Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 1, pp. 9–33
DOI : 10.47974/JSMS-1013

A generalization of Balakrishnan-alpha-skew-normal distribution : Properties, characterisations and applications

S. Shah †
P. J. Hazarika *
S. Chakraborty §
Department of Statistics
Dibrugarh University
Dibrugarh 786004
Assam
India

G. G. Hamedani ‡
Department of Mathematical and Statistical Sciences
Marquette University
Wisconsin
U.S.A.

Abstract

In this article, a generalized form of the Balakrishnan-alpha-skew-normal distribution of Hazarika et al. (2020) is proposed and some of its simple and distributional properties are studied. Two characterizations have also been presented. For model selection criteria namely Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), are used to check appropriateness of the proposed distribution by conducting data fitting experiments and comparing it with some other related distributions. The likelihood ratio test is employed for discriminating between nested models.

Subject Classification: *60E05, 62H10, 62H12.*

Keywords: *Skew distribution, Bimodal distribution, Alpha-skew-normal distribution, Balakrishnan-alpha-skew-normal distribution, AIC, BIC.*

† *E-mail:* **charan.shah90@gmail.com**
E-mail:* **parthajhazarika@gmail.com (Corresponding Author)
§ *E-mail:* **subrata_arya@yahoo.co.in**
‡ *E-mail:* **gholamhoss.hamedani@marquette.edu**

1. Introduction

The applications of skewed distributions arise in different fields of sciences, engineering and medicine. Azzalini (1985) first introduced the path breaking skew-normal distribution by adding a parameter to normal distribution to introduce asymmetry. Many researchers have studied this distribution and its generalization. Some of the notable researchers include Sharafi and Behboodian (2008), Yadegari *et al.* (2008), Wahed and Ali (2001), Arnold and Beaver (2002), Ma and Genton (2004), Hasanalipour and Sharafi (2012), Gupta and Gupta (2004), Jamalizadeh *et al.* (2008), Bahrami *et al.* (2009), Chakraborty *et al.* (2015) among others. A generalization of the skew normal distribution was proposed by Balakrishnan (2002) as a discussant in Arnold and Beaver. They studied its properties and the probability density function (pdf) is given by

$$f_Z(z;\lambda,n)=\phi(z)[\Phi(\lambda z)]^n/C_n(\lambda), \tag{1}$$

where n is a positive integer, $\Phi(.)$ is the cumulative distribution function (cdf) of standard normal random variable and C $C_n(\lambda)=\mathrm{E}(\Phi^n(\lambda U))$, $U\sim \mathrm{N}(0,1)$.

The general formula for the construction of skewed distributions was proposed by Huang and Chen (2007) starting from a symmetric (about 0) pdf $h(.)$ by introducing the concept of skew function $G(.)$ which is a Lebesgue measurable function under the condition that $0\le G(z)\le 1$ and $G(z)+G(-z)=1,\ z\in R,$ almost everywhere. The pdf of the same is given by

$$f(z)=2h(z)G(z), z\in R \tag{2}$$

A new form of skew distribution which has both unimodal as well as bimodal behavior was introduced by Elal-Olivero (2010) and called the alpha-skew-normal distribution. Its density function is given by

$$f(z;\alpha)=\left(\frac{(1-\alpha z)^2+1}{2+\alpha^2}\right)\phi(z), z\in R. \tag{3}$$

Along the same line of eqn. (3), Venegas *et al.* (2016) studied the logarithmic form of alpha-skew-normal distribution and used for modelling chemical data. Louzada et al. (2017) introduced a bivariate extension of the univariate alpha-skew-normal distribution and studied some of its properties. Sharafi *et al.* (2017) introduced a generalization of $ASN(\alpha)$ distribution. Shah *et al.* (2021) introduced the generalization of $ABSN(\alpha,\beta)$ (alpha- beta skew normal distribution) distribution of Shafiei *et al.* (2016). Hazarika *et al.* (2020) introduced a new family of skew distributions with more flexibility than the Azzalini (1985) and the Elal-Olivero (2010) distributions and is known as the Balakrishnan-alpha-skew-normal distribution. Its density function is defined as

$$f(z;\alpha,\beta)=\left(\frac{[(1-\alpha z)^2+1]^2}{4+8\alpha^2+3\alpha^4}\right)\phi(z), z\in R, \alpha\in R. \tag{4}$$

Using the similar approach, Shah *et al.* (2020a, 2020b, 2020c and 2020d) introduced the Balakrishnan-alpha-skew-(logistic, Laplace, truncated Cauchy and log-normal) distributions.

The goals of considering this new distribution are as follows: *First,* this distribution includes three important classes of distribution which are the normal distribution, the skew-normal distribution and the Balakrishnan-alpha-skew-normal distribution. *Second,* this distribution is flexible enough to support both unimodality and bimodality behaviours and has at most two modes. *Third,* for some real life data, this proposed distribution provides a better fitting compared to the other distributions considered here. The rest of this article is summarized as follows: In Section 2 we define the proposed distribution and some of its basic properties are investigated. Distributional properties of the proposed distribution are studied in Section 3. Two characterization of the proposed distribution are given in Section 4. The parameter estimation and real life applications of the proposed distribution along with a simulation study are provided in Section 5. The concluding remarks are given in Section 6.

2. A Generalized Balakrishnan-Alpha-Skew-Normal Distribution

In this section we introduce a new generalized form of the alpha-skew-normal distribution and investigated some of its basic properties.

Definition 1 : If a continuous random variable Z has a density function

$$f(z;\alpha,\lambda)=\frac{[(1-\alpha z)^2+1]^2}{C_2(\alpha,\lambda)}\phi(z)\Phi(\lambda z),\ z\in R, \tag{5}$$

where, $C_2(\alpha,\lambda)=2+4\alpha^2+\frac{3\alpha^4}{2}-b\delta\left(4\alpha+\frac{2\alpha^3(3+2\lambda^2)}{1+\lambda^2}\right)$, then, it is said to be the generalized Balakrishnan-alpha-skew-normal distribution denoted by $GBASN_2(\alpha,\lambda)$ with the skewness parameters $\alpha\in R$ and $\lambda\in R$. The normalizing constant $C_2(\alpha,\lambda)$ is calculated as follows:

$$
\begin{aligned}
C_2(\alpha,\lambda) &= \int_{-\infty}^{\infty} [(1-\alpha z)^2+1]^2 \phi(z)\Phi(\lambda z)\,dz \\
&= \int_{-\infty}^{\infty} (4-8\alpha z+8\alpha^2 z^2-4\alpha^3 z^3+\alpha^4 z^4)\phi(z)\Phi(\lambda z)\,dz \\
&= 2\int_{-\infty}^{\infty} \phi(z;\lambda)\,dz - 4\alpha E(Z_\lambda)+4\alpha^2 E(Z_\lambda{}^2)-2\alpha^3 E(Z_\lambda{}^3)+\frac{\alpha^4}{2}E(Z_\lambda{}^4) \\
&= 2-4\alpha\left(\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}\right)+4\alpha^2-2\alpha^3\left(\frac{6\lambda+4\lambda^3}{\sqrt{2\pi}(\sqrt{1+\lambda^2})^3}\right)+\frac{3\alpha^4}{2} \\
&= 2+4\alpha^2+\frac{3\alpha^4}{2}-\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}\left(4\alpha+\frac{2\alpha^3(3+2\lambda^2)}{1+\lambda^2}\right) \\
&= 2+4\alpha^2+\frac{3\alpha^4}{2}-b\delta\left(4\alpha+\frac{2\alpha^3(3+2\lambda^2)}{1+\lambda^2}\right).
\end{aligned}
$$

where $b=\sqrt{\frac{2}{\pi}}, \delta=\frac{\lambda}{\sqrt{1+\lambda^2}}$ and $\phi(z;\lambda)$ is the density function skew normal distribution of Azzalini (1985), i.e., $Z_\lambda \sim SN(\lambda)$.

2.1 *Properties of* $GBASN_2(\alpha,\lambda)$

- If $\alpha=0$, then we get the skew normal $SN(\lambda)$ distribution of Azzalini (1985) and which is given by $f(z;\lambda)=2\varphi(z)\Phi(\lambda z)$.
- If $\lambda=0$, then we get the Balakrishnan-alpha-skew-normal $BASN_2(\alpha)$ distribution of Hazarika *et al.* (2020) and which is given by

$$f(z;\alpha)=[(1-\alpha z)^2+1]^2\varphi(z)/(4+8\alpha^2+3\alpha^4).$$

- If $\alpha=\lambda=0$, then we get the standard normal distribution given by $f(z)=\varphi(z)$.
- If $\alpha\to\pm\infty$, then we get the generalized bimodal normal $GBN(4)$ distribution given by $f(z;\lambda)=2z^4\varphi(z)\Phi(\lambda z)/3$ (see Hazarika *et al.*, 2020).
- For fixed α, if, $\lambda\to+\infty$ then

$$f(z;\alpha)=\frac{[(1-\alpha z)^2+1]^2}{2+4\alpha^2+\frac{3\alpha^4}{2}-4\sqrt{\frac{2}{\pi}}\alpha(1+\alpha^2)}\varphi(z)I(z>0).$$

and if, $\lambda \to -\infty$

then $$f(z;\alpha) = \frac{[(1-\alpha z)^2+1]^2}{2+4\alpha^2+\frac{3\alpha^4}{2}+4\sqrt{\frac{2}{\pi}}\alpha(1+\alpha^2)}\varphi(z)I(z<0).$$

- If $Z \sim GBASN_2(\alpha,\lambda)$, then $-Z \sim GBASN_2(-\alpha,-\lambda)$.

2.2 *Plots of the pdf*

The pdf of $GBASN_2(\alpha,\lambda)$ distribution for different choices of the parameters α and λ are plotted in Figure 1.

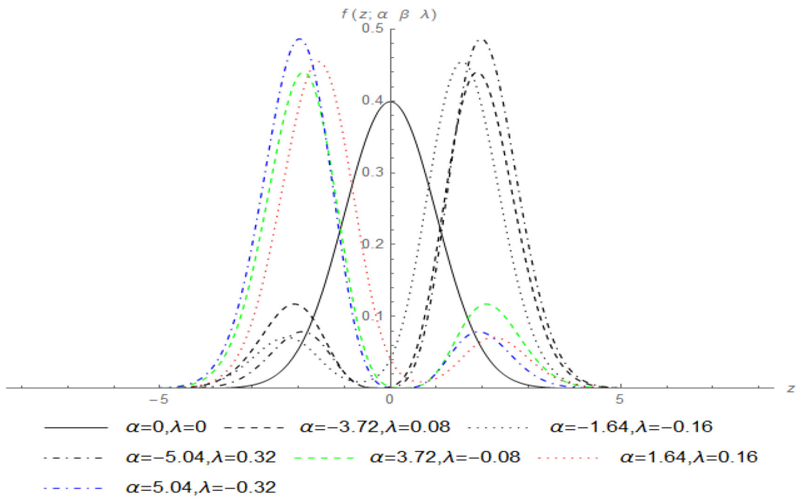


Figure 1

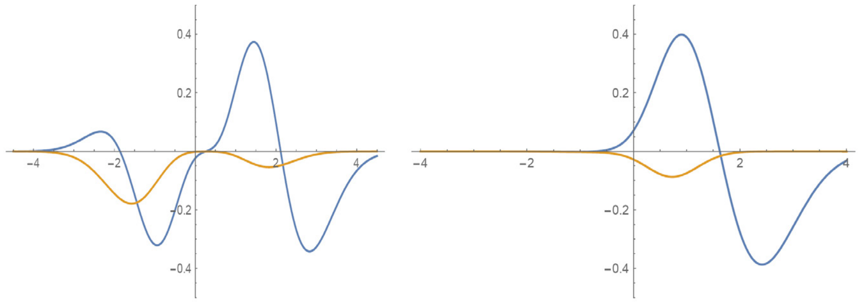
Plots of the pdf of $GBASN_2(\alpha,\lambda)$

2.3 *Mode of $GBASN_2(\alpha,\lambda)$*

Theorem 1: *The $GBASN_2(\alpha,\lambda)$ distribution has at most two modes.*

Proof: Let $Z \sim GBASN_2(\alpha,\lambda)$ distribution then by eqn. (4) and eqn. (5), we have

$$f(z;\alpha,\lambda) = \frac{4+8\alpha^2+3\alpha^4}{C_2(\alpha,\lambda)} f(z;\alpha)\Phi(\lambda z) \tag{6}$$



(a) (b)

Figure 2

(a) The Plot of C1 and C2 for $\alpha = 3.85, \lambda = 0.6$.
(b) The Plot of C1 and C2 for $\alpha = -1.53, \lambda = 1.5$

and $$f'(z;\alpha,\lambda) = \frac{4+8\alpha^2+3\alpha^4}{C_2(\alpha,\lambda)}[f'(z;\alpha)\Phi(\lambda z) + \lambda f(z;\alpha)\phi(\lambda z)]. \tag{7}$$

To show that eqn.(6) has at most two modes, it should be proved that eqn.(7) has one or three answers. To do this, let us apply a graphical approach, so

$$f'(z;\alpha,\lambda) = F_1(z) - F_2(z).$$

Where, $F_1(z) = \frac{4+8\alpha^2+3\alpha^4}{C_2(\alpha,\lambda)} f'(z;\alpha)\Phi(\lambda z)$ and

$$F_2(z) = -\frac{4+8\alpha^2+3\alpha^4}{C_2(\alpha,\lambda)} \lambda f(z;\alpha)\phi(\lambda z).$$

Set $f'(z;\alpha,\lambda) = 0$, hence $F_1(z) = F_2(z)$. (8)

Since eqn. (5) vanishes out of (-5, 5), we draw the curves of $C1: y = F_1(z)$ for $\alpha = 3.85$, $\lambda = 0.6$ and $C2: y = F_2(z)$ for $\alpha = -1.53$, $\lambda = 1.5$. As it can be seen in Figure 2 that the two curves have at least one and at most three intersection points and the values of z of these points are the roots of eqn. (8). Since $\lim_{z\to\pm\infty} f(z;\alpha,\lambda) = 0$, then if eqn. (8) has one answer it should be the mode of the eqn. (6) and if eqn. (8) has three answers, then eqn.(6) should have two modes and hence the $GBASN_2(\alpha,\lambda)$ distribution has at least one and at most two modes.

3. Distributional Properties

3.1 *Cumulative Distribution Function*

Theorem 2: *The cdf of $GBASN_2(\alpha,\lambda)$ distribution is given by*

$$F_Z(z) = \frac{1}{2(1+\lambda^2)^{3/2} C(\alpha,\lambda)}[-b\delta\alpha^2\varphi(z\sqrt{1+\lambda^2})\{-8(1+\lambda^2)+\alpha(5\alpha-4z+\alpha z^2$$
$$+\lambda^2(-4z+3\alpha+\alpha z^2)\}+(1+\lambda^2)\{-4b\lambda\alpha(1+\alpha^2)Erf(z\sqrt{1+\lambda^2}/\sqrt{2})$$
$$-\alpha(-8+8\alpha z-4\alpha^2(2+z^2)+\alpha^3 z(3+z^2))\sqrt{1+\lambda^2}\,\Phi(z;\lambda)$$
$$-4b\delta\alpha^3\Phi(z\sqrt{1+\lambda^2})+(5+3\alpha^4)\sqrt{1+\lambda^2}\,\Phi(z;\lambda)\}] \tag{9}$$

where $Erf(z)$ *gives the error function and is the integral of the Gaussian distribution, given by* $Erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$, $\Phi(z;\lambda)$ *is the cdf of skew normal distribution of Azzalini (1985).*

***Proof*:**

$$F_Z(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{[(1-\alpha t)^2+1]^2}{C(\alpha,\lambda)}\varphi(t)\Phi(\lambda t)dt$$
$$= \frac{1}{C(\alpha,\lambda)}\int_{-\infty}^{z}[4-8\alpha t+8\alpha^2 t^2-4\alpha^3 t^3+\alpha^4 t^4]\varphi(t)\Phi(\lambda t)dt$$
$$= \frac{1}{C(\alpha,\lambda)}\left[\int_{-\infty}^{z} 4\phi(t)\Phi(\lambda t)dt - 8\alpha\int_{-\infty}^{z} t\phi(t)\Phi(\lambda t)dt + 8\alpha^2\int_{-\infty}^{z} t^2\phi(t)\Phi(\lambda t)dt\right.$$
$$\left.-4\alpha^3\int_{-\infty}^{z} t^3\phi(t)\Phi(\lambda t)dt + \alpha^4\int_{-\infty}^{z} t^4\phi(t)\Phi(\lambda t)dt\right]. \tag{10}$$

Now, we have

$$\int_{-\infty}^{z} 4\phi(t)\Phi(\lambda t)dt = 2\Phi(z;\lambda),$$
$$\int_{-\infty}^{z} t\phi(t)\Phi(\lambda t)dt = -\varphi(z)\Phi(\lambda z) + \frac{\lambda\, Erf\left((z\sqrt{1+\lambda^2})/\sqrt{2}\right)}{2\sqrt{2\pi}\sqrt{1+\lambda^2}},$$
$$\int_{-\infty}^{z} t^2\phi(t)\Phi(\lambda t)dt = -z\varphi(z)\Phi(\lambda z) + \frac{b\,\delta\,\varphi\left(z\sqrt{1+\lambda^2}\right)}{2\sqrt{1+\lambda^2}},$$
$$\int_{-\infty}^{z} t^3\phi(t)\Phi(\lambda t)dt = -(2+z^2)\phi(z)\Phi(\lambda z) + \frac{\lambda\, Erf\left((z\sqrt{1+\lambda^2})/\sqrt{2}\right)}{\sqrt{2\pi}\sqrt{1+\lambda^2}} - z\frac{b\,\delta\,\phi\left(z\sqrt{1+\lambda^2}\right)}{2\sqrt{1+\lambda^2}} +$$
$$\frac{b\,\delta\,\Phi\left(z\sqrt{1+\lambda^2}\right)}{2\sqrt{1+\lambda^2}},$$
$$\int_{-\infty}^{z} t^4\phi(t)\Phi(\lambda t)dt = -z(3+z^2)\phi(z)\Phi(\lambda z) - \frac{b\,\delta\,\phi\left(z\sqrt{1+\lambda^2}\right)(5+z^2+\lambda^2(3+z^2))}{2\left(\sqrt{1+\lambda^2}\right)^3} + \frac{3}{2}\Phi(z;\lambda).$$

Putting the above results in eqn. (10), we get the desired result.

3.2 Moments

Theorem 3: *The k^{th} order moment of $GBASN_2(\alpha,\lambda)$ distribution is given by*

$$E(Z^k)=\frac{1}{C(\alpha,\lambda)}[2E(Z_\lambda^{\ k})-4\alpha E(Z_\lambda^{\ k+1})+4\alpha^2E(Z_\lambda^{\ k+2})-2\alpha^3E(Z_\lambda^{\ k+3})+\frac{\alpha^4}{2}E(Z_\lambda^{\ k+4})]$$

where $E(Z_\lambda^{\ k})$ is the k^{th} moment of $Z_\lambda \sim SN(\lambda)$. (11)

Proof:

$$E(Z^k)=\int_{-\infty}^{\infty} z^k\frac{[(1-\alpha z)^2+1]^2}{C(\alpha,\lambda)}\varphi(z)\Phi(\lambda z)\,dz$$

$$=\frac{1}{C(\alpha,\lambda)}\Bigg[\int_{-\infty}^{\infty}4z^k\varphi(z)\Phi(\lambda z)dz-8\alpha\int_{-\infty}^{\infty}z^{k+1}\varphi(z)\Phi(\lambda z)\,dz+8\alpha^2\int_{-\infty}^{\infty}z^{k+2}\varphi(z)\Phi(\lambda z)\,dz-$$

$$4\alpha^3\int_{-\infty}^{\infty}z^{k+3}\varphi(z)\Phi(\lambda z)\,dz+\alpha^4\int_{-\infty}^{\infty}z^{k+4}\varphi(z)\Phi(\lambda z)\,dz\Bigg]$$

$$=\frac{1}{C(\alpha,\lambda)}\left[2E(Z_\lambda^{\ k})-4\alpha E(Z_\lambda^{\ k+1})+4\alpha^2E(Z_\lambda^{\ k+2})-2\alpha^3E(Z_\lambda^{\ k+3})+\frac{\alpha^4}{2}E(Z_\lambda^{\ k+4})\right]$$

The moments of $Z\sim GBASN_2(\alpha,\lambda)$ can be obtained by applying the moments of $Z_\lambda \sim SN(\lambda)$ (Henze, 1986). For example, the following results are obtained for $k=1,2,3,4$ as

$$E(Z)=\frac{1}{C_2(\alpha,\lambda)}\left[-2\alpha(2+3\alpha^2)+b\delta\left(\frac{2}{(1+\lambda^2)^2}+\frac{8\alpha^2d_1+\alpha^4d_2}{2(1+\lambda^2)^2}\right)\right],$$

$$E(Z^2)=\frac{1}{C_2(\alpha,\lambda)}\left[2+12\alpha^2+\frac{15\alpha^4}{2}-b\delta\left(\frac{2\alpha(d_3+\alpha^2d_2)}{(1+\lambda^2)^2}\right)\right],$$

$$E(Z^3)=\frac{1}{C_2(\alpha,\lambda)}\left[-6\alpha(2+5\alpha^2)+b\delta\left(\frac{(6+4\lambda^2)}{(1+\lambda^2)}+\frac{16\alpha^2(d_2)}{(1+\lambda^2)^2}+\frac{6\alpha^4(d_4)}{(1+\lambda^2)^3}\right)\right],$$

$$E(Z^4)=\frac{1}{C_2(\alpha,\lambda)}\left[6+60\alpha^2+\frac{105\alpha^4}{2}-b\delta\frac{2\alpha(d_5+3\alpha^2d_4)}{(1+\lambda^2)^3}\right],$$

$$Var(Z)=-\frac{[\{b\delta(4+8d_1\alpha^2+d_2\alpha^4)-(8\alpha+12\alpha^3)(1+\lambda^2)^2\}^2-2(1+\lambda^2)^2C_2(\alpha,\lambda)\{-4\alpha b\delta(d_3+d_2\alpha^2)+(4+24\alpha^2+15\alpha^4)(1+\lambda^2)^2\}]}{4(1+\lambda^2)^4[C_2(\alpha,\lambda)]^2}.$$

where $d_1=3+5\lambda^2+2\lambda^4$, $d_2=15+20\lambda^2+8\lambda^4$, $d_3=6+10\lambda^2+4\lambda^4$, $d_4=35+70\lambda^2+56\lambda^4+16\lambda^6$ and $d_5=30+70\lambda^2+56\lambda^4+16\lambda^6$.

Remark 1: The numerical optimization of the mean $E(Z)$ and the variance $Var(Z)$ with respect to the parameters α and λ gives the following bounds for the mean and the variance as $-2.12769\le E(Z)\le 2.12769$ and $0.262814\le Var(Z)\le 5.$ This can also be visualized from Figure 3(a) - 3(b).

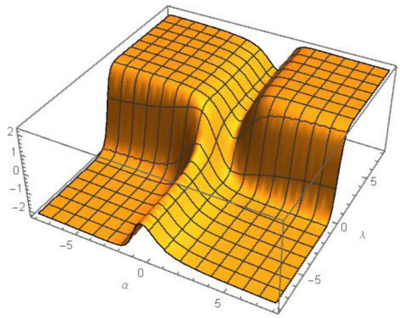


Figure 3(a)

Plot of mean

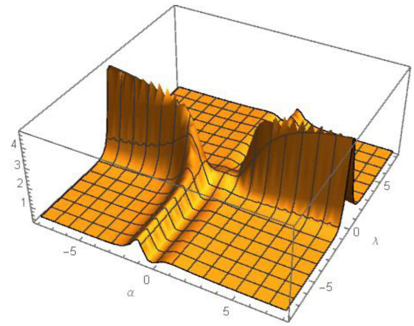


Figure 3(b)

Plot of variance

Remark 2: If we take the limit $\alpha \to \pm\infty$ in the moments of $GBASN_2(\alpha, \lambda)$ distribution, we obtained the moments of $GBN(4)$ distribution as

$$E(Z) = \frac{b\delta d_2}{3(1+\lambda^2)^2} \text{ and } Var(Z) = \frac{45\pi(1+\lambda^2)^5 - 2\lambda^2 d_2{}^2}{9\pi(1+\lambda^2)^5}.$$

Remark 3: By taking the limit $\lambda \to +\infty\, or -\infty$ in the moments of $GBASN_2(\alpha, \lambda)$ distribution, then we get

when $\lambda \to +\infty$ then $E(Z) = \frac{4\alpha(2+3\alpha^2) - 4b(1+4\alpha^2+2\alpha^4)}{8\alpha b(1+\alpha^2) - (2+\alpha^2)(2+3\alpha^2)}$ and

$$Var(Z) = \frac{-8\alpha\sqrt{2\pi}(2+\alpha^2)^2(1+3\alpha^2) - 32(1-2\alpha^4)^2 + \pi(2+3\alpha^2)(2+5\alpha^2)(4+3\alpha^4)}{\pi[-8\alpha b(1+\alpha^2) + (2+\alpha^2)(2+3\alpha^2)]}.$$

and when $\lambda \to +\infty$ then $E(Z) = -\frac{4\alpha(2+3\alpha^2) + 4b(1+4\alpha^2+2\alpha^4)}{8\alpha b(1+\alpha^2) + (2+\alpha^2)(2+3\alpha^2)}$ and

$$Var(Z) = \frac{8\alpha\sqrt{2\pi}(2+\alpha^2)^2(1+3\alpha^2) - 32(1-2\alpha^4)^2 + \pi(2+3\alpha^2)(2+5\alpha^2)(4+3\alpha^4)}{\pi[8\alpha b(1+\alpha^2) + (2+\alpha^2)(2+3\alpha^2)]}.$$

Remark 4: The skewness and kurtosis of $GBASN_2(\alpha, \lambda)$ distribution are obtained respectively, by using the formulae as

$$\beta_1 = \frac{(E(Z^3) - 3E(Z^2)E(Z) + 2[E(Z)]^3)^2}{(E(Z^2) - [E(Z)]^2)^3}$$

and
$$\beta_2 = \frac{E(Z^4) - 4E(Z^3)E(Z) + 6E(Z^2)[E(Z)]^2 - 3[E(Z)]^4}{(E(Z^2) - [E(Z)]^2)^2}.$$

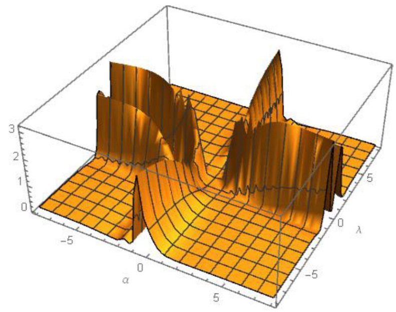


Figure 4(a)
Plot of skewness

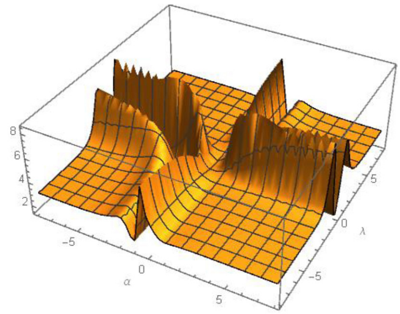


Figure 4(b)
Plot of kurtosis

These cannot be simplified as the expression are very vast. The following bounds for the skewness and kurtosis can be calculated using numerically optimizing β_1 and β_2 with respect to the parameters α and λ as $0 \le \beta_1 \le 3.18813$ and $1.4 \le \beta_2 \le 8.53099.$ This can also be visualized from Figure 4(a) - 4(b).

Remark 5: The skewness and kurtosis of $GBN(4)$ distribution can be obtained easily by taking limit $\alpha \to \pm\infty$ in the results of $GBASN_{\,2}(\alpha, \lambda)$ distribution as

$$\beta_1 = \frac{2\lambda^2[4\lambda^2 d_2{}^3 - 27\pi(1+\lambda^2)^4 d_6]^2}{[45\pi(1+\lambda^2)^5 - 2\lambda^2 d_2{}^2]^3} \text{ and } \beta_2 = \frac{3[945\pi^2(1+\lambda^2)^{10} - 4\lambda^4 d_2{}^4 + 36\pi\lambda^2(1+\lambda^2)^4 d_7]}{[45\pi(1+\lambda^2)^5 - 2\lambda^2 d_2{}^2]^2}.$$

Where $d_6 = 40 + 105\lambda^2 + 84\lambda^4 + 24\lambda^6$ and $d_7 = 75 + 625\lambda^2 + 1160\lambda^4 + 960\lambda^6 + 384\lambda^8 + 64\lambda^{10}$.

4. Characterizations of $GBASN_{\,2}$ (α, λ) distribution

In this section two characterisations of the proposed distribution are presented (see Hamedani et al., 2021). The first one uses conditional distribution and is very useful in generating random sample. The second characterization is based on the ratio of two truncated moments which is based on a theorem due to Glanzel (1987), see Proposition 1 of Appendix.

Theorem 4: *If* $W \sim BASN_2(\alpha)$ *and* $X \sim N(0,1)$ *are independent, then* $W \,|\, \{\lambda W > X\} \sim GBASN_{\,2}(\alpha, \lambda)$.

Proof: Let $Z = W \mid \{\lambda W > X\}$. Now, we write

$$P(Z \le z) = P(W \le z \mid \lambda W > X) = \frac{P(W \le z, \lambda W > X)}{P(\lambda W > X)}$$

So, by the eqn. (4), we have

$$P(W \le z, \lambda W > X) = \int_{-\infty}^{z} \frac{[(1-\alpha u)^2+1]^2}{4+8\alpha^2+3\alpha^4} \varphi(u)\Phi(\lambda u)du \text{ and}$$

$$P(\lambda W > X) = \int_{-\infty}^{\infty} \frac{[(1-\alpha u)^2+1]^2}{4+8\alpha^2+3\alpha^4} \varphi(u)\Phi(\lambda u)du = \frac{1}{4+8\alpha^2+3\alpha^4} C_2(\alpha,\lambda)$$

Therefore,

$$P(Z \le z) = \frac{\int_{-\infty}^{z} \frac{[(1-\alpha u)^2+1]^2}{4+8\alpha^2+3\alpha^4} \varphi(u)\Phi(\lambda u)du}{\frac{1}{4+8\alpha^2+3\alpha^4} C_2(\alpha,\lambda)} = \int_{-\infty}^{z} \frac{[(1-\alpha u)^2+1]^2}{C_2(\alpha,\lambda)} \varphi(u)\Phi(\lambda u)du$$

and the density function of $Z = W \mid \{\lambda W > X\}$ is

$$f_Z(z) = \frac{[(1-\alpha z)^2+1]^2}{C_2(\alpha,\lambda)} \varphi(z)\Phi(\lambda z)$$

that is, $W \mid \{\lambda W > X\} \sim GBASN_2(\alpha,\lambda)$.

Remark 6: One can simulate the random number from the proposed distribution using Theorem 4 and applying the acceptance-rejection technique under the following steps:

a) Generate $X \sim N(0,1)$ and $W \sim BASN_2(\alpha)$ independently.

b) If $Z = W \mid \{\lambda W > X\}$ put $Z = W$. Otherwise, go to step (a).

Theorem 5: *Let* $Z : \Omega \to R$ *be a continuous random variable and let* $q_1(z) = \frac{\varphi(\lambda z)}{\varphi(z)}[(1-\alpha z)^2+1]^{-2}$ *and* $q_2(z) = q_1(z)\,\Phi^2(\lambda z)$ *for* $z \in R$. *The random variable Z has pdf (5) if and only if the function* ξ *defined in Proposition 1 (see Appendix) is of the form*

$$\xi(z) = \frac{1}{2}\{1+\Phi^2(\lambda z)\},\ z \in R.$$

Proof: Suppose the random variable Z has pdf (5), then

$$(1-F(z))E[q_1(Z) \mid Z \ge z] = \frac{1}{2\lambda C_2(\alpha,\lambda)}\{1-\Phi^2(\lambda z)\},\ z \in R$$

and

$$(1-F(z))E[q_1(Z)\,|\,Z\ge z]=\frac{1}{4\lambda C_2(\alpha,\lambda)}\{1-\Phi^4(\lambda z)\},\ z\in R$$

Further,

$$\xi(z)q_1(z)-q_2(z)=\frac{q_1(z)}{2}\{1-\Phi^2(\lambda z)\}>0,\ \text{for } z\in R.$$

Conversely, if ξ is of the form, then

$$s'(z)=\frac{\xi'(z)q_1(z)}{\xi(z)q_1(z)-q_2(z)}=\frac{2\lambda\varphi(\lambda z)\Phi(\lambda z)}{1-\Phi^2(\lambda z)},\ z\in R,$$

and consequently

$$s(z)=-\log(1-\Phi^2(\lambda z)),\ z\in R.$$

Now, according to Proposition 1, Z has density (5).

Corollary 1: *Let* $Z:\Omega\to R$ *be a continuous random variable and let* $q_1(z)$ *be as in Theorem 4. The random variable Z has pdf (4) if and only if there exist function* q_2 *and* ξ *defined in Theorem 4 satisfying the following differential equation*

$$\frac{\xi'(z)q_1(z)}{\xi(z)q_1(z)-q_2(z)}=\frac{2\lambda\varphi(\lambda z)\Phi(\lambda z)}{1-\Phi^2(\lambda z)},\ z\in R.$$

Corollary 2: *The general solution of the differential equation in Corollary 1 is*

$$\xi(z)=\{1-\Phi^2(\lambda z)\}^{-1}\left[-\int 2\lambda\,\varphi(\lambda z)\,\Phi(\lambda z)(q_1(z))^{-1}q_2(z)dz+D\right],$$

where D is a constant.

We like to point out that one set of functions satisfying the above differential equation is given in Theorem 4 with $D=\frac{1}{2}$. Clearly, there are two other triplets (q_1,q_2,ξ) which satisfy conditions of Proposition 1.

5. Parameter Estimation and Applications

5.1 *Location Scale Extension*

The location and scale extension of $GBASN_2(\alpha,\lambda)$ distribution is derived as follows. If $Z\sim GBASN_2(\alpha,\lambda)$ then $Y=\mu+\sigma Z$ is said to be the location (μ) and scale (σ) extension of Z and has the pdf given by

$$f_y(y;\mu,\sigma,\alpha,\lambda)=\frac{1}{C(\alpha,\lambda)}\left[\left\{1-\alpha\left(\frac{y-\mu}{\sigma}\right)\right\}^2+1\right]^2\varphi\left(\frac{y-\mu}{\sigma}\right)\Phi\left[\lambda\left(\frac{y-\mu}{\sigma}\right)\right] \tag{12}$$

where, $\mu\in R$, $\alpha\in R$, $\alpha\in R$ and $\sigma>0$. Symbolically it is denoted by $Y\sim GBASN_2(\mu,\sigma,\alpha,\lambda)$.

5.2 *Maximum Likelihood Estimation*

Let $y_1,y_2,\ldots,y_n$ be a random sample from the distribution of the random variable $Y\sim GBASN_2(\mu,\sigma,\alpha,\lambda)$, so that the log-likelihood function for $\theta=(\mu,\sigma,\alpha,\lambda)$ is given by

$$\begin{aligned} l(\theta)&=2\sum_{i=1}^{n}\log\left[\left\{1-\alpha\left(\frac{y_i-\mu}{\sigma}\right)\right\}^2+1\right]-n\log C_2(\alpha,\lambda)-n\log(\sigma)-\frac{n}{2}\log(2\pi)\\ &\quad-\sum_{i=1}^{n}\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2-\sum_{i=1}^{n}\log\left[\Phi\left(\frac{\lambda(y_i-\mu)}{\sigma}\right)\right] \end{aligned} \tag{13}$$

Taking Partial derivatives of eqn. (14) w.r.t. the parameters $\theta=(\mu,\sigma,\alpha,\lambda)$, the following normal equations are obtained:

$$\frac{\partial l(\theta)}{\partial\mu}=\sum_{i=1}^{n}\frac{(y_i-\mu)}{\sigma^2}+2\sum_{i=1}^{n}\frac{2\alpha D}{\sigma(1+D^2)}+\frac{2\lambda}{\sigma}\sum_{i=1}^{n}W\left(\frac{\lambda(y_i-\mu)}{\sigma}\right)$$

$$\frac{\partial l(\theta)}{\partial\sigma}=-\frac{n}{\sigma}+\sum_{i=1}^{n}\left(\frac{(y_i-\mu)^2}{\sigma^3}\right)+2\sum_{i=1}^{n}\frac{2\alpha(y_i-\mu)D}{\sigma^2(1+D^2)}+\frac{2\lambda}{\sigma^2}\sum_{i=1}^{n}(y_i-\mu)W\left(\frac{\lambda(y_i-\mu)}{\sigma}\right)$$

$$\frac{\partial l(\theta)}{\partial\alpha}=-\frac{n\left\{8\alpha+6\alpha^3-b\delta\left(4+\frac{6\alpha^2(3+2\lambda^2)}{1+\lambda^2}\right)\right\}}{C_2(\alpha,\lambda)}-\frac{4}{\sigma}\sum_{i=1}^{n}\frac{(y_i-\mu)D}{(1+D^2)}$$

$$\frac{\partial l(\theta)}{\partial\lambda}=\frac{n\left(\frac{2b\alpha(2+3\alpha^2+2\lambda^2)}{(1+\lambda^2)^{5/2}}\right)}{C_2(\alpha,\lambda)}-\frac{2}{\sigma}\sum_{i=1}^{n}(y_i-\mu)W\left(\frac{\lambda(y_i-\mu)}{\sigma}\right)$$

where $D=1-\frac{\alpha(y_i-\mu)}{\sigma}$ and $W(.)=\frac{\varphi(.)}{\Phi(.)}$.

The solutions of the above equations give the maximum likelihood estimators, which can be found using data and numerical computation done in GenSA package in R programming language.

5.3 *Simulation for assessing MLE*

By applying the acceptance-rejection technique (see Remark 6), a simulation study is conducted by taking sample sizes $n=100$, 300 and 500 with different combinations of the true values of the scale and location parameters μ and σ for fixed values $\alpha(=-1,0,1)$ and $\lambda(=-2,-1,0,1,2)$ and taking number of replications $(r) = 1000$. For each sample size we have computed MLEs using GenSA package in R and then $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ and $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$ where θ is the true value of the parameter.

From the simulation results presented in Tables 1, 2, and 3, it is observed that the estimated Bias and MSE for different choices of parameter values gradually decreases to zero with increase in sample size as expected.

5.4 *Real Life Applications*

Here we have considered two datasets which are related to N latitude degrees in 69 samples from world lakes, which appear in Column 5 of the Diversity data set in website: http://users.stat.umn.edu/sandy/courses/8061/datasets/lakes.lsp and the body mass index (BMI) of 202 Australian athletes (Cook and Weisberg, 1994).

Table 1

Results of Simulation

α	λ	n	$\mu = 0$		$\sigma = 1$		α		λ	
			Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
-1	-2	100	0.0064	0.0059	0.0100	0.0111	0.0202	0.0138	0.029996	0.0159
		300	-0.0069	0.0061	-0.0094	0.0089	-0.0173	0.0126	-0.0171	0.0126
		500	-0.0040	0.0051	-0.0081	0.0080	0.0107	0.0096	-0.0098	0.0099
	-1	100	0.0102	0.0093	0.0096	0.0081	0.0194	0.0147	-0.0152	0.0123
		300	0.0093	0.0084	-0.0120	0.0071	-0.0111	0.0100	0.0150	0.0097
		500	-0.0081	0.0067	-0.0091	0.0080	0.0043	0.0097	-0.0121	0.0090
	0	100	-0.0061	0.0089	0.0101	0.0123	-0.0087	0.0095	0.0099	0.0127
		300	0.0053	0.0070	-0.0085	0.0082	0.0087	0.0074	-0.0098	0.0096
		500	0.0042	0.0051	-0.0080	0.0080	0.0080	0.0071	-0.0081	0.0089
	1	100	0.0121	0.0113	0.0141	0.0104	0.0157	0.0138	-0.0142	0.0133
		300	-0.0103	0.094	-0.0099	0.0083	-0.0150	0.0096	0.0111	0.0099
		500	0.0060	0.0083	0.0089	0.0083	-0.0120	0.0096	0.0097	0.0092
	2	100	0.0246	0.0157	0.0173	0.0122	0.0274	0.0151	0.0202	0.0167
		300	-0.0153	0.0141	0.0130	0.0090	-0.0210	0.0123	-0.0193	0.0134
		500	0.0110	0.0106	-0.0094	0.0090	0.0173	0.0099	0.0106	0.0101

Contd...

0	-2	100	-0.0183	0.0150	0.0201	0.0147	0.0141	0.0129	0.0170	0.0113
		300	-0.0191	0.0094	0.0173	0.0100	-0.0107	0.0111	0.0164	0.0150
		500	0.0096	0.0090	-0.0104	0.0095	-0.0093	0.0090	0.0112	0.0099
	-1	100	-0.0143	0.0101	-0.0170	0.0113	0.0119	0.0110	0.0345	0.0456
		300	0.0042	0.0061	0.0090	0.0101	-0.0112	0.0096	-0.0217	0.0334
		500	-0.0081	0.0053	-0.0091	0.0088	-0.0099	0.0090	0.0200	0.0217
	0	100	0.0092	0.0121	-0.0099	0.0081	0.0094	0.0147	0.0347	0.0310
		300	-0.0072	0.0061	-0.0095	0.0078	0.0094	0.0097	-0.0210	0.0241
		500	0.0041	0.0057	-0.0084	0.0062	0.0057	0.0069	-0.0156	0.0176
	1	100	0.0093	0.0114	0.0101	0.0120	0.0173	0.0196	0.0241	0.0190
		300	-0.0100	0.0106	-0.0097	0.0084	-0.0156	0.0153	0.0210	0.0200
		500	-0.0071	0.0092	0.0090	0.0081	0.0131	0.0101	-0.0190	0.0099
	2	100	0.0110	0.0099	0.0170	0.0134	0.0314	0.0222	0.0143	0.0107
		300	0.0096	0.0099	-0.0119	0.0101	-0.0202	0.0312	-0.0131	0.0110
		500	-0.0081	0.0080	0.0104	0.0096	-0.0159	0.0146	0.0094	0.0091
1	-2	100	-0.0210	0.0142	0.0139	0.0110	-0.0169	0.0151	0.0224	0.0170
		300	-0.0161	0.0131	-0.0127	0.0103	-0.0117	0.0111	-0.0183	0.0143
		500	0.0099	0.0090	0.0099	0.0093	0.0090	0.0112	-0.0113	0.0111
	-1	100	0.0132	0.0119	0.0111	0.0097	0.0262	0.0221	-0.0324	0.0142
		300	-0.0112	0.0121	0.0098	0.0081	-0.0204	0.0170	-0.0214	0.0111
		500	-0.0100	0.0099	0.0090	0.0072	-0.0116	0.0095	0.0099	0.0134
	0	100	-0.0101	0.0127	-0.0093	0.0092	0.0159	0.0122	0.0194	0.0138
		300	-0.0091	0.0080	-0.0083	0.0076	-0.0099	0.0106	-0.0153	0.0142
		500	0.0084	0.0081	0.0074	0.0075	-0.0121	0.0095	0.0111	0.0100
	1	100	-0.0192	0.0140	0.0126	0.0101	0.0217	0.0177	0.0200	0.0174
		300	0.0129	0.0096	-0.0099	0.0090	-0.0140	0.0130	0.0138	0.0117
		500	0.0099	0.0096	-0.0095	0.0090	-0.0112	0.0102	-0.0107	0.0099
	2	100	-0.0241	0.0150	0.0199	0.0140	0.0194	0.0201	0.0202	0.0190
		300	-0.0173	0.0162	-0.0120	0.0110	-0.0107	0.0147	0.0178	0.0143
		500	0.0108	0.0121	0.0093	0.0097	0.0101	0.0098	-0.0130	0.0141

Table 2

Results of Simulation

α	λ	n	$\mu = -2$		$\sigma = 2$		α		λ	
			Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
-1	-2	100	-0.0423	0.0340	0.0493	0.0445	0.0479	0.0339	0.0339	0.0229
		300	-0.0341	0.0210	-0.0331	0.0245	-0.0455	0.0400	-0.0210	0.0174
		500	-0.0156	0.0110	0.0246	0.0167	0.0313	0.0197	-0.0176	0.0140
	-1	100	0.0355	0.0222	0.0341	0.0221	0.0559	0.0334	-0.0476	0.0347
		300	-0.0277	0.0156	0.0256	0.0143	-0.0476	0.0276	0.0241	0.0214
		500	-0.0122	0.0096	0.0211	0.0102	-0.0321	0.0176	0.0173	0.0176
	0	100	0.0111	0.0096	0.0173	0.0124	0.0241	0.0187	0.0311	0.0222
		300	-0.0096	0.0081	-0.0103	0.0121	-0.0202	0.0191	-0.0290	0.0200
		500	0.0091	0.0064	-0.0096	0.0117	0.0193	0.0101	-0.0190	0.0164
	1	100	0.0376	0.0353	-0.0355	0.0231	0.0553	0.0444	0.0657	0.0556
		300	-0.0251	0.0210	0.0246	0.0120	-0.0496	0.0356	0.0491	0.0334
		500	-0.0121	0.0111	0.0114	0.0099	-0.0217	0.0210	0.0231	0.0141
	2	100	-0.0556	0.0341	0.0310	0.0210	0.0595	0.0463	0.0496	0.0389
		300	0.0412	0.0261	-0.0301	0.0143	-0.0440	0.0407	-0.0553	0.0421
		500	-0.0243	0.0153	-0.0210	0.0096	-0.0210	0.0160	0.0335	0.0216
0	-2	100	0.0151	0.0164	0.0423	0.0397	0.0414	0.0333	0.0656	0.0446
		300	-0.0162	0.0137	-0.0190	0.0244	-0.0392	0.0300	-0.0429	0.0321
		500	-0.0093	0.0110	-0.0100	0.0097	0.0210	0.0120	-0.0312	0.0172
	-1	100	0.0202	0.0173	0.0340	0.0257	0.0321	0.0221	0.0493	0.0319
		300	-0.0107	0.0134	-0.0213	0.0200	-0.0222	0.0192	-0.0403	0.0346
		500	-0.0046	0.0121	0.0130	0.0147	0.0130	0.0110	0.0222	0.0170
	0	100	0.0201	0.0223	0.0323	0.0221	0.0298	0.0341	0.0443	0.0396
		300	-0.0109	0.0193	-0.0177	0.0142	0.0210	0.0213	-0.0410	0.0321
		500	-0.0097	0.0112	-0.0121	0.0099	0.0110	0.0099	0.0228	0.0160
	1	100	0.0341	0.0223	0.0393	0.0245	0.0404	0.0343	-0.0513	0.0441
		300	-0.0268	0.0202	-0.0300	0.0242	-0.0411	0.0219	0.0406	0.0367
		500	-0.0187	0.0146	-0.0202	0.0130	-0.0141	0.0105	0.0310	0.0214
	2	100	0.0347	0.0218	0.0443	0.0346	0.0550	0.0417	0.0555	0.0431
		300	-0.0216	0.0145	-0.0360	0.0290	-0.0431	0.0382	-0.0432	0.0288
		500	0.0172	0.0117	-0.0220	0.0151	-0.0342	0.0213	-0.0310	0.0264

Contd...

1	-2	100	0.0555	0.0431	0.0444	0.0396	0.0499	0.0389	-0.0626	0.0521
		300	-0.0430	0.0393	-0.0421	0.0303	-0.0398	0.0288	-0.0437	0.0483
		500	-0.0217	0.0107	-0.0321	0.0112	-0.0218	0.0193	-0.0331	0.0201
	-1	100	0.0479	0.0334	0.0662	0.0400	0.0423	0.0314	0.0550	0.0432
		300	-0.0407	0.0299	-0.0431	0.0326	-0.0401	0.0353	-0.0321	0.0320
		500	-0.0209	0.0126	-0.0327	0.0201	-0.0218	0.0217	-0.0210	0.0180
	0	100	-0.0348	0.0298	0.0491	0.0362	0.0464	0.0334	0.0436	0.0366
		300	0.0223	0.0170	-0.0330	0.0309	-0.0394	0.0309	-0.0315	0.0218
		500	0.0193	0.0109	-0.0243	0.0192	0.0210	0.0187	-0.0282	0.0201
	1	100	0.0424	0.0316	0.0534	0.0440	0.0421	0.0389	-0.0431	0.0324
		300	-0.0331	0.0310	-0.0474	0.0317	0.0316	0.0288	-0.0359	0.0309
		500	0.0210	0.0119	-0.0217	0.0194	-0.0300	0.0121	0.0200	0.0186
	2	100	0.0529	0.0424	0.0535	0.0443	0.0667	0.0550	0.0396	0.0248
		300	-0.0449	0.0330	-0.0493	0.0340	0.0521	0.0500	-0.0207	0.0219
		500	-0.0210	0.0192	0.0330	0.0295	0.0211	0.0121	0.0149	0.0180

Table 3

Results of Simulation

α	λ	n	$\mu = 2$		$\sigma = 3$		α		λ	
			Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
-1	-2	100	0.0513	0.0347	0.0610	0.0411	0.0209	0.0342	0.0396	0.0298
		300	-0.0340	0.0219	0.0321	-0.0290	0.0341	0.0291	-0.0276	0.0160
		500	-0.0209	0.0206	-0.0232	0.0211	-0.0200	0.0206	0.0211	0.0122
	-1	100	0.0474	0.0330	0.0433	0.0319	-0.0606	0.0414	0.0447	0.0333
		300	0.0394	0.0241	0.0370	-0.0364	0.0513	0.0310	-0.0460	0.0209
		500	-0.0220	0.0187	0.0217	0.0219	0.0251	0.0192	-0.0293	0.0181
	0	100	0.0359	0.0443	0.0556	0.0443	0.0555	0.0420	0.0500	0.0391
		300	-0.0214	0.0361	-0.0476	0.0349	-0.0321	0.0330	0.0427	0.0332
		500	0.0149	0.0238	0.0343	0.0221	0.0302	0.0214	-0.0331	0.0217
	1	100	0.0476	0.0331	0.0417	0.0339	0.0656	0.0550	0.0505	0.0340
		300	-0.0534	0.0281	0.0421	0.0341	-0.0521	0.0431	-0.0493	0.0413
		500	-0.0343	0.0275	-0.0326	0.0210	0.0431	0.0300	0.0281	0.0201
	2	100	0.0497	0.0330	0.0626	0.0416	0.0596	0.0437	0.0673	0.0516
		300	-0.0400	0.0279	-0.0453	0.0341	-0.0433	0.0303	-0.0595	0.0464
		500	-0.0292	0.0146	-0.0370	0.0219	0.0337	0.0220	-0.0340	0.0201

Contd...

0	-2	100	0.0439	0.0341	0.0591	0.0444	0.0494	0.0339	0.0476	0.0391
		300	0.0447	0.0395	-0.0454	0.0372	-0.0334	0.0337	0.0410	0.0330
		500	-0.0210	0.0231	-0.0324	0.0206	0.0316	0.0172	0.0343	0.0216
	-1	100	0.0373	0.0221	0.0318	0.0299	0.0459	0.0401	0.0496	0.0337
		300	-0.0248	0.0210	-0.0301	0.0200	-0.0333	0.0307	-0.0431	0.0356
		500	-0.0191	0.0120	-0.0214	0.0163	0.0208	0.0153	0.0298	0.0298
	0	100	0.0159	0.0107	0.0251	0.0200	0.0341	0.0260	0.0551	0.0443
		300	-0.0107	0.0106	-0.0163	0.0143	-0.0269	0.0263	0.0434	0.0371
		500	0.0098	0.0099	-0.0107	0.0101	-0.0159	0.0100	0.0218	0.0153
	1	100	0.0246	0.0180	0.0223	0.0301	0.0331	0.0222	0.0538	0.0444
		300	-0.0188	0.0163	-0.0167	0.0222	0.0310	0.0261	0.0439	0.0301
		500	0.0172	0.0160	0.0121	0.0119	0.0287	0.0190	0.0227	0.0140
	2	100	0.0176	0.0217	0.0319	0.0202	-0.0490	0.0396	0.0463	0.0349
		300	0.0214	0.0143	-0.0261	0.0204	-0.0321	0.0264	-0.0347	0.0246
		500	-0.0153	0.0106	0.0200	0.0159	0.0172	0.0157	-0.0210	0.0119
1	-2	100	0.0651	0.0410	0.0656	0.0551	0.0660	0.0521	0.0596	0.0419
		300	-0.0514	0.0343	-0.0497	0.0347	0.0453	0.0436	-0.0443	0.0340
		500	0.0227	0.0159	0.0310	0.0281	-0.0221	0.0310	-0.0276	0.0262
	-1	100	0.0556	0.0576	0.0490	0.0341	-0.0560	0.0463	-0.0499	0.0520
		300	-0.0413	0.0420	-0.0444	0.0229	0.0431	0.0322	0.0378	0.0248
		500	0.0300	0.0152	0.0249	0.0220	0.0363	0.0238	0.0210	0.0221
	0	100	0.0444	0.0363	0.0406	0.0399	0.0494	0.0423	0.0532	0.0427
		300	-0.0303	0.0277	0.0339	0.0257	-0.0400	0.0322	-0.0437	0.0400
		500	0.0191	0.0201	-0.0240	0.0210	-0.0372	0.0260	0.0311	0.0232
	1	100	0.0539	0.0470	0.0533	0.0421	0.0548	0.0498	0.0666	0.0536
		300	-0.0476	0.0360	-0.0407	0.0332	-0.0503	0.0338	-0.0500	0.0420
		500	-0.0293	0.0196	-0.0313	0.0281	0.0363	0.0307	0.0410	0.0234
	2	100	-0.0667	0.0538	0.0673	0.0552	0.0565	0.0443	0.0550	0.0421
		300	0.0543	0.0419	-0.0521	0.0392	0.0452	0.0320	-0.0333	0.0392
		500	0.0416	0.0300	0.0344	0.0228	0.0334	0.0221	-0.0203	0.0210

Now, we compare our proposed distribution $GBASN_2(\mu,\sigma,\alpha,\lambda)$ with the normal distribution, $N(\mu,\sigma^2)$, the logistic distribution, $LG(\mu,\beta)$, Laplace distribution, $La(\mu,\beta)$, the skew-normal distribution, $SN(\mu,\sigma,\lambda)$ of Azzalini (1985), the skew-logistic distribution, $SLG(\mu,\beta,\lambda)$ of Wahed and Ali (2001), the skew-Laplace distribution, $SLa(\mu,\beta,\lambda)$ of Nekoukhou and Alamatsaz (2012), the alpha-skew-normal distribution, $ASN(\mu,\sigma,\alpha)$ of Elal-Olivero (2010), the alpha-skew-Laplace distribution,

Table 4

MLE's, log-likelihood, AIC and BIC for N latitude degrees in 69 samples from world lakes.

Distributions	μ	σ	λ	α	β	$\log L$	AIC	BIC
$N(\mu,\sigma^2)$	45.165	9.549	--	--	--	-253.599	511.198	515.666
$LG(\mu,\beta)$	43.639	--	--	--	4.493	-246.65	497.290	501.758
$SN(\mu,\sigma,\lambda)$	35.344	13.7	3.687	--	--	-243.04	492.072	498.774
$BSN(\mu,\sigma,\beta)$	54.47	5.52	--	--	0.74	-242.53	491.060	497.760
$SLG(\mu,\beta,\lambda)$	36.787	--	2.828	--	6.417	-239.05	490.808	490.808
$La(\mu,\beta)$	43.0	--	--	--	5.895	-239.25	482.496	486.964
$ASLG(\mu,\alpha,\beta)$	49.087	--	--	0.861	3.449	-237.35	480.702	487.404
$SLa(\mu,\beta,\lambda)$	42.3	--	0.255	--	5.943	-236.90	479.799	486.501
$ASLa(\mu,\alpha,\beta)$	42.3	--	--	-0.22	5.44	-236.08	478.159	484.861
$ASN(\mu,\sigma,\alpha)$	52.147	7.714	--	2.042	--	-235.37	476.739	483.441
$ABSN(\mu,\sigma,\alpha,\beta)$	53.28	9.772	--	2.943	-0.292	-234.36	476.719	485.655
$GASN(\mu,\sigma,\alpha,\lambda)$	56.319	8.544	-0.672	12.052	--	-230.531	469.062	477.998
$GBASN_2(\mu,\sigma,\alpha,\lambda)$	55.997	6.664	-0.314	3.877	--	-225.851	459.702	**468.638**

$ASLa(\mu,\alpha,\beta)$ of Harandi and Alamatsaz (2013), the alpha-skew-logistic distribution, $ASLG(\mu,\alpha,\beta)$ of Hazarika and Chakraborty (2014), the alpha-beta-skew-normal distribution, $ABSN(\mu,\sigma,\alpha,\beta)$ and beta-skew-normal $BSN(\mu,\sigma,\beta)$ distribution of Shafiei *et al.* (2016), and the generalized alpha-skew-normal distribution, $GASN(\mu,\sigma,\alpha,\lambda)$ of Sharafi *et al.* (2017). The MLE of the parameters are obtained by using numerical optimization routine. AIC and BIC are used for model comparison.

From Tables 4 and 5, it is observed that the proposed generalized Balakrishnan - alpha - skew - normal $GBASN_2(\mu,\sigma,\alpha,\lambda)$ distribution provides best fit to the dataset under consideration in terms of all criteria, namely the log-likelihood, the AIC as well as the BIC. The plots of observed (in histogram) and expected densities (lines) considering only the top three competitors of the proposed model presented in Figure, 5 and 6, also confirm our findings.

Table 5

MLE's, log-likelihood, AIC and BIC for body mass index (BMI) of 202 Australian athletes.

Distributions	μ	σ	λ	α	β	$\log L$	AIC	BIC
$N(\mu,\sigma^2)$	22.956	2.857	--	--	--	-498.668	1001.336	1007.953
$La(\mu,\beta)$	22.749	--	--	--	2.123	-494.08	992.160	998.7765
$BSN(\mu,\sigma,\beta)$	22.528	2.694	--	--	-0.058	-492.88	991.760	1001.685
$ASLa(\mu,\alpha,\beta)$	22.350	--	--	-0.14	2.07	-492.601	991.202	1001.127
$SLa(\mu,\beta,\lambda)$	22.350	--	0.865	--	2.084	-492.461	990.922	1000.847
$LG(\mu,\beta)$	22.787	--	--	--	1.529	-491.462	986.924	993.5405
$SN(\mu,\sigma,\lambda)$	19.969	4.133	2.313	--	--	-490.099	986.198	996.1228
$ASLG(\mu,\alpha,\beta)$	21.933	--	--	-0.201	1.475	-489.094	984.188	994.1128
$ASN(\mu,\sigma,\alpha)$	24.834	2.653	--	0.994	--	-488.69	983.380	993.3048
$ABSN(\mu,\sigma,\alpha,\beta)$	23.998	2.853	--	0.817	-0.131	-486.743	981.486	994.7191
$SLG(\mu,\beta,\lambda)$	20.717	--	1.401	--	1.975	-487.311	980.622	990.5468
$GASN(\mu,\sigma,\alpha,\lambda)$	26.879	3.109	-0.667	2.066	--	-485.788	979.576	992.8091
$GBASN_2(\mu,\sigma,\alpha,\lambda)$	26.022	2.686	0.149	0.905	--	-484.574	977.148	**990.3811**

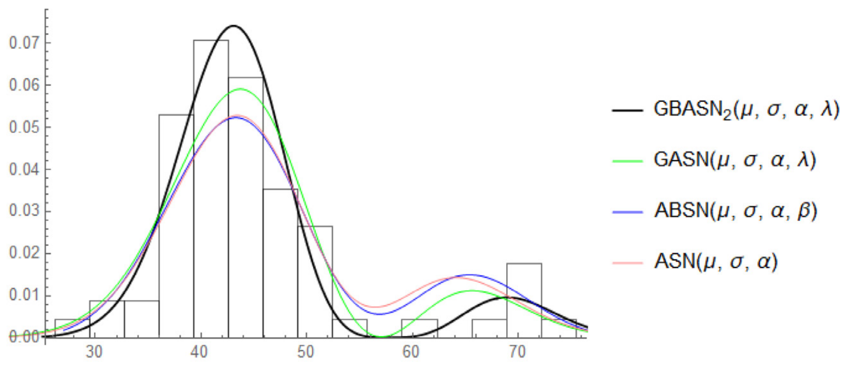


Figure 5

Plots of observed and expected densities of some distributions for N latitude degrees in 69 samples from world lakes.

5.4.1 Likelihood Ratio test

Since $N(\mu,\sigma^2)$, $SN(\lambda,\mu,\sigma)$ and $GBASN_2(\mu,\sigma,\alpha,\lambda)$ distributions are nested models, the likelihood ratio (LR) test is used to discriminate between them. The LR test is carried out to test the following hypothesis as shown in Table 6 below.

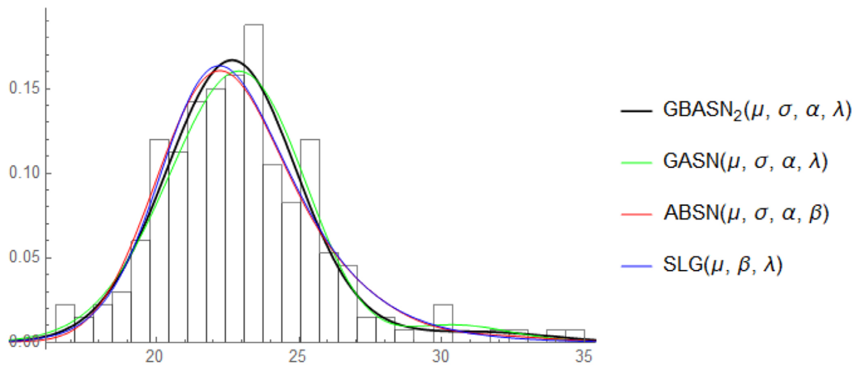


Figure 6

Plots of observed and expected densities of some distributions for Body Mass Index (BMI) of 202 Australian athletes.

Table 6

The values of LR test statistic for different hypothesis.

Hypothesis	LR test statistic values		Degrees of Freedom	Critical value
	Dataset 1	Dataset 2		
$H_0 : \alpha = 0$ vs $H_1 : \alpha \neq 0$	34.378	11.05	3	7.815
$H_0 : \alpha = \lambda = 0$ vs $H_1 : \alpha \neq 0, \lambda \neq 0$	55.496	28.188	2	5.991

The values of LR test statistic exceed the corresponding critical value at 5% level of significance for the given degrees of freedom. Thus, there is evidence in support of the alternative hypothesis. Therefore, we may conclude that the sampled data comes from $GBASN_2(\mu, \sigma, \alpha, \lambda)$ distribution and not from other distributions considered here.

6. Conclusions

In this article, a generalized-Balakrishnan-alpha-skew-normal distribution which has at most two modes, is introduced and some of its basic and distributional properties along with some extensions are studied. The numerical results of the modelling of a real life dataset considered here have shown that the proposed distribution $GBASN_2(\alpha,\lambda)$ provides better fitting, in terms of AIC and BIC values, in comparison to the other known distributions. The cdf of the $GBASN_2(\alpha,\lambda)$ is quite messy, function of the cdf of standard normal cdf. As a future work one may attempt to approximate this cdf with good accuracy using the similar approximation available for standard normal. (See Omar and Jawaher, 2021 among others).

Acknowledgment

Authors would like to acknowledge the comments and suggestions of the reviewers and the editor towards the improvement of the paper.

Appendix

Proposition 1: Let (Ω, F, P) be a given probability space and let $H=[\mathrm{a},\mathrm{b}]$ be an interval for some $d<b$ $(a=-\infty,\ \ b=\infty$ might as well be allowed). Let $Z:\Omega\to H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions defined on *H* such that

$$\mathrm{E}[\mathrm{q}_2(\mathrm{Z})|\mathrm{Z}\geq \mathrm{z}]=\mathrm{E}[\mathrm{q}_1(\mathrm{Z})|\mathrm{Z}\geq \mathrm{z}]\xi(\mathrm{z}), \quad \mathrm{z}\in H,$$

is defined with some real function η. Assume that $q_1, q_2 \in C^1(H)$, $\xi\in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set *H*. Finally, assume that the equation $\xi q_1 = q_2$ has no real solution in the interior of *H*. Then *F* is uniquely determined by the functions q_1, q_2 and ξ, particularly

$$F(z)=\int_a^z C\left|\frac{\xi'(u)}{\xi(u)\mathrm{q}_1(u)-\mathrm{q}_2(u)}\right|\exp(-\mathrm{s}(u))du$$

where the function *s* is a solution of the differential equation $\mathrm{s}'=\frac{\xi' q_1}{\xi q_1 - q_2}$ and C is the normalization constant, such that $\int_H dF=1$.

The result, however, holds also when the interval *H* is not closed, since the condition of the Proposition 1 is on the interior of *H*. It should be mentioned that our characterizations do not require a closed form for the cumulative distribution function. Moreover this result is stable in the sense of weak convergence (see, Glanzel, 1990). A further consequence of the stability property of Proposition 1 is the application of this proposition to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions

q_1, q_2 and, specially, ξ should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose ξ as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics. In some cases, one can take $q_1(z)=1$, which reduces the condition of Proposition 1 to $E[q_2(Z)|Z \geq z] = \xi(z), \ \ z \in H$. In general however, employing three functions q_1, q_2 and ξ, expected to enhance the domain of applicability of Proposition 1.

References

[1] Arnold, B. C., and Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Test*, *11*(1), 7-54.

[2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*(2), 171-178.

[3] Bahrami,W., Agahi, H., and Rangin, H. (2009). A Two-Parameter Balakrishnan Skew-Normal Distribution, *J. Statist. Res. Iran*, 6, 231–242.

[4] Balakrishnan, N. (2002). Discussion on "Skew multivariate models related to hidden truncation and/or selective reporting" by B. C. Arnold and R. J. Beaver. *Test 11*, 37-39

[5] Chakraborty, S., Hazarika, P. J., and Ali, M. M. (2015). A multimodal skewed extension of normal distribution: its properties and applications. *Statistics*, *49*(4), 859-877.

[6] Cook, R.D., and Weisberg, S. (1994). *An Introduction to Regression Analysis*, Wiley, New York.

[7] Elal-Olivero, D. (2010). Alpha-skew-normal distribution. *Proyecciones (Antofagasta)*, *29* (3), 224-240.Wahed, A., & Ali, M. M. (2001). The skew-logistic distribution. *J. Statist. Res*, *35*(2), 71-80.

[8] Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families. *Mathematical Statistics and Probability Theory* (pp. 75-84). Springer, Dordrecht.

[9] Glanzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics*, *21*(4), 613-618.

[10] Gupta, R.C., and Gupta, R.D. (2004). Generalized Skew Normal Model, *Test, 13*, 501-524.

[11] Hamedani, G. G., Korkmaz, M. Ç., & Yousof, H. M. (2021). The type I quasi lambert family: properties, characterizations and different estimation methods. P*akistan Journal of Statistics and Operation Research*, *17* (3), 545-558.

[12] Harandi, S. S., and Alamatsaz, M. H. (2013). Alpha–Skew–Laplace distribution. *Statistics and Probability Letters*, *83*(3), 774-782.

[13] Hasanalipour, P., and Sharafi, M. (2012). A New Generalized Balakrishnan Skew Normal Distribution, *Stat Papers*, *53*, 219-228.

[14] Hazarika, P. J., and Chakraborty, S. (2014). Alpha-Skew-Logistic Distribution. *IOSR Journal of Mathematics*, *10*(4), 36-46.

[15] Hazarika P. J., Shah, S. and Chakraborty, S. (2020). Balakrishnan Alpha Skew Normal Distribution: Properties and Applications. *Malaysian Journal of Science*, *39* (2), 71-91.

[16] Henze, N. (1986). A Probabilistic Representation of the Skew Normal Distribution, *Scand. J. Stat*., *13*, 271–275.

[17] Huang, W. J., and Chen, Y. H. (2007). Generalized skew-Cauchy distribution. *Statistics and Probability Letters*, *77*(11), 1137-1147.

[18] Jamalizadeheh, A., Behbooddian, J., and J. Balakrrishnan, N (2008). A Two Parameter Generalized Skew Normal Distribution, *Stat. Prob. Lett*., *78*, 1722-1726.

[19] Ma, Y., and Genton, M.G. (2004). Flexible Class of Skew Symmetric Distributions, *Scand J Statist*., *31*, 459–468.

[20] Louzada, F., Ara, A. and Fernandes, G. (2017). The bivariate alpha-skew-normal distribution. *Communications in Statistics-Theory and Methods*, *46*(14), 7147-7156.

[21] Nekoukhou, V., and Alamatsaz, M.H., (2012), A family of skew-symmetric-Laplace distributions. *Statistical Papers*, *53*(3), 685-696.

[22] Omar E., and Jawaher A. (2021) An accurate approximation for the standard normal distribution function, *Journal of Information and Optimization Sciences*, *42* (1), 17-27.

[23] Shafiei, S., Doostparast, M., and Jamalizadeh, A. (2016). The alpha–beta skew normal distribution: Properties and applications. *Statistics*, *50*(2), 338-349.

[24] Shah, S., Chakraborty, S., and Hazarika, P. J. (2020a). The Balakrishnan Alpha Skew Logistic Distribution: Properties and Applications. *International Journal of Applied Mathematics and Statistics*, *59*(1), 76-92.

[25] Shah, S., Hazarika, P. J., and Chakraborty, S. (2020b). A New Alpha Skew Laplace Distribution: Properties and Its Applications. *International Journal of Agricultural and Statistical Sciences, 16*(1), 1-10.

[26] Shah, S., Hazarika, P. J., and Chakraborty, S. (2020c). The Balakrishnan Alpha Skew Truncated Cauchy Distribution with Applications in Modelling Currency Exchange Rate. *Sains Malaysiana*, *49*(10), 2565-2571.

[27] Shah, S., Hazarika, P. J., Chakraborty, S., and Ali, M. M. (2020d). The Log-Balakrishnan-Alpha-Skew-Normal Distribution and Its Applications. *Pakistan Journal of Statistics and Operation Research*, *16*(1), 109-117.

[28] Shah, S., Hazarika, P. J., Chakraborty, S., and Ali, M. M. (2021). A Generalized-Alpha-Beta-Skew Normal Distribution with Applications. *Annals of Data Science*, https://doi.org/10.1007/s40745-021-00325-0.

[29] Sharafi, M., and Behboodian, J. (2008). The Balakrishnan skew–normal density. *Statistical Papers*, *49*(4), 769-778.

[30] Sharafi, M., Sajjadnia, Z., and Behboodian, J. (2017). A new generalization of alpha-skew-normal distribution. *Communications in Statistics-Theory and Methods*, *46*(12), 6098-6111.

[31] Venegas, O., Bolfarine, H., Gallardo, D. I., Vergara-Fernández, A., and Gómez, H. W. (2016). A Note on the Log Alpha Skew Normal Model with Geochemical Applications. *Appl. Math*, *10*(5), 1697-1703.

[32] Wahed, A., and Ali, M. M. (2001). The skew-logistic distribution. *J. Statist. Res*, *35*(2), 71-80.

[33] Yadegari, I., Gerami, A., and Khaledi, M.J. (2008). A generalization of the Balakrishnan skew normal distribution. *Stat. Probab. Lett*. *78*: 1165-1167.

Received, April 2022

Revised, August 2022