-1039.
22. Kumar, C. S. and Anusree, M.R., (2013). A Generalized Two-Piece Skew Normal Distribution and Some of its Properties, Statistics, 47(6), 1370-1380.
23. Liseo, B., (1990). The Skew-Normal Class of Densities: Aspects of Inference From the Bayesian Point of View. Statistica, 50(1), 71-82.
24. MA, Y. and Genton, M.G., (2004). Flexible Class of Skew-Symmetric Distribution. Scandinavian Journal of Statistics, 31(3), 459-468.
25. Maleki, M. and Nematollahi, A. R., (2016). Bayesian Approach to Epsilon-Skew-Normal Family. Communications in Statistics-Theory and Methods, 46 (15), 7546-7561.
26. Mameli, V. and Musio, M., (2013). A Generalization of the Skew-Normal Distribution: the Beta-Skew-Normal Distribution, Communications in Statistics-Theory and Methods, 42, 2229-2242.
27. Martinez, E. H., Varela, H., Gomez, H. W. and Bolfarine, H., (2008). A Note on the Likelihood and Moments of the Skew-Normal Distribution, Stat Oper Res Trans, 32(1), 57-66.
28. Mukhopadhyay, S. and Vidakovic, B., (1995). Efficiency of Linear Bayes Rules for a Normal Mean: skewed priors class, The Statistician, 44, 389-397.
29. Nadarajah, S. and Kotz, S., (2003). Skewed distributions generated by the normal kernel. Statistics and Probability Letters, 65, 269-277.
30. Nekoukhou, V. and Alamatsaz, M. H., (2012). A family of skew-symmetric-Laplace distributions. Statistical papers, 53, 685-696.
31. O¡¯Hagan, A., and Leonard, T., (1976). Bayes Estimation Subject to Uncertainty About Parameter Constraints. Biometrika, 63, 201-203.
32. Pewsey, A., (2000). Problems of Inference for Azzalini¡¯s Skew-Normal Distribution. Journal of Applied Statistics, 27(7), 859-870.
33. Rao, C. R., (1965). On Discrete Distributions Arising out of Methods of Ascertainment, Sankhya, 27, 311-324.
34. Rasekhi, M., Chinipardaz, R. and Alavi, S. M. R., (2015). A Flexible Generalization of the Skew Normal Distribution Based on a Weighted Normal Distribution. Statistical Methods and Application, 25 (3), 375-394.
35. Sewell, M. and Young, C., (1997). Are Echinoderm Egg Size Distributions Bimodal. Biological Bulletin, 193, 297-205.
36. Sharafi, M. and Behboodian, J., (2007). The Balakrishnan Skew-Normal Density. Statistical Papers, 49, 769-778.
37. Wang, J., Boyer, J. and Genton, M. G., (2004). A Skew-Symmetric Representation of Multivariate Distributions, Statistica Sinica, 14, 1259-1270.
38. Yadegari, I., Gerami, A. and Khaledi, M. J., (2008). A Generalized of the Balakrishnan Skew-Normal Distribution. Statistics and Probablity Letters, 78, 1165-1167.