Mathematical Forum ISSN: 0972-9852
Vol. 29, 2022

LOG-BALAKRISHNAN-ALPHA-BETA-SKEW-NORMAL DISTRIBUTION

S. SHAH[1], P. J. HAZARIKA[2] AND S. CHAKRABORTY[3]

[1,2,3] *Department of Statistics, Dibrugarh University, Dibrugarh, 786004, Assam, India*

Email: [1]charan.shah90@gmail.com, [2]parthajhazarika@gmail.com

[3]subrata_arya@yahoo.co.in

Received on: 03/03/2021
Accepted on: 14/12/2021

Abstract

In this study, a logarithmic form of the Balakrishnan alpha beta skew normal distribution of Shah et al. (2021) is proposed by implementing the methodology adopted by Shah et al. (2020b) and investigated some of its basic properties. Advantages of the proposed distribution relative to some related distributions including the log-normal are checked through data fitting experiment using maximum likelihood parameter estimation. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are adapted for model selection. Further the likelihood ratio testis used for discriminating between the log-normal and the proposed distributions. Our finding clearly indicated the superiority of the proposed model.

Keywords: Skew Distribution, Alpha Beta Skew Normal Distribution, Bimodal Distribution, Likelihood Ratio Test, AIC, BIC.

2010 AMS classification: 60E05, 62E10

1. Introduction

Since the normal distribution is commonly employed for examining the symmetric data, skew normal distribution is for studying asymmetric data but for analyzing the positive and skewed data the gamma, the Weibull and the log-normal distributions are commonly used. In fact, the log-normal model has been found quite useful for fitting chemical element concentration using many data sets present positive asymmetry as studied by Ahrens (1953, 1954a, 1954b). Vistelius (1960) showed that the chemical element concentrations in soil samples follow asymmetric distribution. Therefore, it is important to modify the skew normal distribution to deal with geochemical data as in Mateu-Figueras *et al.* (2004) where log-skew-normal distribution is used. This distribution was also used by Azzalini *et al.* (2003), for family income data.

Azzalini (1985) first introduced the skew-normal distribution by adding an additional parameter to introduce asymmetry and the probability density function (pdf) of the same is given by

$$f(z;\lambda) = 2\varphi(z)\Phi(\lambda z)\ ,\ \ z \in R \tag{1}$$

where $\lambda \in R$ is the skewness parameter, $\phi(.)$ and $\Phi(.)$ are respectively, the pdf and cumulative distribution function (cdf) of standard normal distribution. The general formula for the construction of skew-symmetric distributions was given by Huang and Chen (2007) by introducing the concept of skew function $G(.)$ and the pdf of the same is given by

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

where $G(.)$ (skew function) is the Lebesgue measurable function such that $0 \le G(z) \le 1$ and $G(z)+G(-z)=1,\ \ z \in R$, almost everywhere. Obviously, by selecting different skew functions, one can construct many numbers of skewed distributions with unimodal, bimodal and multimodal behaviors (see for example, Elal-Olivero (2010), Hazarika and Chakraborty (2014), Chakraborty *et al.* (2012, 2014, and 2015), Hazarika *et al.* (2020), Shah and Hazarika (2019), Shah *et al.* (2020a, 2020b,2020c) and among others). Shah *et al.* (2020b) introduced the Log-Balakrishnan-alpha-skew-normal $LBASN_2(\alpha)$ distribution for studying unimodal as well as bimodal data and the pdf is given by

$$f_Z(z;\alpha,\beta) = \frac{[(1-\alpha\, y)^2+1]^2}{C_2(\alpha)}\frac{\phi(y)}{z},\ \ z>0 \tag{3}$$

where $C_2(\alpha) = 4 + 8\alpha^2 + 3\alpha^4$ and $y = Log(z)$.

Shah *et al.* (2021) proposed a family of skew distributions to provide more flexibility and called it the Balakrishnan-alpha-beta-skew-normal distribution, denoted by $BABSN_2(\alpha,\beta)$ with pdf given by

$$f_Z(z;\alpha,\beta) = \frac{[(1-\alpha z - \beta z^3)^2 + 1]^2}{C_2(\alpha,\beta)}\phi(z),\ z \in R \tag{4}$$

where $C_2(\alpha)$ is defined above, $\alpha \in R$ and $\beta \in R$ are the skewness parameters and $C_2(\alpha,\beta) = C_2(\alpha) + 60\alpha^3\beta + 12\alpha\beta(4 + 315\beta^2) + 630\alpha^2\beta^2 + 15\beta^2(8 + 693\beta^2)$,

In the present article the methodology of Shah *et al.* (2021) and Shah *et al.* (2020b) has been implemented to propose a Log-Balakrishnan-alpha-beta-skew-normal distribution which is flexible enough to support both unimodal, bimodal and multimodal behaviors for positive support and investigate some of its distributional properties. To exhibit the applicability of the proposed distribution, the three real life datasets are consider which give better fitting when compared to some other known distributions. The article is organized as follows. In Section 2, the Log-Balakrishnan–alpha–beta–skew–normal distribution is defined and discusses some of its important distributional properties. The random number generation of Log–Balakrishnan–alpha–beta–skew–normal distribution is discussed in Section 3. Section 4 discusses the parameter estimation and maximum likelihood estimation of the proposed distribution. In Section 5, some numerical examples based on real life data and Likelihood ratio test are provided. Finally, the article is ended with conclusions in Section 6.

2. Log–Balakrishnan–Alpha–Beta–Skew–Normal Distribution

Here we define a new family of distribution known as Log Balakrishnan alpha beta normal distribution and give some of its basic properties.

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

$$f_Z(z;\alpha,\beta) = \frac{[(1-\alpha y - \beta y^3)^2 + 1]^2}{C_2(\alpha,\beta)}\frac{\phi(y)}{z},\ z > 0,\ y = Log(z) \tag{5}$$

then it is said to follow Log–Balakrishnan–alpha–beta–skew–normal distribution with skewness parameters $\alpha \in R$ and $\beta \in R$. In the rest of this article, we shall refer the distribution in the Eqn. (5) as $LBABSN_2(\alpha,\beta)$. Where, $\varphi(.)$ is the pdf of standard normal distribution, $C_2(\alpha)$ is defined above and

$C_2(\alpha,\beta) = C_2(\alpha) + 60\alpha^3\beta + 12\alpha\beta(4+315\beta^2) + 630\alpha^2\beta^2 + 15\beta^2(8+693\beta^2)$.

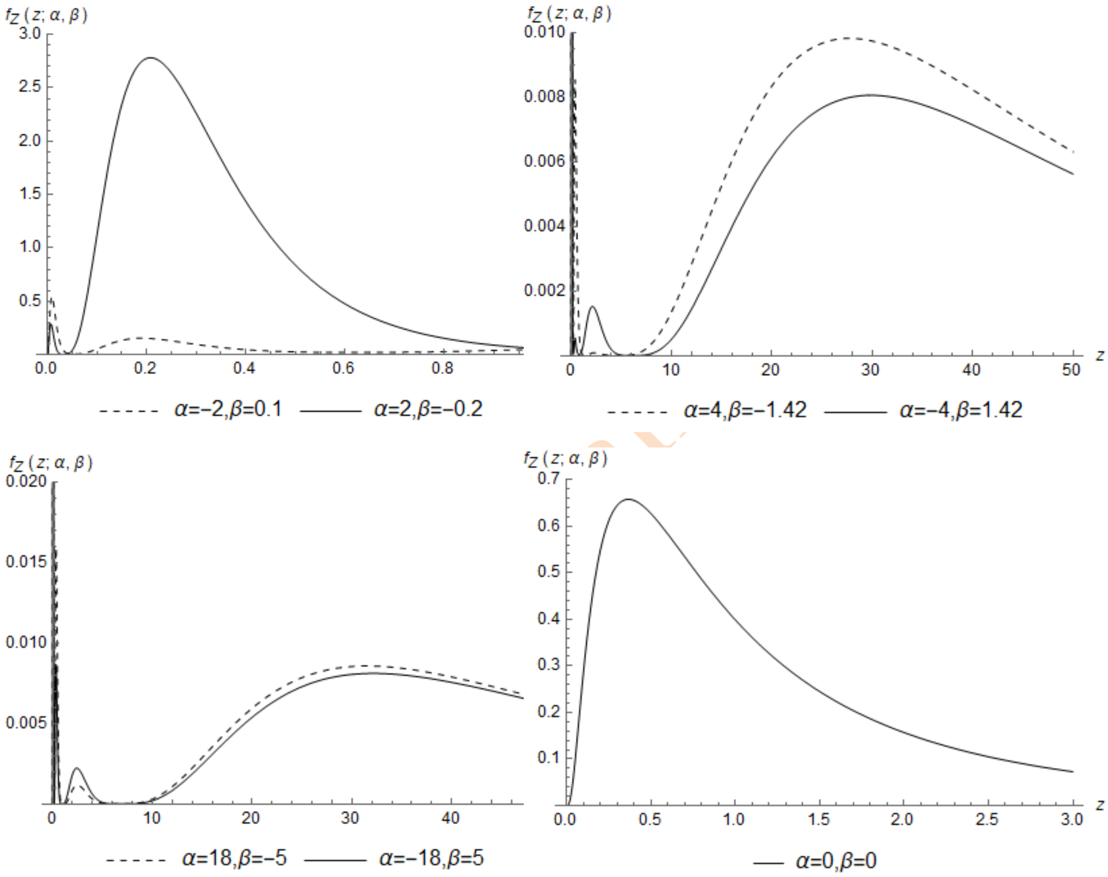

Figure 1: Plots of the pdf of $LBABSN_2(\alpha,\beta)$ distribution for different choices of the parameters α and β.

2.1. Special cases of $LBABSN_2(\alpha,\beta)$:

Some of the special cases of $LBABSN_2(\alpha,\beta)$ distribution are obtained as follows:

- When $\beta = 0$, we get the $LBASN_2(\alpha)$ distribution of Shah *et al.* (2020b) and is given by $f(z;\alpha) = \dfrac{[(1-\alpha\, y)^2 + 1]}{(4 + 8\alpha^2 + 3\alpha^4)} \dfrac{\varphi(y)}{z}$.
- If $\alpha = 0$, we get $f(z;\beta) = \dfrac{[(1-\beta\, z^3)^2 + 1]^2}{[4 + 15\beta^2(8 + 693\beta^2)]} \dfrac{\phi(y)}{z}$.

  This distribution is known as the Log–Balakrishnan–beta–skew–normal $LBBSN_2(\beta)$ distribution.
- For $\alpha = \beta = 0$, we get the standard Log-normal $LN(0,1)$ distribution and is given by $f(z) = \phi(y)/z$.
- If $\alpha \to \pm\infty$, we get the bimodal-log-normal $LBLN(4)$ distribution (see Hazarika *et al.* 2020) given by $f(z) = y^4 \varphi(y)/3z$.
- If $\beta \to \pm\infty$, we get the bimodal-log-normal $LBLN(12)$ distribution (see Hazarika *et al.* 2020) given by $f(z) = y^{12}\phi(y)/10395z$.

  Additionally, if $Z \sim LBABSN_2(\alpha,\beta)$, then $-Z \sim LBABSN_2(-\alpha,-\beta)$.

The pdfs of $LBABSN_2(\alpha,\beta)$ distribution for different choices of the parameters α and β are plotted in Figure 1, show that the proposed distribution is flexible and supports unimodal, bimodal and multimodal data and with skewness (right tailed).

2.2. Cumulative Distribution Function

Theorem 1: *The cdf of* $LBABSN_2(\alpha,\beta)$ *distribution is given by*

$$F_Z(z;\alpha,\beta) = \frac{\phi(y)}{C(\alpha,\beta)} \left[\begin{array}{l} 8(b_2+\alpha) + (4+b_1)e^{\frac{y^2}{2}}\sqrt{2\pi}\Phi(y) - b_1 y + 4b_2 y^2 - b_3 y^3 + 12\beta(\alpha^2 + \\ 6\alpha\beta + 16\beta^2)y^4 - b_4\beta y^5 + 4\beta^2(3\alpha + 8\beta)y^6 - 3\beta^2(2\alpha^2 + 12\alpha\beta + \\ 33\beta^2)y^7 + 4\beta^3 y^8 - \beta^3(4\alpha + 11\beta)y^9 - \beta^4 y^{11} \end{array} \right] \quad (6)$$

where $z > 0,\ y = Log(z)$, $\varphi(.)$, $\Phi(.)$, *and* $C_2(\alpha,\beta)$ *are defined above.*

$b_1 = 3\alpha^4 + 60\alpha^3\beta + 120\beta^2 + 10395\beta^4 + 12\alpha\beta(4 + 315\beta^2) + \alpha^2(8 + 630\beta^2)$,

$b_2 = \alpha^3 + 12\alpha^2\beta + 72\alpha\beta^2 + 2(\beta + 96\beta^3)$,

$b_3 = \alpha^4 + 20\alpha^3\beta + 210\alpha^2\beta^2 + 4\alpha\beta(4 + 315\beta^2) + 5\beta^2(8 + 693\beta^2)$, and

$b_4 = 4\alpha^3 + 8\beta + 42\alpha^2\beta + 252\alpha\beta^2 + 693\beta^3$.

Proof: see Appendix A.

The above cdf is plotted in Figure 2 to investigate variation in its shape with respect to the parameters α and β.

Corollary 1: In particular, by taking the limit $\alpha \to \pm\infty$ of $F_Z(z;\alpha,\beta)$ in Eqn. (6), we get the cdf of $LBLN(4)$ distribution as $F_Z(z) = \Phi(y) - \frac{\phi(y)[3y + y^3]}{3}$.

Again, in particular, by taking the limit $\beta \to \pm\infty$ of $F_Z(z;\alpha,\beta)$ in Eqn. (6), we get the cdf of $LBLN(12)$ distribution as

$$F_Z(z) = \Phi(y) - \frac{\varphi(y)[10395y + 3465y^3 + 693y^5 + 99y^7 + 11y^9 + y^{11}]}{10395}.$$

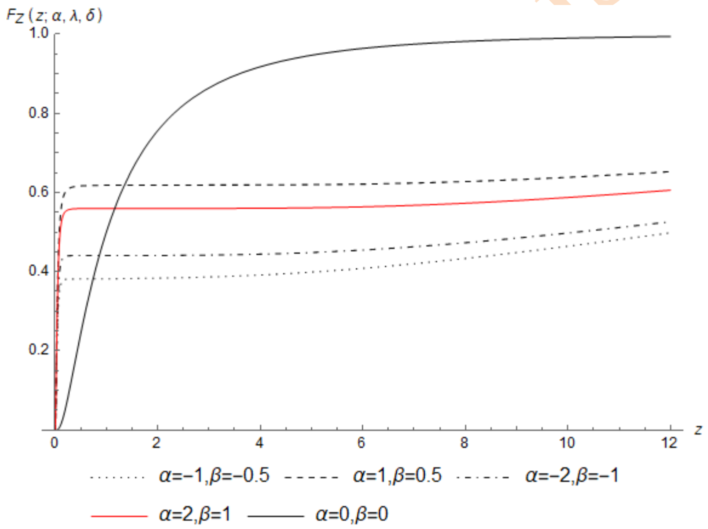

Figure 2: Plots of the cdf of $LBABSN_2(\alpha,\beta)$ distribution for different choices of the parameters α and β.

2.3. Mode

Theorem 2: *The $LBABSN_2(\alpha,\beta)$ distribution has at most four modes.*

Proof: see Appendix B.

2.4. Moments

Theorem 3: *The* n^{th} *order moment of* $LBABSN_2(\alpha,\beta)$ *distribution is given by*

$$E(Z^n)=\frac{e^{\frac{n^2}{2}}}{C_2(\alpha,\beta)}\begin{bmatrix}4-8n\alpha+8\alpha^2(1+n^2)-12n\alpha^2\beta c_1-12n\alpha\beta^2c_2+\\ 6\alpha^2\beta^2c_3-4n\beta^3c_4+4\alpha\beta^3c_5+\beta^4c_6-4n(3+n^2)(\alpha^3\\ +2\beta)+4\beta c_7(\alpha^3+2\beta)+(3+6n^2+n^4)(\alpha^4+16\alpha\beta)\end{bmatrix} \quad (7)$$

where $c_1=15+10n^2+n^4, c_2=105+105n^2+21n^4+n^6,$

$c_3=105+420n^2+210n^4+28n^6+n^8,$

$c_4=945+1260n^2+378n^4+36n^6+n^8,$

$c_5=945+4725n^2+3150n^4+630n^6+45n^8+n^{10},$

$c_6=10395+62370n^2+51975n^4+13860n^6+1485n^8+66n^{10}+n^{12}$ and

$c_7=15+45n^2+15n^4+n^6.$

Proof: see Appendix C.

From the above equation of n^{th} order moment, we get

$$E(Z)=\frac{2\sqrt{e}[2+5\alpha^4+8\alpha^3(-1+19\beta)+4d_1\alpha^2+4d_2\alpha+4d_3\beta]}{C_2(\alpha,\beta)}$$

$$E(Z^2)=\frac{e^2[4+43\alpha^4+4\alpha^3(-14+499\beta)+2d_4\alpha^2+4d_5\alpha+d_6\beta]}{C_2(\alpha,\beta)}$$

$$E(Z^3)=\frac{2e^{9/2}[69\alpha^4+24\alpha^3(-3+197\beta)+4d_7\alpha^2+12d_8\alpha+2d_9]}{C_2(\alpha,\beta)}$$

$$E(Z^4)=\frac{e^8[4+355\alpha^4+4\alpha^3(-76+8671\beta)+2d_{10}\alpha^2+4d_{11}\alpha+d_{12}\beta]}{C_2(\alpha,\beta)}$$

where $d_1=2-39\beta+573\beta^2, d_2=-1+20\beta-348\beta^2+4748\beta^3,$

$d_3=-4+76\beta-1310\beta^2+17519\beta^3, d_4=20-852\beta+21579\beta^2,$

$d_5=-4+172\beta-5550\beta^2+123109\beta^3,$

$d_6=-112+3992\beta-116744\beta^2+2430355\beta^3,$

$d_7 = 10 - 837\beta + 35901\beta^2$, $d_8 = -1 + 92\beta - 5220\beta^2 + 185364\beta^3$,

$d_9 = 1 - 72\beta + 4728\beta^2 - 227124\beta^3 + 7264350\beta^4$,

$d_{10} = 68 - 10344\beta + 722427\beta^2$,

$d_{11} = -8 + 1420\beta - 135084\beta^2 + 7461121\beta^3$, and

$d_{12} = -608 + 69368\beta - 5293840\beta^2 + 254388667\beta^3$.

Variance can be easily obtained as

$$Var(Z) = E(Z^2) - [E(Z)]^2$$

$$= \frac{e\left[\begin{array}{l} -4\{2 + 4d_2\alpha + 4d_1\alpha^2 + 5\alpha^4 + 4d_3\beta + 8\alpha^3(-1+19\beta)\}^2 + eC_2(\alpha,\beta)\{4 + 4d_5\alpha + \\ 2d_4\alpha^2 + +43\alpha^4 + d_6\beta + 4\alpha^3(-14+499\beta)\} \end{array}\right]}{[C_2(\alpha,\beta)]^2}$$

By numerically optimizing $E(Z)$ and $Var(Z)$ with respect to α and β, we get the approximate bounds for mean and variance as $0.4972 \le E(Z) \le 33.6709$ and $1.0131 \le Var(Z) \le 2736.75$ respectively. The plots of the mean and the variance are given respectively, in Figure 3 and Figure 4 to study their behaviors. These plots also verify these bounds.

Remark 1: By taking the limit $\alpha \to \pm\infty$ in the moments of $LABSN(\alpha,\beta)$ distribution, we can derive the moments of $LBN(2)$ distribution as $E(Z) \to 5.4957$, $Var(Z) \to 75.7067$.Again, by taking the limit $\beta \to \pm\infty$ in the moments of $LABSN(\alpha,\beta)$ distribution, we can derive the moments of $LBN(6)$ distribution as $E(Z) \to 22.23$, $Var(Z) \to 1233.43$.

2.5. Skewness and Kurtosis

The skewness (β_1) and kurtosis (β_2) of $LBABSN_2(\alpha,\beta)$ distribution are respectively, given by

$$\beta_1 = \frac{4[2C_2(\alpha,\beta)D_1^3 - 3e^2 D_1 D_2 + C_2(\alpha,\beta)e^4 D_3]^2}{C_2(\alpha,\beta)\sqrt{e}[-2D_1^2 + e^{3/2}D_2]^3}$$

where $D_1 = 2 + 5\alpha^4 + 8\alpha^3(-1+19\beta) + 4\alpha^2 d_1 + 4\alpha d_2 + 4\beta d_3$,

$D_2 = 4 + 43\alpha^4 + 4\alpha^3(-14 + 499\beta) + 2\alpha^2 d_4 + 4\alpha\, d_5 + \beta\, d_6$,

$D_3 = 64\alpha^4 + 24\alpha^3(-3 + 197\beta) + 4\alpha^2 d_7 + 12\alpha\, d_8 + 2\, d_9$ and

$d_i, i = 1, 2, \ldots, 9$ is defined above.

And, $\beta_2 = \dfrac{-6C_2(\alpha,\beta)\,D_1^4 + 12\,e^2 D_1^2 D_2 + e^{15/2} C_2(\alpha,\beta)\,D_4 - 16e^{9/2} D_1 D_3}{\sqrt{e}[-2\,D_1^2 + e^{3/2}\,D_2]^2}$

where $D_4 = 4 + 355\alpha^4 + 4\alpha^3(-76 + 8671\beta) + 2\alpha^2 d_{10} + 4\alpha\, d_{11} + \beta\, d_{12}$ and

$d_i, i = 1, 2, \ldots, 12$ are defined above.

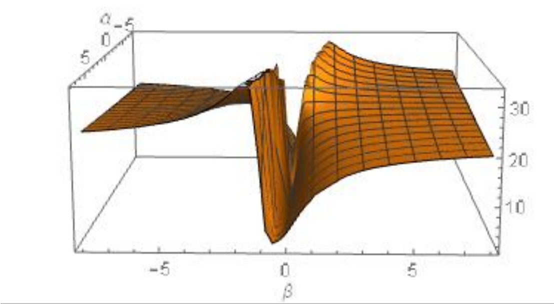

Figure 3: Plots of mean

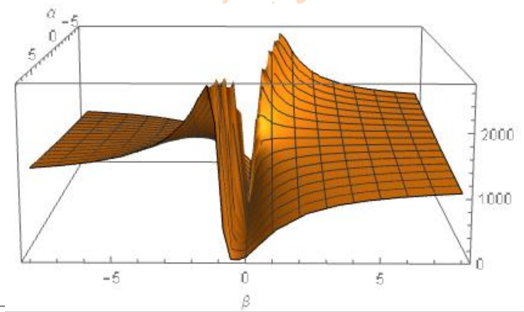

Figure 4: Plots of variance

Optimizing β_1 and β_2 numerically with respect to α and β, the approximate bounds for skewness and kurtosis are seen as $9.41 \le \beta_1 \le 14722.3$ and $25.68 \le \beta_2 \le 7428.63$. Skewness and kurtosis are plotted in Figure 5 and Figure 6 respectively to study their behaviors and also verify these bounds.

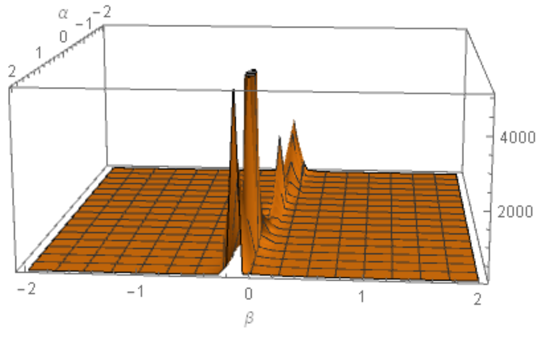

Figure 5: Plots of skewness

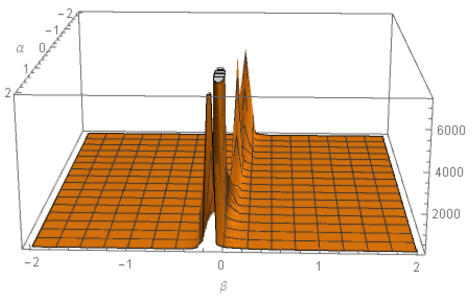

Figure 6: Plots of kurtosis

Remark 2: In limit $\alpha \to \pm\infty$ the results of $LBABSN_2(\alpha,\beta)$ distribution, reduces to that of $BLN(4)$ distribution as $\beta_1 \to 17.1334,\ \beta_2 \to 48.5346$. Again, in limit $\beta \to \pm\infty$ in the results of $BABSN_2(\alpha,\beta)$ distribution, gives for $BLN(12)$ distribution as $\beta_1 \to 13.3692,\ \beta_2 \to 36.1299$.

3. Random Number Generation

Here we express the $LBABSN_2(\alpha,\beta)$ distribution into a sum of asymmetric and an asymmetric component and then generate a random number from $LABSN(\alpha,\beta)$ distribution.

Lemma1: The pdf in the Eqn. (5) of $LBABSN_2(\alpha,\beta)$ can be represented as the sum of two functions

$$f_Z(z;\alpha,\beta) = \frac{4+y^2(\alpha+\beta y^2)^2(8+y^2(\alpha+\beta y^2)^2)}{z\,C_2(\alpha,\beta)}\phi(y) - \frac{4y(\alpha+\beta y^2)(2+y^2(\alpha+\beta y^2)^2)}{z\,C_2(\alpha,\beta)}\phi(y) \quad (8)$$

In Eqn. (8), the 1st part is symmetric while the second part is asymmetric one and the symmetric part is denoted by $SCLBABSN_2(\alpha,\beta)$. For $\alpha=\beta=0$, $Z \sim LN(0,1) = \phi(y)/z$.

Remark 3: To generate data from $LBABSN_2(\alpha,\beta)$ distribution, we use the acceptance-rejection algorithm (Von Neumann, 1951) as follows:

Let $f(x)$ be the pdf of $X \sim LBABSN_2(\alpha,\beta)$ and $f_1(x)$ is the pdf of $Z \sim SCLBABSN_2(\alpha,\beta)$. Thus

$$M = Sup\left[\frac{f(x)}{f_1(x)}\right] = \frac{1}{3}(3+2\sqrt{2})$$

To generate an $X \sim LBABSN_2(\alpha,\beta)$, we shall carry out the following steps:

a) Generate $Z \sim SCLBABSN_2(\alpha,\beta)$ (For generating random sample from $SCLBABSN_2(\alpha,\beta)$ distribution one can used the *Metropolis–Hastings*(MH) algorithm (Metropolis *et al.*, 1953; Hastings, 1970).
b) Generate $U \sim Uniform(0,1)$ independently from Z .

c) If $U < \dfrac{1}{M}\dfrac{f(Z)}{f_1(Z)} = \dfrac{3[(1-\alpha Y-\beta Y^3)^2+1]^2}{(3+2\sqrt{2})[4+Y^2(\alpha+\beta Y^2)^2(8+Y^2(\alpha+\beta Y^2)^2)]}$, set $X = Z$ accept; otherwise, go back to step one (reject).

By the acceptance-rejection method, any choice of this random variable will be accepted with probability $\dfrac{1}{M}$, i.e., $P\left(U < \dfrac{1}{M}\dfrac{f(Z)}{f_1(Z)}\right) = \dfrac{1}{M} = \dfrac{3}{(3+2\sqrt{2})}$. Thus, the number of trails is geometric with $p = \dfrac{1}{M}$, the expected value for this number is $M = \dfrac{(3+2\sqrt{2})}{3} = 1.9428$.

4. Parameter Estimation of $LBABSN_2(\alpha,\beta)$ Distribution

When $Z \sim LABSN(\alpha,\beta)$ distribution then for $Y = \mu + \sigma Z$ we get the location (μ) and scale (σ) extension of $LABSN(\alpha,\beta)$ distribution with pdf

$$f_Y(y;\mu,\sigma,\alpha,\beta) = \frac{\left[\left(1-\alpha\left(\frac{y-\mu}{\sigma}\right)-\beta\left(\frac{y-\mu}{\sigma}\right)^3\right)^2+1\right]}{z\, C_2(\alpha,\beta)}\phi\left(\frac{y-\mu}{\sigma}\right) \tag{9}$$

Where $y = Log(z)$, $z > 0$, $\alpha \in R$, $\beta \in R$ and $C_2(\alpha,\beta)$ is defined above and denote it by $Y \sim LBABSN_2(\mu,\sigma,\alpha,\beta)$.

4.1. Maximum Likelihood Estimation

The log-likelihood function of the random sample $y_1, y_2, \ldots, y_n$ from $Y \sim LBABSN_2(\mu,\sigma,\alpha,\beta)$ distribution with parameters $\theta = (\mu,\sigma,\alpha,\beta)$ is given by

$$l(\theta) = 2\sum_{i=1}^{n}\log\left[\left\{1-\alpha\left(\frac{y_i-\mu}{\sigma}\right)-\beta\left(\frac{y_i-\mu}{\sigma}\right)^3\right\}^2+1\right] - n\log C_2(\alpha,\beta) - n\log(\sigma) - \frac{n}{2}\log(2\pi) - \sum_{i=1}^{n} y_i - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i-\mu}{\sigma}\right)^2 \tag{10}$$

By differentiating Eqn. (10) partially with respect to the parameters $\theta = (\mu,\sigma,\alpha,\beta)$, we get the following likelihood equations as follows:

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

$$\frac{\partial l(\theta)}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{(y_i-\mu)^2}{\sigma^3} + 2\sum_{i=1}^{n} \frac{2\left(\frac{\alpha(y_i-\mu)}{\sigma^2} + \frac{3\beta(y_i-\mu)^3}{\sigma^4}\right) b_i}{(1+b_i^2)}$$

$$\frac{\partial l(\theta)}{\partial \alpha} = -\frac{n}{C_2(\alpha,\beta)} \begin{bmatrix} 12\alpha^3 + 180\alpha^2\beta + 12\beta(4+315\beta^2) + \\ 2\alpha(8+630\beta^2) \end{bmatrix} + 2\sum_{i=1}^{n} -\frac{2(y_i-\mu)b_i}{\sigma(1+b_i^2)}$$

$$\frac{\partial l(\theta)}{\partial \beta} = -\frac{n}{C_2(\alpha,\beta)} \begin{bmatrix} 60\alpha^3 + 1260\alpha^2\beta + 7560\alpha\beta^2 + 20790\beta^3 + \\ 12\alpha(4+315\beta^2) + 30\beta(8+693\beta^2) \end{bmatrix} + 2\sum_{i=1}^{n} -\frac{2(y_i-\mu)^3 b_i}{\sigma^3(1+b_i^2)}$$

where $y_i = Log(z_i)$, $b_i = \left(1 - \frac{\alpha(y_i-\mu)}{\sigma} - \frac{\beta(y_i-\mu)^3}{\sigma^3}\right)$.

Solutions of the above system of likelihood equations is derived by numerical maximization of Eqn. (10) with respect to the parameters $\theta = (\mu, \sigma, \alpha, \beta)$ to compute the maximum likelihood estimator for the parameters.

5. Applications in data modeling

Here we have considered two datasets for checking the usefulness of our proposed distribution. Dataset 1 is the exchange rate data of the United Kingdom Pound to the United States Dollar from 1800 to 2003 which is obtained from the website http://www.globalfindata.com, and Dataset 2 consists of the velocities of 82 distant galaxies, diverging from our own galaxy available at http://www.stats.bris.ac.uk/~peter/mixdata. The summary statistics of the datasets are given in Table 1 below.

Table 1: Summary Statistic for the Datasets.

Datasets	Median	Mean	SD	Skewness	Kurtosis
1	4.753	4.117	1.384	0.026	5.282
2	20.834	20.831	4.568	-0.431	5.259

The proposed $LBABSN_2(\mu,\sigma,\alpha,\beta)$ distribution is compared with the log-normal $LN(\mu,\sigma^2)$ distribution, the log-skew-normal $LSN(\mu,\sigma,\lambda)$ distribution, the log-alpha-skew-normal distribution $LASN(\mu,\sigma,\alpha)$ of Venegas *et al.* (2016), the log-Balakrishnan-alpha-skew-normal $LBASN_2(\mu,\sigma,\alpha)$ distribution of Shah *et al.* (2020b) and the log-Balakrishnan-beta-skew-normal $LBBSN(\mu,\sigma,\beta)$ distribution. Using GenSA package (see GenSA package version-1.0.3, Xiang *et al.* 2013) from R software, the MLE of the parameters are computed. AIC and BIC are used for model comparison. Table 2 and 3 shows the MLE's, log-likelihood, AIC and BIC of the distributions. The graphical representations of the results are given in Figure 7 and 8.

Table 2: MLE's, log-likelihood, AIC and BIC for Dataset 1.

Parameters → Distribution ↓	μ	σ	λ	α	β	$\log L$	AIC	BIC
$LN(\mu,\sigma^2)$	1.344	0.406	--	--	--	-379.953	763.906	770.542
$LSN(\mu,\sigma,\lambda)$	1.774	0.592	-4.041	--	--	-350.678	707.356	717.311
$LBASN_2(\mu,\sigma,\alpha)$	1.201	0.199	--	-3.809	--	-342.201	690.402	700.356
$LBBSN_2(\mu,\sigma,\beta)$	1.503	0.266	--	--	0.077	-341.237	688.474	698.428
$LASN(\alpha,\mu,\sigma)$	1.218	0.255	--	-4.029	--	-326.841	659.682	669.637
$LBABSN_2(\mu,\sigma,\alpha,\beta)$	1.062	0.149	--	0.416	-0.229	**-302.686**	**613.372**	**626.645**

From Table 2 and 3, it is seen that the proposed Log-Balakrishnan-alpha-beta-skew-normal $LBABSN_2(\mu,\sigma,\alpha,\beta)$ distribution provides better fit to the data set 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 (lines) densities presented in Figure 7 and 8, also confirm our findings.

5.1. Likelihood Ratio Test

Since $LN(\mu,\sigma^2)$, $LBASN_2(\mu,\sigma,\alpha)$ and $LBABSN_2(\mu,\sigma,\alpha,\beta)$ are nested models, the likelihood ratio (LR) test is used to discriminate between them.

The results are presented in Table 4.

Table 3: MLE's, log-likelihood, AIC and BIC for Dataset 2.

Parameters → / Distribution ↓	μ	σ	λ	α	β	$\log L$	AIC	BIC
$LN(\mu,\sigma^2)$	3.007	0.258	--	--	--	-251.937	507.874	512.687
$LSN(\mu,\sigma,\lambda)$	3.262	0.363	-2.735	--	--	-242.505	491.011	498.231
$LBBSN_2(\mu,\sigma,\beta)$	3.068	0.207	--	--	0.052	-240.642	487.284	494.504
$LASN(\alpha,\mu,\sigma)$	2.806	0.214	--	-1.992	--	-230.730	467.460	474.681
$LBASN_2(\mu,\sigma,\alpha)$	2.739	0.186	--	-1.973	--	-217.716	441.432	448.652
$LBABSN_2(\mu,\sigma,\alpha,\beta)$	2.905	0.156	--	-1.369	0.201	**-211.703**	**431.406**	**441.033**

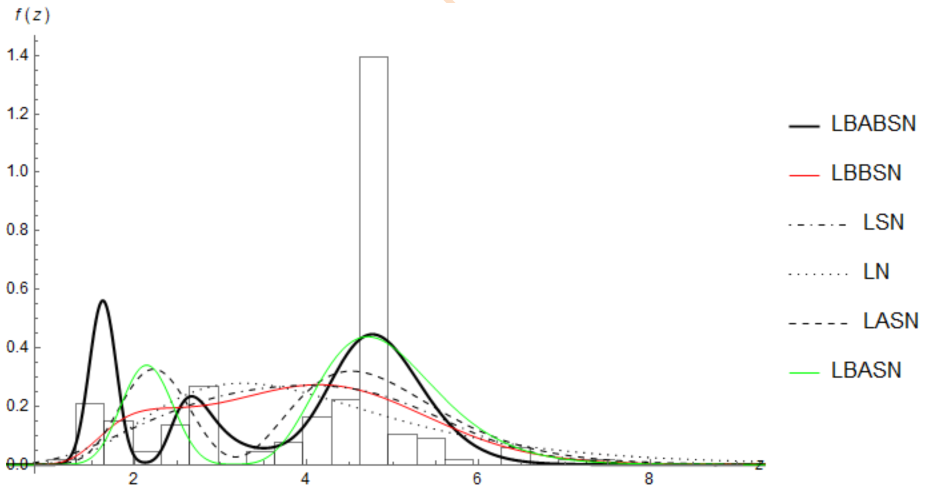

Figure 7: Observed and expected densities for Dataset 1

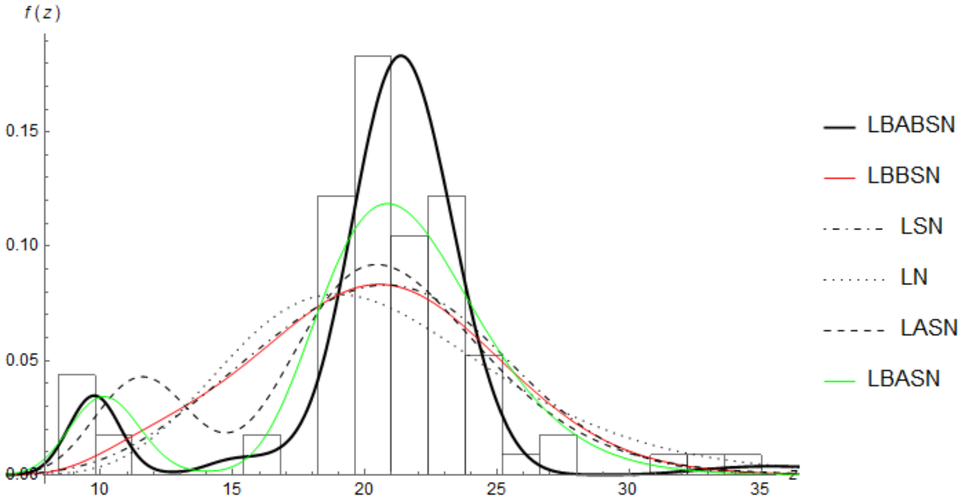

Figure 8: Observed and expected densities for Dataset 2.

Table 4: The findings of the LR test statistic for different Datasets.

Hypothesis	LR test statistic values		Degrees of Freedom	Critical values at 5%
	Dataset 1	Dataset 2		
$H_0: \beta = 0$ vs $H_1: \beta \neq 0$	79.03	12.026	1	3.841
$H_0: \alpha = 0, \beta = 0$ vs $H_1: \alpha \neq 0, \beta \neq 0$	154.534	80.468	2	5.991

In table 4, we observe that the values of LR test statistic exceed the corresponding critical value at 5% level of significance. Thus, there is enough evidence in support of the alternative hypothesis and we may conclude that the sampled data comes from $BABSLG_2(\mu, \sigma, \alpha, \beta)$ distribution and not from other distribution considered.

6. Conclusion

The log-Balakrishnan-alpha-beta-skew-normal distribution is constructed which includes unimodal, bimodal as well as multimodal shapes and some of its properties are studied. Our findings adequately supported the proposed $LBABSN_2(\mu,\sigma,\alpha,\beta)$ distribution as the better fitted model to the datasets under consideration in terms of AIC and BIC values. The plots of observed and expected densities for three datasets presented also confirm our findings. Furthermore, there is scope of extending the present work by considering the Logistic and the Laplace distributions and extending to bivariate as future work.

Acknowledgement: We are thankful to the unknown reviewer for constructive as well as creative suggestions for the preparation of manuscript.

References

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

[2] Azzalini, A., Dal Cappello T., and Kotz S., (2003), Log-skew-normal and log-skew-t distributions as models for family income data,Journal of Income Distribution,11(3-4), 12-20.

[3] Ahrens, L., (1953), A fundamental law of geochemistry, Nature, 172(4390), 1148.

[4] Ahrens, L., (1954a), The lognormal distribution of the elements (a fundamental law of geochemistry and its subsidiary),Geochimica et Cosmochimicaacta, 5(2), 49-73.

[5] Ahrens, L., (1954b), The lognormal distribution of the elements (2), Geochimica et Cosmochimicaacta,6(2-3), 121-131.

[6] Chakraborty, S., Hazarika, P. J., and Ali, M. M., (2012), A New Skew Logistic Distribution and Its Properties, Pak. J. Statist, 28(4), 513-524.

[7] Chakraborty, S., Hazarika, P. J., and Ali, M. M., (2014), A Multimodal Skew Laplace Distribution. Pak. J. Statist, 30(2), 253-264.

[8] 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.

[9] Elal-Olivero, D., (2010), Alpha-skew-normal distribution, Proyecciones (Antofagasta), 29(3), 224-240.

[10] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1), 97–109.

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

[12] 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.

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

[14] Mateu-Figueras, G., Pawlosky-Glanh, and Barcelo-Vidal, C., (2004), The natural law in geochemistry: Lognormal or log skew-normal?,“32th International Geological Congress”, International Union of Soil Sciences, Florence, Italy, 1849-1858.

[15] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. The Journal of Chemical Physics, 21(6): 1087–1092.

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

[17] 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.

[18] Shah, S., and Hazarika, P. J., (2019), The Alpha-Beta-Skew-Logistic Distribution And Its Applications, arXiv preprint arXiv:1909.03651.

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

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

[21] 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.

[22] Vistelius, A. B., (1960), The skew frequency distributions and the fundamental law of the geochemical processes, The Journal of Geology, 68 (1), 1-22.

[23] Von Neumann, J., (1951), Various techniques used in connection with random digits. Monte Carlo methods, National Bureau of Standards, Applied Mathematics Series, 12, 36-38.

[24] Xiang, Y., Gubian, S., Suomela, B., and Hoeng, J., (2013), Generalized Simulated Annealing for Global Optimization: The GenSA Package, R Journal, 5(1), 13-28.

Appendix

A: Proof of cdf

$$F_Z(z;\alpha,\beta)=P(Z\le z) = \int_0^z \frac{[(1-\alpha\, y-\beta\, y^3)^2+1]^2}{z\, C_2(\alpha,\beta)}\, \varphi(y)\, dz$$

$$=\frac{1}{C_2(\alpha,\beta)}\int_0^z \left[\begin{array}{l} 4-8\alpha y+8\alpha^2 y^2-4(\alpha^3-2\beta)y^3+\alpha(\alpha^3+16\beta)y^4-12\alpha^2\beta y^5+ \\ 4\beta(\alpha^3+2\beta)y^6-12\alpha\beta^2 y^7+6\alpha^2\beta^2 y^8-4\beta^3 y^9+4\alpha\beta^3 y^{10}+\beta^4 y^{12} \end{array}\right]\frac{\varphi(y)}{z}\, dz$$

$$=\frac{1}{C_2(\alpha,\beta)}\Bigg[4\int_0^z \frac{\phi(y)}{z}dz-8\alpha\int_0^z \frac{y\,\phi(y)}{z}dz+8\alpha^2\int_0^z \frac{y^2\,\phi(y)}{z}dz-4\,(\alpha^3+2\beta)\int_0^z \frac{y^3\,\phi(y)}{z}dz+$$

$$\alpha(\alpha^3+16\beta)\int_0^z \frac{y^4\,\phi(y)}{z}dz-12\alpha^2\beta\int_0^z \frac{y^5\,\phi(y)}{z}dz+4\beta(\alpha^3+2\beta)\int_0^z \frac{y^6\,\phi(y)}{z}dz-12\alpha\beta^2$$

$$\int_0^z \frac{y^7\,\phi(y)}{z}dz+6\alpha^2\beta^2\int_0^z \frac{y^8\,\phi(y)}{z}dz-4\beta^3\int_0^z \frac{y^9\,\phi(y)}{z}dz+4\alpha\beta^3\int_0^z \frac{y^{10}\,\phi(y)}{z}dz+\beta^4\int_0^z \frac{y^{12}\,\phi(y)}{z}dz\Bigg] \qquad \text{(A1)}$$

Now, we have

$$\int_0^z \frac{\varphi(y)}{z}dz=\Phi(y),\ \int_0^z \frac{y\,\varphi(y)}{z}dz=-\varphi(y),\ \int_0^z \frac{y^2\phi(y)}{z}dz=\Phi(y)-y\,\phi(y),$$

$$\int_0^z \frac{y^3\varphi(y)}{z}dz=-(2+y^2)\varphi(y),\ \int_0^z \frac{y^4\phi(y)}{z}dz=3\Phi(y)-z(3y+y^3)\phi(y),$$

$$\int_0^z \frac{y^5\varphi(y)}{z}dz=-(8+4y^2+y^4)\varphi(y),\ \int_0^z \frac{y^6\phi(y)}{z}dz=15\Phi(y)-(15+5y^2+y^4)\phi(y),$$

$$\int_0^z \frac{y^7\phi(y)}{z}dz=-(48+24y^2+6y^4+y^6)\phi(y),$$

$$\int_0^z \frac{y^8\phi(y)}{z}dz=105\Phi(y)-(105y+35y^3+7y^5+y^7)\phi(z),$$

$$\int_0^z \frac{y^9\varphi(y)}{z}dz=-(384+192y^2+48y^4+8y^6+y^8)\varphi(y),$$

$$\int_0^z \frac{y^{10}\varphi(y)}{z}dz = 945\Phi(y) - (945y + 315y^3 + 63y^5 + 9y^7 + y^9)\varphi(y),$$

$$\int_0^z \frac{y^{12}\varphi(y)}{z}dz = 10395\Phi(y) - (10395y + 3465y^3 + 693y^5 + 99y^7 + 11y^9 + y^{11})\varphi(y).$$

Putting these values in Eqn. (A1), we get the desired result in Eqn. (5).

B: Proof of Mode

Differentiating Eqn. (4) with respect to z, we have

$$f_Z'(z;\alpha,\beta) = \frac{\partial f_Z(z;\alpha,\beta)}{\partial z} = \frac{\partial}{\partial z}\frac{[(1-\alpha y - \beta y^3)^2 + 1]^2}{z C_2(\alpha,\beta)}\frac{\varphi(y)}{z};\ y = Log(z)$$

$$= \frac{1}{C_2(\alpha,\beta)}\frac{\partial}{\partial z}[(1-\alpha y - \beta y^3)^2 + 1]^2 \frac{\phi(y)}{z}$$

$$= \frac{1}{C_2(\alpha,\beta)}\frac{\partial}{\partial z}\begin{bmatrix} 4 - 8\alpha y + 8\alpha^2 y^2 - 4(\alpha^3 - 2\beta)y^3 + \alpha(\alpha^3 + 16\beta)y^4 - 12\alpha^2\beta y^5 + \\ 4\beta(\alpha^3 + 2\beta)y^6 - 12\alpha\beta^2 y^7 + 6\alpha^2\beta^2 y^8 - 4\beta^3 y^9 + 4\alpha\beta^3 y^{10} + \beta^4 y^{12} \end{bmatrix}\frac{\varphi(y)}{z}$$

$$= \frac{1}{C_2(\alpha,\beta)}\left[4\frac{\partial}{\partial z}\left\{\frac{\phi(y)}{z}\right\} - 8\alpha\frac{\partial}{\partial z}\left\{\frac{y\phi(y)}{z}\right\} + 8\alpha^2\frac{\partial}{\partial z}\left\{\frac{y^2\phi(y)}{z}\right\} - 4(\alpha^3 - 2\beta)\frac{\partial}{\partial z}\left\{\frac{y^3\phi(y)}{z}\right\} + \alpha(\alpha^3\right.$$
$$+16\beta)\frac{\partial}{\partial z}\left\{\frac{y^4\phi(y)}{z}\right\} - 12\alpha^2\beta\frac{\partial}{\partial z}\left\{\frac{y^5\phi(y)}{z}\right\} + 4\beta(\alpha^3 + 2\beta)\frac{\partial}{\partial z}\left\{\frac{y^6\phi(y)}{z}\right\} - 12\alpha\beta^2\frac{\partial}{\partial z}\left\{\frac{y^7\phi(y)}{z}\right\}$$
$$\left.+6\alpha^2\beta^2\frac{\partial}{\partial z}\left\{\frac{y^8\phi(y)}{z}\right\} - 4\beta^3\frac{\partial}{\partial z}\left\{\frac{y^9\phi(y)}{z}\right\} + 4\alpha\beta^3\frac{\partial}{\partial z}\left\{\frac{y^{10}\phi(y)}{z}\right\} + \beta^4\frac{\partial}{\partial z}\left\{\frac{y^{12}\phi(y)}{z}\right\}\right] \quad \text{(B1)}$$

Now, we have

$$\frac{\partial}{\partial z}\left\{\frac{\varphi(y)}{z}\right\} = -\frac{(y+1)}{z^2}\varphi(y),\ \frac{\partial}{\partial z}\left\{\frac{y\phi(y)}{z}\right\} = -\frac{(y^2+y-1)}{z^2}\phi(y),$$

$$\frac{\partial}{\partial z}\left\{\frac{y^2\varphi(y)}{z}\right\} = -\frac{y(y^2+y-2)}{z^2}\varphi(y), \frac{\partial}{\partial z}\left\{\frac{y^3\phi(y)}{z}\right\} = -\frac{y^2(y^2+y-3)}{z^2}\phi(y),$$

$$\frac{\partial}{\partial z}\left\{\frac{y^4\phi(y)}{z}\right\} = -\frac{y^3(y^2+y-4)}{z^2}\phi(y), \frac{\partial}{\partial z}\left\{\frac{y^5\phi(y)}{z}\right\} = -\frac{y^4(y^2+y-5)}{z^2}\phi(y),$$

$$\frac{\partial}{\partial z}\left\{\frac{y^6\varphi(y)}{z}\right\}=-\frac{y^5(y^2+y-6)}{z^2}\varphi(y),\ \frac{\partial}{\partial z}\left\{\frac{y^7\phi(y)}{z}\right\}=-\frac{y^6(y^2+y-7)}{z^2}\phi(y),$$

$$\frac{\partial}{\partial z}\left\{\frac{y^8\varphi(y)}{z}\right\}=-\frac{y^7(y^2+y-8)}{z^2}\varphi(y),\ \frac{\partial}{\partial z}\left\{\frac{y^9\phi(y)}{z}\right\}=-\frac{y^8(y^2+y-9)}{z^2}\phi(y),$$

$$\frac{\partial}{\partial z}\left\{\frac{y^{10}\varphi(y)}{z}\right\}=-\frac{y^9(y^2+y-10)}{z^2}\varphi(y),\ \frac{\partial}{\partial z}\left\{\frac{y^{12}\phi(y)}{z}\right\}=-\frac{y^{11}(y^2+y-12)}{z^2}\phi(y).$$

Putting these values in Eqn. (B1), we get

$$f_Z'(z;\alpha,\beta)=-\frac{\varphi(y)}{z^2C_2(\alpha,\beta)}[\{2-2\alpha y+\alpha^2y^2-2\beta y^3+2\alpha\beta y^4+\beta^2y^6\}\{2+4\alpha-2(-1+\alpha+2\alpha^2)y+(-2\alpha+\alpha^2+12\beta)y^2+(\alpha^2-2\beta-16\alpha\beta)y^3+2\beta(-1+\alpha)y^4+2\beta(\alpha-6\beta)y^5+\beta^2y^6+\beta^2y^7\}] \quad \text{(B2)}$$

Since the Eqn. (B2) has at most seven zeros. Thus, the function $f_Z(z;\alpha,\beta)$ can have at most four modes.

C: Proof of n^{th} moment

$$E(Z^n)=\int_0^\infty z^n\frac{[(1-\alpha y-\beta y^3)^2+1]^2}{zC_2(\alpha,\beta)}\phi(y)\,dz$$

$$=\frac{1}{C_2(\alpha,\beta)}\int_0^\infty z^n\begin{bmatrix}4-8\alpha y+8\alpha^2y^2-4(\alpha^3-2\beta)y^3+\alpha(\alpha^3+16\beta)y^4-12\alpha^2\beta y^5+\\4\beta(\alpha^3+2\beta)y^6-12\alpha\beta^2y^7+6\alpha^2\beta^2y^8-4\beta^3y^9+4\alpha\beta^3y^{10}+\beta^4y^{12}\end{bmatrix}\frac{\phi(z)}{z}dz$$

$$=\frac{1}{C_2(\alpha,\beta)}\Bigg[4\int_0^\infty z^n\frac{\varphi(y)}{z}dz-8\alpha\int_0^\infty z^n\frac{y\varphi(y)}{z}dz+8\alpha^2\int_0^\infty z^n\frac{y^2\varphi(y)}{z}dz-4(\alpha^3+2\beta)\int_0^\infty z^n\frac{y^3\varphi(y)}{z}dz+\alpha(\alpha^3+16\beta)\int_0^\infty z^n\frac{y^4\varphi(y)}{z}dz-12\alpha^2\beta\int_0^\infty z^n\frac{y^5\varphi(y)}{z}dz+4\beta(\alpha^3+2\beta)\int_0^\infty z^n\frac{y^6\varphi(y)}{z}dz-12\alpha\beta^2\int_0^\infty z^n\frac{y^7\varphi(y)}{z}dz+6\alpha^2\beta^2\int_0^\infty z^n\frac{y^8\varphi(y)}{z}dz-4\beta^3\int_0^\infty z^n\frac{y^9\varphi(y)}{z}dz+4\alpha\beta^3\int_0^\infty z^n\frac{y^{10}\varphi(y)}{z}dz+\beta^4\int_0^\infty z^n\frac{y^{12}\varphi(y)}{z}dz\Bigg] \quad \text{(C1)}$$

Now, we have

$$\int_0^\infty z^n \frac{\phi(y)}{z} dz = e^{\frac{n^2}{2}}, \int_0^\infty z^n \frac{y\varphi(y)}{z} dz = n e^{\frac{n^2}{2}}, \int_0^\infty z^n \frac{y^2\varphi(y)}{z} dz = (n^2+1) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^3\varphi(y)}{z} dz = n(n^2+3) e^{\frac{n^2}{2}}, \int_0^\infty z^n \frac{y^4\phi(y)}{z} dz = (n^4+6n^2+3) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^5\phi(y)}{z} dz = n(n^4+10n^2+15) e^{\frac{n^2}{2}}, \int_0^\infty z^n \frac{y^6\varphi(y)}{z} dz = (n^6+15n^4+45n^2+15) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^7\varphi(y)}{z} dz = n(n^6+21n^4+105n^2+105) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^8\varphi(y)}{z} dz = (n^8+28n^6+210n^4+420n^2+105) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^9\varphi(y)}{z} dz = n(n^8+36n^6+378n^4+1260n^2+945) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^{10}\varphi(y)}{z} dz = (n^{10}+45n^8+630n^6+3150n^4+4725n^2+945) e^{\frac{n^2}{2}},$$

$$\int_0^\infty z^n \frac{y^{12}\phi(y)}{z} dz = (n^{12}+66n^{10}+1485n^8+13860n^6+51975n^4+62370n^2+10395) e^{\frac{n^2}{2}}.$$

Putting these values in Eqn. (C1), we get the desired result in Eqn. (6).