Communications in Statistics - Theory and Methods

A Generalization of the Skew-Normal Distribution: The Beta Skew-Normal

Valentina Mameli & Monica Musio

Dipartimento di Matematica, Università di Cagliari, Cagliari, Italy
Published online: 29 Apr 2013.

To cite this article: Valentina Mameli & Monica Musio (2013): A Generalization of the Skew-Normal Distribution: The Beta Skew-Normal, Communications in Statistics - Theory and Methods, 42:12, 2229-2244

To link to this article: http://dx.doi.org/10.1080/03610926.2011.607530

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
A Generalization of the Skew-Normal Distribution: 
The Beta Skew-Normal

VALENTINA MAMELI AND MONICA MUSIO
Dipartimento di Matematica, Università di Cagliari, Cagliari, Italy

We consider a new generalization of the skew-normal distribution introduced by Azzalini (1985). We denote this distribution Beta skew-normal (BSN) since it is a special case of the Beta generated distribution (Jones, 2004). Some properties of the BSN are studied. We pay attention to some generalizations of the skew-normal distribution (Bahrami et al., 2009; Sharafi and Behboodian, 2008; Yadegari et al., 2008) and to their relations with the BSN.

Keywords Balakrishnan skew-normal; Beta skew-normal; Order statistics; Skew-normal distribution.

Mathematics Subject Classification 60E05; 62E15.

1. Introduction
The skew-normal distribution (SN), introduced by Azzalini (1985), has been studied and generalized extensively. The aim of this article is to introduce a new family of distributions, which generalizes the skew-normal, that is flexible enough to support both unimodal and bimodal shape. This new family, called Beta skew-normal (BSN), arises naturally when we consider the distributions of order statistics of the skew-normal distribution. The BSN can also be obtained as a special case of the Beta generated distribution (Jones, 2004). In this article, we pay attention to three other generalizations of the skew-normal distribution: the Balakrishnan skew-normal (SNB) (Balakrishnan, 2002; [as a discussant of Arnold and Beaver, 2002]; Gupta and Gupta, 2004; Sharafi and Behboodian, 2008), the generalized Balakrishnan skew-normal (GBSN) (Yadegari et al., 2008), and a two-parameter generalization of the Balakrishnan skew-normal (TBSN) (Bahrami et al., 2009). The above three extensions are related to the Beta skew-normal distribution for particular values of the parameters.

Given a random sample \(X_1, \ldots, X_n\) from a distribution \(F(x)\), in general the distribution of the related order statistics does not belong to the family of \(F(x)\). In this article, we show that the maximum between the \(X_i\)’s from a Balakrishnan
skew-normal with parameters $m$ and 1, denoted by $X_i \sim SNB_m(1)$, is still a Balakrishnan skew-normal with parameters $k$ and 1, where $k$ is a function of $m$ and $n$.

This article is organized as follows. After describing briefly, in Sec. 2, the skew-normal distribution, its generalizations, and listing their most important properties, in Sec. 3 we present some generalizations of the Beta distribution. In the last section, we define the Beta skew-normal distribution and present its properties and some special cases. In particular, the BSN contains the Beta half-normal distribution (Pescim et al., 2010) as limiting case. Besides, we investigate its shape properties. We derive its moment-generating function and we also compute numerically the first moment, variance, skewness, and kurtosis. We present two different methods which allow to simulate a BSN distribution. We explore its relationships with the other generalizations of the skew-normal and we show that the distributions of order statistics from the skew-normal distribution are Beta skew-normal and are log-concave. Furthermore, in this section we give some results concerning the SNB distribution. In particular, we derive the exact distributions of the largest order statistic from $SNB_m(1)$ and the shortest order statistic from $SNB_m(-1)$.

2. The Skew-Normal Density and its Generalizations

The present section recalls some important definitions and properties about the skew-normal distribution and some of its extensions.

2.1. The Skew-Normal Density

The skew-normal distribution refers to a parametric class of probability distributions which includes the standard normal as a special case. A random variable $Z$ is said to be skew-normal with parameter $\lambda$, if its density function is

$$
\phi(z; \lambda) = 2\phi(z)\Phi(\lambda z), \quad \text{with } z, \lambda \in \mathbb{R},
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution, respectively. We denote a random variable $Z$ with the above density by $Z \sim SN(\lambda)$. The parameter $\lambda$ controls skewness. The standard normal distribution is a skew-normal distribution with $\lambda = 0$. We remind some properties of the SN distribution.

**Properties of $SN(\lambda)$**

a. As $\lambda \to \infty$, $\phi(z; \lambda)$ tends to the half-normal density.

b. If $Z$ is a $SN(\lambda)$ random variable, then $-Z$ is a $SN(-\lambda)$ random variable.

c. If $Z \sim SN(\lambda)$, then $Z^2 \sim \chi^2(1)$.

d. The density (1) is strongly unimodal, i.e., $\log \phi(z; \lambda)$ is a concave function of $z$.

The corresponding distribution function is

$$
\Phi(z; \lambda) = 2 \int_{-\infty}^{z} \int_{-\infty}^{\lambda t} \phi(t)\phi(u) du dt = \Phi(z) - 2T(z, \lambda),
$$

where $T(z, \lambda)$ is Owen’s function. The properties of this function are:

1. $-T(z, \lambda) = T(z, -\lambda)$;
Using the properties of Owen’s function, we have immediately the following ones.

Property 1. \( 1 - \Phi(-z; \lambda) = \Phi(z; -\lambda) \).

Property 2. \( \Phi(z; 1) = \Phi(z)^2 \).

Property 3. \( \Phi(z; \lambda) + \Phi(z; -\lambda) = 2\Phi(z) \).

The class of skew-normal distributions can be generalized by the inclusion of the location and scale parameters which we identify as \( \xi \) and \( \psi > 0 \). Thus, if \( X \sim SN(\lambda) \) then \( Y = \xi + \psi X \) is a skew-normal variable with parameters \( \xi, \psi, \lambda \). We denote \( Y \) by \( Y \sim SN(\xi, \psi, \lambda) \).

2.2. The Balakrishnan Skew-Normal Density and its Generalization

Balakrishnan (2002) proposed a generalization of the standard skew-normal distribution as follows.

Definition 2.1. A random variable \( X \) has Balakrishnan skew-normal distribution, denoted by \( SNB_n(\lambda) \), if it has the following density function, with \( n \in \mathbb{N} \):

\[
f_n(x; \lambda) = c_n(\lambda)\phi(x)\Phi(\lambda x)^n, \quad x \in \mathbb{R}, \lambda \in \mathbb{R}.
\]  

(2)

The coefficient \( c_n(\lambda) \), which is a function of \( n \) and the parameter \( \lambda \), is given by

\[
c_n(\lambda) = \frac{1}{\int_{-\infty}^{\infty} \phi(x)\Phi(\lambda x)^n dx} = \frac{1}{E(\Phi(\lambda U)^n)},
\]

where \( U \sim N(0, 1) \).

For \( n = 0 \) and \( n = 1 \), the above density reduces to the standard normal and the skew-normal distribution, respectively.

For \( n = 2 \) a random variable \( X \) with the density (2) is denoted by \( X \sim NSN(\lambda) \) with \( c_2(\lambda) = \frac{\pi}{\arctan \sqrt{1+2\lambda^2}} \) (Sharafi and Behboodian, 2006).

The class of Balakrishnan skew-normal can be generalized by the inclusion of the location and scale parameters which we identify as \( \mu \) and \( \sigma > 0 \). Thus, if \( X \sim SNB_n(\lambda) \) then \( Y = \mu + \sigma X \) is a Balakrishnan skew-normal variable with parameters \( \mu, \sigma, \lambda \). We denote \( Y \) by \( Y \sim SNB_n(\mu, \sigma, \lambda) \).

Remark 2.1. Sharafi and Behboodian (2008) showed that for \( \lambda = 1 \), (2) is the density function of the \((n + 1)\)th order statistic \( X_{n+1} \) in a sample of size \( n + 1 \) from \( N(0, 1) \). Moreover, for \( \lambda = -1 \), (2) is the density function of the first-order statistic \( X_1 \) in a sample of size \( n + 1 \) from \( N(0, 1) \).

Recently, Yadegari et al. (2008) introduced the following generalization of the Balakrishnan skew-normal distribution and explained some important properties of this distribution.
Definition 2.2. A random variable $X$ is said to have a generalized Balakrishnan skew-normal distribution, denoted by $GBSN_{n,m}(\lambda)$, with parameters $n, m \in \mathbb{N}$ and $\lambda \in \mathbb{R}$, if its density function has the following form:

$$f_{n,m}(x; \lambda) = \frac{1}{C_{n,m}(\lambda)} \phi(x) \Phi(\lambda x)^n (1 - \Phi(\lambda x))^m, \quad x \in \mathbb{R},$$

where $C_{n,m}(\lambda) = \sum_{i=0}^{m} \binom{m}{i} (-1)^i \int_{-\infty}^{\infty} \phi(x) \Phi(\lambda x)^{n+i} dx$.

For $m = 0$ this density reduces to the Balakrishnan skew-normal.

Remark 2.2. Let $X_1, \ldots, X_n$ be a random sample from a $N(0, 1)$. Then the $j$th order statistic is a $GBSN_{j-1,n-j}(1)$, with $j = 1, \ldots, n$. In this case we have that

$$C_{j-1,n-j}(1) = \sum_{i=0}^{n-j} \binom{n-j}{i} (-1)^i \int_{-\infty}^{\infty} \phi(x) \Phi(x)^j \Phi(x)^{n-j} dx = \frac{n!}{(j-1)!(n-j)!}.$$ (4)

Bahrami et al. (2009) discussed a two-parameter generalized skew-normal distribution which includes the skew-normal, the Balakrishnan skew-normal and the generalized Balakrishnan skew-normal as special cases.

Definition 2.3. A random variable $Z$ has a two-parameter Balakrishnan skew-normal distribution with parameters $\lambda_1, \lambda_2 \in \mathbb{R}$, denoted by $TBSN_{n,m}(\lambda_1, \lambda_2)$, if its probability density function (pdf) is

$$f_{n,m}(z; \lambda_1, \lambda_2) = \frac{1}{c_{n,m}(\lambda_1, \lambda_2)} \phi(z) \Phi(\lambda_1 z)^n \Phi(\lambda_2 z)^m, \quad z \in \mathbb{R},$$

and $n, m$ are non-negative integer numbers. The coefficient $c_{n,m}(\lambda_1, \lambda_2)$, which is a function of the parameters $n, m, \lambda_1$, and $\lambda_2$, is given by

$$c_{n,m}(\lambda_1, \lambda_2) = E \left[ \Phi(\lambda_1 X)^n \Phi(\lambda_2 X)^m \right],$$

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

The following properties are direct consequences of definition (2.3).

Properties of $TBSN_{n,m}(\lambda_1, \lambda_2)$:
1. $TBSN_{1,1}(\lambda_1, 0) = SN(\lambda_1)$ and $TBSN_{1,1}(0, \lambda_2) = SN(\lambda_2)$;
2. $TBSN_{n,m}(\lambda, \lambda) = SNB_{n+m}(\lambda)$;
3. $TBSN_{n,m}(\lambda_1, 0) = SNB_{n}(\lambda_1)$ and $TBSN_{n,m}(0, \lambda_2) = SNB_{m}(\lambda_2)$;
4. $TBSN_{n,m}(\lambda_1, -\lambda_2) = GBSN_{n,m}(\lambda_1)$ and $TBSN_{n,m}(-\lambda_1, \lambda_2) = GBSN_{n,m}(\lambda_2)$;
5. $TBSN_{n,m}(0, 0) = TBSN_{0,0}(\lambda_1, \lambda_2) = N(0, 1)$.

Remark 2.3. Let $Z_1, \ldots, Z_n$ i.i.d. $N(0, 1)$ and $Z_{(1)} \leq Z_{(2)} \leq \cdots \leq Z_{(n)}$ be the corresponding order statistics, then $Z_{(j)} \sim TBSN_{j-1,n-j}(1, -1)$.

The location-scale two-parameter Balakrishnan skew-normal distribution is defined as the distribution of $Y = \mu + \sigma X$, where $X \sim TBSN_{n,m}(\lambda_1, \lambda_2)$. Hence, $\mu \in \mathbb{R}$.
and \( \sigma > 0 \), are the location and scale parameters, respectively. We denote \( Y \) by \( Y \sim TBSN_{n,m}(\mu, \sigma, \lambda_1, \lambda_2) \).

In the rest of the article, we denote by \( \phi(\cdot; \lambda) \) the density function of \( SN(\lambda) \) and by \( f_{n,m}(\cdot; \lambda_1, \lambda_2) \) the density function of \( TBSN_{n,m}(\lambda_1, \lambda_2) \).

3. Some Extensions of the Beta Distribution

In this section we introduce two families of distributions which generalize the Beta one.

3.1. The Generalized Beta of the First Type

We recall the following definition due to McDonald (1984).

**Definition 3.1.** A variable \( X \) is said to have a generalized Beta distribution of the first kind with positive parameters \( a, b, p, \) and \( q \) if its density is given by

\[
g(x; a, b, p, q) = \frac{px^{ap-1} \left(1 - \left(\frac{x}{q}\right)^p\right)^{b-1} q^{ap} B(a, b)}{\left[1 - (\frac{x}{q})^p\right]^{a-1}}, \quad \text{with } 0 \leq x \leq q. \tag{6}
\]

If \( p = 1 \) and \( q = 1 \) the variable \( X \) is a Beta of the first kind with parameters \( a \) and \( b \).

For \( q = 1 \) and \( a = 1 \), the variable \( X \) is said to have a *Kumaraswamy distribution* with parameters \( p \) and \( b \) (Kumaraswamy, 1980).

3.2. The Beta Generated Distribution

Starting from the Beta distribution, Jones (2004) defined a new family of probability distributions, called Beta generated distribution. Following the notation of Jones, the class of Beta generated distributions is defined as follows.

**Definition 3.2.** Let \( F(\cdot) \) be a continuous distribution function with density function \( f(\cdot) \). The univariate family of distributions generated by \( F(\cdot) \), called Beta generated distribution, with parameters \( a, b > 0 \), has pdf

\[
g^B_F(x; a, b) = \frac{1}{B(a, b)} (F(x))^{a-1} (1 - F(x))^{b-1} f(x), \tag{7}
\]

where \( B(a, b) \) is the complete Beta function.

Thus, this family of distributions has distribution function given by:

\[
G^B_F(x; a, b) = I_{F(x)}(a, b), \quad a, b > 0, \tag{8}
\]

where the function \( I_{F(x)} \) denotes the incomplete Beta ratio defined by

\[
I_x(a, b) = \frac{B_x(a, b)}{B(a, b)}, \tag{9}
\]

**Downloaded by [Universita Degli Studi di Cagliari] at 02:02 02 May 2013**
with
\[ B_y(a, b) = \int_0^y z^{a-1} (1 - z)^{b-1} \, dz, \quad 0 < y \leq 1, \quad (10) \]
the incomplete Beta function. Replacing (9) and (10) in (8), we get that this family of distributions has distribution function
\[ G^B_F(x; a, b) = \frac{1}{B(a, b)} \int_0^{F(x)} z^{a-1} (1 - z)^{b-1} \, dz. \quad (11) \]

**Remark 3.1.** Let \( f(\cdot) \) unimodal and continuously differentiable, if \( a, b \geq 1 \) then \( g^B_F(\cdot, a, b) \) is also unimodal. The strong unimodality, i.e., log-concavity, of \( f(\cdot) \) implies strong unimodality of \( g^B_F(\cdot, a, b) \).

Eugene et al. (2002) studied in detail the family of Beta-normal distribution (BN) and discussed its properties. Recently, Pescim et al. (2010) proposed the Beta half-normal (BHN) distribution to extend the half-normal (HN) distribution. We now recall the definitions of Beta-normal distribution and Beta half-normal distribution.

### 3.2.1. Beta-normal distribution.

When in (7) \( F(x) \) is the normal distribution function with parameters \( \mu \) and \( \sigma \) we have the Beta-normal family with distribution function
\[ G^B_{\Phi(z)}(x; a, b) = \frac{1}{B(a, b)} \int_0^{\Phi(z)} z^{a-1} (1 - z)^{b-1} \, dz, \quad (12) \]
and corresponding pdf
\[ g^B_{\Phi(z)}(x; a, b) = \frac{1}{B(a, b)} \left( \Phi\left( \frac{x - \mu}{\sigma} \right) \right)^{a-1} \left( 1 - \Phi\left( \frac{x - \mu}{\sigma} \right) \right)^{b-1} \sigma^{-1} \phi\left( \frac{x - \mu}{\sigma} \right), \quad (13) \]
where \( \sigma^{-1} \phi\left( \frac{x - \mu}{\sigma} \right) \) and \( \Phi\left( \frac{x - \mu}{\sigma} \right) \) are the normal density and distribution with parameters \( \mu \) and \( \sigma \).

A random variable \( X \) with Beta-normal distribution with vector of parameters \( \xi = (0, 1, a, b) \) is denoted by \( X \sim \text{BN}(a, b) \).

### 3.2.2. Beta half-normal distribution.

Let \( F(x) = 2\Phi(x) - 1, \) with \( x > 0 \), the distribution function of the half-normal distribution. By using \( F(x) \) in (7), the density function of the Beta half-normal distribution (BHN) is given by
\[ g^B_{2\Phi(x)-1}(x; a, b) = \frac{2^b}{B(a, b)} (2\Phi(x) - 1)^{a-1} (1 - \Phi(x))^{b-1} \phi(x), \quad x > 0, \quad (14) \]
and the relative distribution function is
\[ G^B_{2\Phi(x)-1}(x; a, b) = \frac{2^b}{B(a, b)} \int_0^x (2\Phi(z) - 1)^{a-1} (1 - \Phi(z))^{b-1} \phi(z) \, dz, \quad x > 0. \quad (15) \]
When $X$ is a random variable following the BHN distribution, it is denoted by $X \sim BHN(a, b)$.

4. A New Generalization of the Skew-Normal Distribution:
   The Beta Skew-Normal

In this section we define the Beta skew-normal class and we present some of its properties. This new distribution, which contains as special case the skew-normal one, is flexible enough to support unimodal and bimodal shape. Replacing in (7) $F(x)$ by $\Phi(x; \lambda)$, we obtain the Beta skew-normal distribution, with distribution function given by

$$G^B_{\Phi(x; \lambda)}(x; \lambda, a, b) = \frac{1}{B(a, b)} \int_0^x \Phi(x; \lambda) z^{a-1} (1-z)^{b-1} dz,$$  \hspace{1cm} (16)

and pdf

$$g^B_{\Phi(x; \lambda)}(x; \lambda, a, b) = \frac{2}{B(a, b)} (\Phi(x; \lambda))^{a-1} (1-\Phi(x; \lambda))^{b-1} \phi(x) \Phi(\lambda x).$$  \hspace{1cm} (17)

Throughout this article, we denote the Beta skew-normal distribution with vector of parameters $\xi = (\lambda, a, b)$ by $BSN(\lambda, a, b)$. The class of the Beta skew-normal can be generalized by the inclusion of the location and scale parameters which we identify as $\mu$ and $\sigma > 0$. Thus, if $X \sim BSN(\lambda, a, b)$ then $Y = \mu + \sigma X$ is a Beta skew-normal with vector of parameters $\xi = (\mu, \sigma, \lambda, a, b)$. We denote $Y$ by $Y \sim BSN(\mu, \sigma, \lambda, a, b)$.

We now present some properties concerning the $BSN(\lambda, a, b)$.

Properties of $BSN(\lambda, a, b)$.

a. $g^B_{\Phi(x; \lambda)}(x; \lambda, 1, 1) = \phi(x; \lambda)$ for all $x \in \mathbb{R}$, i.e., $BSN(\lambda, 1, 1) = SN(\lambda)$.

b. $g^B_{\Phi(x; 0)}(x; 0, a, b) = g^B_{\Phi(x; 0)}(x; a, b)$ for all $x \in \mathbb{R}$, i.e., $BSN(0, a, b) = BN(a, b)$.

c. $g^B_{\Phi(x; 0)}(x; 0, 1, 1) = \phi(x)$ for all $x \in \mathbb{R}$, i.e., $BSN(0, 1, 1) = N(0, 1)$.

d. $g^B_{\Phi(x; 1)}(x; 1, \frac{1}{2}, 1) = \phi(x)$ for all $x \in \mathbb{R}$, i.e., $BSN(1, \frac{1}{2}, 1) = N(0, 1)$.

e. $g^B_{\Phi(x; -1)}(x; -1, \frac{1}{2}) = \phi(x)$ for all $x \in \mathbb{R}$, i.e., $BSN(-1, \frac{1}{2}) = N(0, 1)$.

f. If $X \sim BSN(\lambda, a, b)$ then $-X \sim BSN(-\lambda, b, a)$.

g. If $X \sim BSN(\lambda, a, b)$ then $Y = \Phi(X; \lambda)$ is a Beta$(a, b)$.

h. If $X \sim BSN(\lambda, a, b)$ then $Y = 1 - \Phi(X; \lambda)$ is a Beta$(b, a)$.

i. As $\lambda \to +\infty$, $g^B_{\Phi(x; \lambda)}(x; \lambda, a, b)$ tends to the Beta half-normal density.

Remark 4.1. Properties from a to e establish that the family of $BSN(\lambda, a, b)$ contains the standard normal distribution, skew-normal distribution, and Beta-normal distribution as special cases.

Proof. The proof follows directly from (17) and from elementary properties of the skew-normal distribution.

The BSN distribution is easily simulated using Property g as follows: if $Y$ has a Beta distribution with parameters $a$ and $b$, then the variable $X = \Phi^{-1}(Y; \lambda)$
has $BSN(\lambda, a, b)$ distribution, where $\Phi^{-1}(\cdot; \lambda)$ is the quantile function of the skew-normal distribution. In Fig. 1 are plotted random samples generated by the BSN distribution for some $a$, $b$, $\lambda$ with the respective curve of the density function obtained using the R-package “sn” (see Azzalini, 2010).

From this plot we can observe that, for values of $a$ and $b$ close to zero, the distribution can be bimodal. The behavior of $BSN$ is highlighted in Figs. 2 and 3. From Remark 3.1 we know that, if $a, b \geq 1$, the density (17) is strongly unimodal, i.e., $\log g_{\phi(x; \lambda)}(x; \lambda, a, b)$ is a concave function of $x$ (see Fig. 2). We don’t have general results for $a$ and/or $b < 1$. A numerical study has shown that, when the two parameters $a$ and $b$ are closed to zero $(0,10,0,20)$, the density can be bimodal (see Fig. 3). Numerically, Eugene et al. (2002) observed that the BN is bimodal when both parameters $a$ and $b$ are less than 0.214.

By Property f we can deduce the following Proposition.

**Proposition 4.1.** Let $X \sim BSN(\lambda, a, b)$ and $Y \sim BSN(-\lambda, b, a)$. We have the following statements:
1. $E_X(X) = -E_Y(Y)$;
2. $\text{var}_X(X) = \text{var}_Y(Y)$;
3. $\gamma_1(X) = -\gamma_1(Y)$;
4. $\gamma_2(X) = \gamma_2(Y)$;

with $\gamma_1$ and $\gamma_2$ we indicate the skewness and the kurtosis, respectively.

Now we find the moment generating function of $X$ which has density (17).

Figure 2. The Beta skew-normal $BSN(\lambda, a, b)$ for values of $a, b \geq 1$.

Figure 3. The Beta skew-normal $BSN(\lambda, a, b)$ for values of $a, b < 1$. 
Property 4.1. The moment generating function of $X \sim BSN(\lambda, a, b)$ is given by

$$M_X(t) = \frac{2}{B(a, b)} e^{\frac{\lambda}{2} t^2} E_Z \left[ (\Phi(Z; \lambda))^{a-1} (1 - \Phi(Z; \lambda))^{b-1} \Phi(\lambda Z) \right], \quad (18)$$

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

We have the following recursion formula:

Property 4.2. Let $k \in \mathbb{N}$ and $k \geq 1$. If $X \sim BSN(\lambda, a, b)$, with $a, b > 1$ then

$$E_X(X^k) = (k - 1) E_X(X^{k-2}) + \lambda E_X \left[ X^{k-1} \frac{\phi(\lambda X)}{\Phi(\lambda X)} \right]$$
$$+ (a + b - 1) E_V \left[ U^{k-1} \phi(U; \lambda) \right] - (a + b - 1) E_V \left[ V^{k-1} \phi(V; \lambda) \right],$$

where $U \sim BSN(\lambda, a - 1, b)$ and $V \sim BSN(\lambda, a, b - 1)$ are independent random variables.

Proof. The proof follows easily from application of the formula for integration by parts and by using the well note $\frac{\partial \Phi(x)}{\partial x} = -x \phi(x)$ (see Arnold et al., 1992).

The Beta skew-normal density is in general asymmetric (see Figs. 2 and 3). We have a partial result concerning symmetry.

Proposition 4.2. If $a = b$ and $BSN(\lambda, a, b)$ is symmetric about 0 then $\lambda = 0$.

Proof. We consider the density of a random variable $X \sim BSN(\lambda, a, a)$:

$$g_{\Phi(x; \lambda)}(-x; \lambda, a, a) = \frac{2}{B(a, a)} \phi(x) \Phi(-\lambda x)(1 - \Phi(x; -\lambda))^{a-1} (\Phi(x; -\lambda))^{a-1}$$

this is equal to $g_{\Phi(x; \lambda)}(x; \lambda, a, a)$ if $\Phi(\lambda x) = \Phi(-\lambda x)$ and $\Phi(x; \lambda) = \Phi(x; -\lambda)$. However, for Property 3 we find that $\Phi(x; \lambda) = \Phi(x)$ which implies that $\lambda = 0$.

Remark 4.2. Eugene et al. (2002) showed that the $BN(a, b) = BSN(0, a, b)$ is symmetric about 0 when $a = b$.

Moments of the BSN cannot be evaluated exactly. We have computed them numerically using the software R. In Table 1 we reported the values of the mean $\mu_{BSN}$, standard deviation $\sigma_{BSN}$, skewness $\gamma_1$, and kurtosis $\gamma_2$ for different values of the parameters $a, b, \lambda$. From this numerical study we have noted that:

- for fixed values of $a$ and $b$ the mean $\mu_{BSN}$ and skewness $\gamma_1$ are increasing function of $\lambda$;
- for fixed values of $b$ and $\lambda$ the mean $\mu_{BSN}$ and skewness $\gamma_1$ are increasing function of $a$;
- for fixed values of $a$ and $\lambda$ the mean $\mu_{BSN}$ is a decreasing function of $b$.

The distribution of order statistics of i.i.d. random variables is important in many areas of inferential statistics. We will see now that the BSN distribution, with $a$ and $b$ integers, is the distribution of order statistics from a skew-normal distribution.
Table 1
The first moment, standard deviation, skewness, and kurtosis of $BSN(\lambda, a, b)$ for different values of $a$, $b$, and $\lambda$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>$\mu_{BSN}$</th>
<th>$\sigma_{BSN}$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\lambda$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>$-10$</td>
<td>$-1.1579$</td>
<td>$1.4029$</td>
<td>$-1.1329$</td>
<td>$3.7648$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>$-1$</td>
<td>$-0.6501$</td>
<td>$1.9679$</td>
<td>$-0.2378$</td>
<td>$2.7777$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
<td>$1.1580$</td>
<td>$1.4027$</td>
<td>$1.1329$</td>
<td>$3.7632$</td>
</tr>
<tr>
<td>0.25</td>
<td>0.5</td>
<td>$-10$</td>
<td>$-1.5906$</td>
<td>$1.3469$</td>
<td>$-0.7185$</td>
<td>$2.7580$</td>
</tr>
<tr>
<td>0</td>
<td>0.9631</td>
<td>1</td>
<td>$-0.9631$</td>
<td>$1.9061$</td>
<td>$-0.0849$</td>
<td>$2.8029$</td>
</tr>
<tr>
<td>1</td>
<td>0.1772</td>
<td>10</td>
<td>$1.6339$</td>
<td>$1.3974$</td>
<td>$0.8434$</td>
<td>$3.2655$</td>
</tr>
<tr>
<td>0.5</td>
<td>0.25</td>
<td>$-10$</td>
<td>$-0.5447$</td>
<td>$0.8727$</td>
<td>$-1.5061$</td>
<td>$5.2003$</td>
</tr>
<tr>
<td>0</td>
<td>0.9625</td>
<td>1</td>
<td>$-0.9625$</td>
<td>$1.9051$</td>
<td>$0.0819$</td>
<td>$2.7927$</td>
</tr>
<tr>
<td>1</td>
<td>1.4411</td>
<td>10</td>
<td>$1.6339$</td>
<td>$1.3974$</td>
<td>$0.8434$</td>
<td>$3.2655$</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>$-10$</td>
<td>$-0.8979$</td>
<td>$0.8874$</td>
<td>$-0.9703$</td>
<td>$3.3176$</td>
</tr>
<tr>
<td>0</td>
<td>0.5882</td>
<td>1</td>
<td>$-0.5882$</td>
<td>$1.2639$</td>
<td>$-0.1811$</td>
<td>$2.9514$</td>
</tr>
<tr>
<td>1</td>
<td>0.5882</td>
<td>10</td>
<td>$1.5253$</td>
<td>$1.2639$</td>
<td>$0.1811$</td>
<td>$2.9514$</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>$-10$</td>
<td>$-1.3018$</td>
<td>$0.9148$</td>
<td>$-0.8262$</td>
<td>$3.4815$</td>
</tr>
<tr>
<td>0</td>
<td>1.1664</td>
<td>1</td>
<td>$-1.1664$</td>
<td>$1.0704$</td>
<td>$-0.3850$</td>
<td>$3.1159$</td>
</tr>
<tr>
<td>1</td>
<td>0.7043</td>
<td>10</td>
<td>$0.9999$</td>
<td>$1.2479$</td>
<td>$-0.1372$</td>
<td>$2.9831$</td>
</tr>
<tr>
<td>10</td>
<td>0.4873</td>
<td>1</td>
<td>$0.9999$</td>
<td>$0.9153$</td>
<td>$1.0703$</td>
<td>$3.7747$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>$\mu_{BSN}$</th>
<th>$\sigma_{BSN}$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\lambda$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>$-10$</td>
<td>$-2.3678$</td>
<td>$0.7314$</td>
<td>$-0.7505$</td>
<td>$3.7967$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>$-1$</td>
<td>$-2.3617$</td>
<td>$0.7389$</td>
<td>$-0.7188$</td>
<td>$3.7849$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>0</td>
<td>$-2.0809$</td>
<td>$0.8033$</td>
<td>$-0.6173$</td>
<td>$3.5736$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>1</td>
<td>$-1.0893$</td>
<td>$0.6117$</td>
<td>$-0.5642$</td>
<td>$3.4799$</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>$-10$</td>
<td>$-0.0182$</td>
<td>$0.1429$</td>
<td>$0.3706$</td>
<td>$3.8635$</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>$-10$</td>
<td>$-0.4873$</td>
<td>$0.5777$</td>
<td>$-1.3200$</td>
<td>$4.8570$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>$-1$</td>
<td>$-0.7043$</td>
<td>$1.2479$</td>
<td>$0.1372$</td>
<td>$2.9831$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$1.1664$</td>
<td>$1.0704$</td>
<td>$0.3086$</td>
<td>$3.1161$</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>$-10$</td>
<td>$-0.7939$</td>
<td>$0.6080$</td>
<td>$-0.9556$</td>
<td>$3.8232$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$-10$</td>
<td>$-0.5642$</td>
<td>$0.8256$</td>
<td>$-0.1369$</td>
<td>$3.0617$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>$-1$</td>
<td>$-1.5388$</td>
<td>$0.5868$</td>
<td>$0.4099$</td>
<td>$3.3314$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$1.8675$</td>
<td>$0.5251$</td>
<td>$0.5005$</td>
<td>$3.4685$</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>$-10$</td>
<td>$-1.8807$</td>
<td>$0.5124$</td>
<td>$0.5744$</td>
<td>$3.5243$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>$-10$</td>
<td>$-1.8807$</td>
<td>$0.5124$</td>
<td>$-0.5744$</td>
<td>$3.5243$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$-1$</td>
<td>$-1.8675$</td>
<td>$0.5251$</td>
<td>$-0.5005$</td>
<td>$3.4685$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>$-1.5388$</td>
<td>$0.5868$</td>
<td>$-0.4099$</td>
<td>$3.3314$</td>
</tr>
<tr>
<td>10</td>
<td>0.0839</td>
<td>1</td>
<td>$0.8256$</td>
<td>$0.1369$</td>
<td>$3.0617$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>$\mu_{BSN}$</th>
<th>$\sigma_{BSN}$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\lambda$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>$-10$</td>
<td>$-1.3018$</td>
<td>$0.9148$</td>
<td>$-0.8262$</td>
<td>$3.4815$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$-1.1664$</td>
<td>$1.0704$</td>
<td>$-0.3850$</td>
<td>$3.1159$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>$-1$</td>
<td>$-1.5388$</td>
<td>$0.5868$</td>
<td>$-0.4099$</td>
<td>$3.3314$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$1.8675$</td>
<td>$0.5251$</td>
<td>$0.5005$</td>
<td>$3.4685$</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>$-10$</td>
<td>$-1.8807$</td>
<td>$0.5124$</td>
<td>$-0.5744$</td>
<td>$3.5243$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>$-10$</td>
<td>$-1.8807$</td>
<td>$0.5124$</td>
<td>$-0.5005$</td>
<td>$3.4685$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$-1$</td>
<td>$-1.5388$</td>
<td>$0.5868$</td>
<td>$-0.4099$</td>
<td>$3.3314$</td>
</tr>
<tr>
<td>10</td>
<td>0.0839</td>
<td>1</td>
<td>$0.8256$</td>
<td>$0.1369$</td>
<td>$3.0617$</td>
<td></td>
</tr>
</tbody>
</table>
Proposition 4.3. Let $X_1, \ldots, X_n$ be a random sample from a $SN(\lambda)$. Then the $j$th order statistic is a $BSN(\lambda, j, n - j + 1)$, where $j = 1, \ldots, n$.

Proof. The proof readily follows using the standard formula of the density of $X_{(j)}$ (see, for example, David and Nagaraya, 2003), the $i$th order statistic of a random sample of size $n$ from the distribution $SN(\lambda)$.

In particular, we have the following Corollaries.

Corollary 4.1. Let $X_1, \ldots, X_n$ be a random sample from a $SN(1)$. Then,

$$X_{(n)} = \max \{X_1, \ldots, X_n\}$$

is a $BSN(1, n, 1)$.

Corollary 4.2. Let $X_1, \ldots, X_n$ be a random sample from a $SN(-1)$. Then,

$$X_{(1)} = \min \{X_1, \ldots, X_n\}$$

is a $BSN(-1, 1, n)$.

Corollary 4.3. Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics from a sample of size $n$ from a $SN(\lambda)$ distribution. Then $X_{(i)}$, $i = 1, \ldots, n$, has log-concave density.

Proof. From Property d of Sec. 2 we know that $X_i$ has a log-concave density. We conclude the proof using the following result due to Gupta and Gupta (2004): Suppose $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics from a sample of size $n$ from a distribution having a log-concave density function. Then $X_{(i)}$, $i = 1, \ldots, n$, has log-concave density.

We now derive other properties of the BSN distribution.

Theorem 4.1. Let $X \sim BSN(\lambda, a, b)$ be independent of a random sample $(Y_1, \ldots, Y_n)$ from $SN(\lambda)$, then $X \mid (Y_{(n)} \leq X) \sim BSN(\lambda, a + n, b)$ and $X \mid (Y_{(1)} \geq X) \sim BSN(\lambda, a, b + n)$, where $Y_{(n)}$ and $Y_{(1)}$ are the largest and the smallest order statistics, respectively.

Proof. If $W = X \mid (Y_{(n)} \leq X)$, then we have

$$P(W \leq w) = \frac{\int_{-\infty}^{w} (\Phi(x; \lambda))^a \Phi(x; \lambda) (\Phi(x; \lambda))^{(a-1)} (1 - \Phi(x; \lambda))^{(b-1)} dx}{P(Y_{(n)} \leq X)}.$$  \hfill (19)

Also,

$$P(Y_{(n)} \leq X) = P(Y_1 \leq X, \ldots, Y_n \leq X)$$

$$= \int_{-\infty}^{\infty} (\Phi(x; \lambda))^a \frac{2}{B(a, b)} \Phi(x; \lambda) (\Phi(x; \lambda))^{(a-1)} (1 - \Phi(x; \lambda))^{(b-1)} dx$$

$$= \frac{B(a + n, b)}{B(a, b)}.$$
Taking derivative from (19) with respect to \( w \), we obtain the \( BSN(\lambda, a + n, b) \) density function, the proof of \( X \mid (Y_{(1)} \geq X) \sim BSN(\lambda, a, b + n) \) is similar.

The following Theorem is a generalization of the above one.

**Theorem 4.2.** Let \( X \sim BSN(\lambda, a, b) \) be independent of \( Y \sim BSN(\lambda, c, 1) \) and of \( Z \sim BSN(\lambda, 1, d) \). Then \( X \mid Y \leq X \sim BSN(\lambda, a + c, b) \) and \( X \mid Z \geq X \sim BSN(\lambda, a, b + d) \), where \( c, d \in \mathbb{R} \).

**Proof.** The proof follows the same general lines as that of the Theorem 4.1.

**Theorem 4.3.** If \( X \sim BSN(\lambda, a, b) \) is independent of \( U_1, \ldots, U_n, V_1, \ldots, V_m \) i.i.d. \( SN(\lambda) \) then \( X \mid \{U_{(n)} \leq X, V_{(1)} \geq X\} \sim BSN(\lambda, a + n, b + m) \), where \( U_{(n)} = \max(U_1, \ldots, U_n) \) and \( V_{(1)} = \min(V_1, \ldots, V_m) \).

**Proof.** The proof is quite similar to the one of Theorem 4.1.

Theorem 4.1 can be used to generate \( X \sim BSN(\lambda, n, 1) \) by extending the acceptance-rejection technique, due to Azzalini, as follows (see Azzalini, 1985; Sharafi and Behboodian, 2008): first we generate a random sample \( T, U_1, U_2, \ldots, U_{n-1} \) from \( SN(\lambda) \), if \( \max(U_1, U_2, \ldots, U_{n-1}) \leq T \) we put \( X = T \). Otherwise, we generate a new random sample, until the above inequality is satisfied.

**4.1. Further Results Concerning the Balakrishnan Skew-Normal and the Beta Skew-Normal**

In this section we present some results concerning the SNB distribution and link the distributions introduced in Sec. 2.2 with the Beta skew-normal. First, we consider two results about the Balakrishnan skew-normal. We study the distribution of the largest order statistic from \( SNB_m(1) \) and subsequently the distribution of the smallest order statistic from \( SNB_m(-1) \). We found that these distributions belong to the family of SNB.

**Proposition 4.4.** Let \( X_1, \ldots, X_n \) be a random sample from a \( SNB_m(1) \). Then

\[
X_{(n)} = \max\{X_1, \ldots, X_n\}
\]

is a \( SNB_k(1) \), where \( k = n(m + 1) - 1 \).

**Proof.** The proof follows easily using the standard formula for the density of \( X_{(n)} \), the largest order statistic of a random sample of size \( n \) from the distribution \( SNB_m(1) \).

In particular the following Corollaries hold.

**Corollary 4.4.** Let \( X_1, \ldots, X_n \) be a random sample from a \( SN(1) \). Then

\[
X_{(n)} = \max\{X_1, \ldots, X_n\}
\]

is a \( SNB_{2n-1}(1) \).
2242 Mameli and Musio

Proof. The skew-normal distribution with parameter \( \lambda = 1 \) is a Balakrishnan skew-normal with parameters \( \lambda = 1 \) and \( m = 1 \).

The same result can be established making use of the well-known result for the density of the largest order statistic from the distribution \( SN(1) \) and Property 2.

Corollary 4.5. Let \( X_1, \ldots, X_n \) be a random sample from a \( SNB_m(-1) \). Then

\[
X_{(1)} = \min \{X_1, \ldots, X_n\}
\]

is a \( SNB_k(-1) \), where \( k = n(m + 1) - 1 \).

It follows immediately from Corollaries (4.1), (4.2), (4.4), and (4.5) that the BSN distribution is related to the previous skew-normal generalizations. In fact, its density simplifies to the Balakrishnan skew-normal when \( b = 1, \lambda = 1 \), and \( a \geq 1 \) integer (or \( a = 1, \lambda = -1 \), and \( b \geq 1 \) integer). Further, if \( \lambda = 0 \) it reduces to the generalized Balakrishnan skew-normal when \( a \) and \( b \) are both integers. These considerations have been summarized in the following Proposition.

Proposition 4.5. The BSN distribution satisfies the following properties:

- \( g^{b}_{\beta(1)}(x; 1, n, 1) = f_{2n-1,m}(x; 1, 0) \) for all \( x \in \mathbb{R} \), i.e., \( BSN(1, n, 1) = \overline{TBSN}_{2n-1,m}(1, 0) \);
- \( g^{b}_{\beta(-1)}(x; -1, 1, m) = f_{n,2m-1}(x; 0, 1) \) for all \( x \in \mathbb{R} \), i.e., \( BSN(-1, 1, m) = TBSN_{n,2m-1}(0, -1) \);
- \( g^{b}_{\beta(0)}(x; 0, n, m) = f_{n-1,m-1}(x; -1, 1) \) for all \( x \in \mathbb{R} \), i.e., \( BSN(0, n, m) = TBSN_{n-1,m-1}(1, -1) \);

where \( n, m \) are positive integer numbers.

Given a random variable \( X \sim BSN(\lambda, a, b) \) we are interested in constructing a random variable \( Y \) with Kumaraswamy distribution. This goal can be achieved using the below properties which follow easily from Property g and Property h of Sec. 4, respectively.

Property 4.3. If \( X \sim BSN(\lambda, 1, b) \) then \( Y = (\Phi(X; \lambda))^{1/b} \) is a \( Kumaraswamy(a, b) \). In particular, if \( X \sim SNB_{2m-1}(-1) \) then \( Y = (1 - \Phi(-X))^{1/b} \) is a \( Kumaraswamy(a, b) \).

Property 4.4. If \( X \sim BSN(\lambda, a, 1) \) then \( Y = (1 - \Phi(X; \lambda))^{1/a} \) is a \( Kumaraswamy(b, a) \). In particular, if \( X \sim SNB_{2m-1}(1) \) then \( Y = (1 - \Phi(X))^{1/a} \) is a \( Kumaraswamy(b, a) \).

Recently, Ferreira and Steel (2006) presented a general approach which allows to generate skew distributions. They show that every univariate continuous skew distribution can be obtained from a “perturbation” of a symmetric one as it explained in the following definition.

Definition 4.1. A distribution \( S \) is said to be a skewed version of the symmetric distribution \( F \), generated by the skewing mechanism \( P \), if its pdf is of the form

\[
s(y | f, p) = f(y)p[F(y)], \quad y \in \mathbb{R},
\]

(20)
where \( f \) and \( F \) are the pdf and cumulative distribution function (cdf) of a symmetric distribution on the real line, respectively, and \( p \) is the pdf of a distribution on \((0, 1)\).

Note that, if \( F \) is the standard normal distribution and \( p \) on \((0, 1)\) is given by

\[
p(u; \lambda, a, b) = \frac{2}{B(a, b)} \Phi(\phi^{-1}(u)) \left( \Phi^{-1}(u); \lambda \right)^{a-1} \left( 1 - \Phi^{-1}(u); \lambda \right)^{b-1}.
\]  

(21)

formula (20) reduces to the Beta skew-normal with parameters \( \lambda, a, b \). Then the pdf of a Beta skew-normal with parameters \( \lambda, a, b \) can be seen as a weighted version of \( \phi(y) \), with weight function given by \( p(\Phi(y); \lambda, a, b) \).

5. Conclusions

In this article, we have studied a class of distributions, referred to as the Beta skew-normal (BSN), which extends the skew-normal and the Beta-normal distribution. For special values of the parameters this distribution also includes the Balakrishnan skew-normal (SNB), the generalized Balakrishnan skew-normal (GBSN), and a two-parameter generalization of the Balakrishnan skew-normal (TBSN).

We provide a mathematical treatment of the new distribution. We derived various properties of the BSN, including the moment generating function, recurrence relations for moments and two methods for simulating. Some results presented in this article, for example Theorems from 4.1–4.3, can be adapted for other distributions belonging to the family of Beta generated distribution, such as the Beta-normal.

Acknowledgments

We thank the Editor and Referees for many helpful comments which greatly improved the quality of the article.

References


