Note on a new generalization of the skew-normal distribution

V. Mameli and M. Musio

Abstract In this paper we define and study a three-parameter distribution, referred to as the Beta skew-normal distribution (*BSN*), which is a generalization of the skew-normal distribution introduced by Azzalini [1]. This family is obtained using the generator approach suggested by Eugene et al. [2] and Jones [3]. Some properties of the proposed distribution are discussed: among others the moment generating function, a recursion formula for its moments and two different methods which allow to simulate a *BSN* distribution. The densities in the family have a symmetric or asymmetric, unimodal or bimodal shape, depending on the values of the parameters. Some of the results presented in this work can be adapted for other distributions belonging to the family of the Beta-generated distribution, such as the Beta-normal (see [2]).

Key words: skew-normal, Beta-generated family, bimodal distributions

1 Introduction

The skew-normal distribution (*SN*), introduced by Azzalini [1], has been studied and generalized extensively. However, a limitation of this family of densities is the fact that the distributions are unimodal. The principal purpose of this paper is to introduce a new generalization of the skew-normal distribution, which can yield unimodal as well as bimodal densities. This new family is called Beta skew-normal (*BSN*). The Beta-normal distribution (*BN*), proposed by Eugene et al. [2], is a special case of the new model. Furthermore, the *BSN* distribution contains the Beta half-normal one [6] as a limiting case. We present some properties of this new distribution. We show that the distributions of order statistics from the skew-normal distribution are Beta skew-normal and log-concave.

V. Mameli and M.Musio
University of Cagliari, via Ospedale 72, e-mail: mameli.valentina@virgilio.it, mmusio@unica.it

2 A new generalization of the skew-normal distribution: the Beta skew-normal

In this section we define the Beta skew-normal class. This new distribution, which contains as a special case the skew-normal one, is flexible enough to support uni-modal and bimodal shapes.

Definition 1. A random variable X is said to have a Beta skew-normal distribution, if its density is given by

$$g^B_{\Phi(x;\lambda)}(x;\lambda,a,b) = \frac{2}{B(a,b)}(\Phi(x;\lambda))^{a-1}(1-\Phi(x;\lambda))^{b-1}\phi(x)\Phi(\lambda x), \; x \in R \quad (1)$$

with $a > 0$, $b > 0$, $\lambda \in R$. We denote the random variable X by $X \sim BSN(\lambda,a,b)$.

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

a.   If $a = 1$ and $b = 1$, then we obtain the $SN(\lambda)$ distribution.
b.   If $\lambda = 0$, then the BSN distribution reduces to the Beta-normal $BN(a,b)$ one.
c.   If $X \sim BSN(0,1,1)$, then X is a standard normal random variable.
d.   If $X \sim BSN(1,\frac{1}{2},1)$, then X is a standard normal random variable.
e.   If $X \sim BSN(-1,1,\frac{1}{2})$, then X is a standard normal random variable.
f.   If $X \sim BSN(\lambda,a,b)$, then $-X \sim BSN(-\lambda,b,a)$.
g.   If $X \sim BSN(\lambda,a,b)$, then $Y = \Phi(X;\lambda) \sim Beta(a,b)$ and
     $Z = 1 - \Phi(X;\lambda) \sim Beta(b,a)$.
h.   As $\lambda \to +\infty$, the BSN density tends to the Beta half-normal distribution.

Properties from c to e establish that the family of $BSN(\lambda,a,b)$ contains the standard normal distribution as a special case under three different parameter sets. Consequently, we have that a probabilistic model based on the BSN distribution is not identifiable under the null hypothesis of normality. It is well known that the classical asymptotic results concerning the likelihood ratio are not true in case of loss of identifiability and the distribution of the LRT statistic is difficult to characterize ([4]). An way to avoid this problem is to redefine the model over the modified parameter space

$$\Theta = \{(\lambda,a,b)|a \geq 1, b \geq 1, \lambda \in R\}.$$

Therefore, under this constraint on the parameter space, H_0 is described by $a = 1$, $b = 1$ and $\lambda = 0$. 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. If $a,b \geq 1$, the density 1 is strongly unimodal [3], i.e.$\log g^B_{\Phi(x;\lambda)}(x;\lambda,a,b)$ is a concave function of x (see Fig. 1(a)). We do not have general results for a and/or $b < 1$. A numerical study has shown that, when at least one of the two parameters a and b is close to zero $(0.10,0.20)$, the density can be bimodal (see Fig. 1(b)). Now we find the moment generating function of X which has density 1.
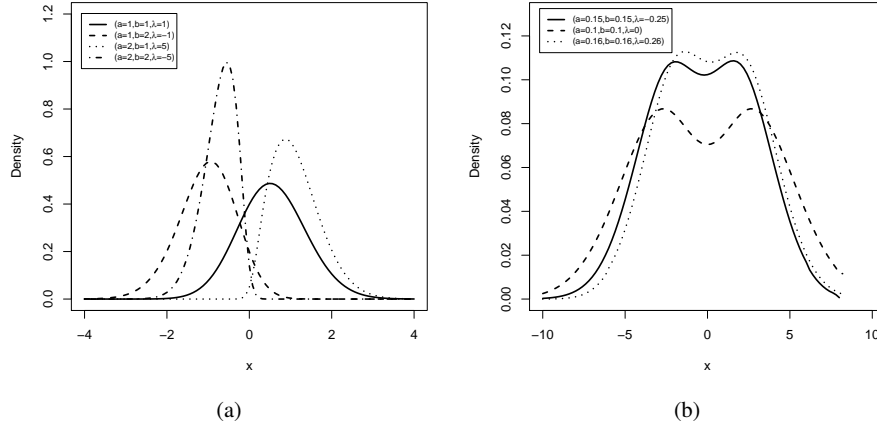
Fig. 1 The *BSN* density function for some values of *a*, *b* and λ

Proposition 1. *The moment generating function of* $X \sim BSN(\lambda, a, b)$ *is given by*

$$M_X(t) = \frac{2}{B(a,b)} e^{\frac{t^2}{2}} E_Z \left((\Phi(Z;\lambda))^{a-1} (1 - \Phi(Z;\lambda))^{b-1} \Phi(\lambda Z) \right), \text{where } Z \sim N(t,1).$$
(2)

We have the following recursive formula.

Proposition 2. *Let* $k \in N$ *and* $k \geq 1$. *If* $X \sim BSN(\lambda, a, b)$, *with* $a, b > 1$ *then*

$$E_X(X^k) = (k-1)E_X(X^{k-2}) + \lambda E_X \left(X^{k-1} \frac{\phi(\lambda X)}{\Phi(\lambda X)} \right) +$$
$$+ (a+b-1)E_U \left(U^{k-1} \phi(U;\lambda) \right) - (a+b-1)E_V \left(V^{k-1} \phi(V;\lambda) \right), (3)$$

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

Moments of the BSN cannot be evaluated exactly in closed form. We have computed them numerically using the software R. From this numerical study we have noted that (see [5]):

- for fixed values of *a* and *b* the mean and skewness are both increasing function of λ;
- for fixed values of *b* and λ the mean and skewness are both increasing function of *a*;
- for fixed values of *a* and λ the mean is a decreasing function of *b* .

The Beta skew-normal density is in general asymmetric (see Figs. 1(a) and 1(b)). We have a partial result concerning symmetry:

Proposition 3. *If $a = b$ and $BSN(\lambda, a, b)$ is symmetric about 0 then $\lambda = 0$.*

We will see now that the *BSN* distribution, with *a* and *b* integers, is the distribution of order statistics from a skew-normal distribution.

Proposition 4. *Let $X_1, \cdots, X_n$ be a random sample from a $SN(\lambda)$. Then the j-th order statistic is a $BSN(\lambda, j, n - j + 1)$, where $j = 1, \cdots, n$.*

In particular, we have the following Corollary:

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

We now derive other properties of the *BSN* distribution of general interest.

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

Theorem 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 $X \,|\, Y \leq X \sim BSN(\lambda, a + c, b)$ and $X \,|\, Z \geq X \sim BSN(\lambda, a, b + d)$, where c, d are positive real numbers.*

Theorem 3. *If $X \sim BSN(\lambda, a, b)$ is independent of $U_1, \cdots, U_n, V_1, \cdots, V_m$ i.i.d. $SN(\lambda)$ then $X \,|\, (U_{(n)} \leq X, V_{(1)} \geq X) \sim BSN(\lambda, a + n, b + m)$, where $U_{(n)} = \max(U_1, \cdots, U_n)$ and $V_{(1)} = \min(V_1, \cdots, V_m)$.*

The above three Theorems can be generalized to Jones' family of distributions. Theorem 1 can be used to generate $X \sim BSN(\lambda, n, 1)$: we generate a random sample $U_1, U_2, \cdots, U_n$ from $SN(\lambda)$, then we take the random variable X defined as $X = \max(U_1, U_2, \cdots, U_n)$ which has a $BSN(\lambda, n, 1)$ distribution.

References

1. Azzalini, A.: A Class of Distributions Which Includes the Normal Ones. Scandinavian Journal of Statistics, 12(2), 171–178 (1985).
2. Eugene, N., Lee, C., Famoye, F.: Beta-normal distribution and its applications. Communications in Statistics - Theory and methods, 31(4), 497–512 (2002).
3. Jones, M. C.: Families of distributions of order statistics. Test, 13(1), 1–43 (2004).
4. Lindsay, B. G.: Mixture Models: Theory, Geometry and Applications. IMS, Hayward, CA (1995).
5. Mameli, V., Musio, M.: A generalization of the skew-normal distribution: the beta skew-normal. Communications in Statistics - Theory and Methods, to appear, doi:10.1080/03610926.2011.607530
6. Pescim, R. R. Demtrio, C. G. B., Cordeiro G. M. E., Ortega, M. M.,Urbanoa, M. R.: The beta generalized half-normal distribution. Computational Statistics and Data Analysis, 54(4), 945–957 (2010).