Pak. J. Statist.
2010 Vol. 26(2), 379-396

A BIMODAL EXPONENTIAL POWER DISTRIBUTION

Mohamed Y. Hassan[1] and **Rafiq H. Hijazi**[2]
Department of Statistics
United Arab Emirates University
P.O. Box 17555, Al-Ain, U.A.E.
Email: [1]myusuf@uaeu.ac.ae
[2]rhijazi@uaeu.ac.ae

ABSTRACT

In this paper, a new generalized bimodal power exponential distribution is proposed. The Characterization of this distribution is investigated and special cases are considered. The problem of obtaining the maximum-likelihood estimates for the parameters of the distribution is studied and the information matrix is derived. A simulation study to assess the properties of the maximum likelihood estimators is carried out. Finally, two illustrative examples are presented.

KEY WORDS

Bimodal exponential power distribution; characterization; meteorology.

1. INTRODUCTION

Data that exhibit bimodal behavior arises in many different disciplines. In medicine, urine mercury excretion has two peaks, see for example, Ely et al. (1999). In material characterization, a study conducted by Dierickx et al. (2000), grain size distribution data reveals a bimodal structure. In meteorology, Zhang et al. (2003) indicated that, water vapor in tropics, commonly have bimodal distributions. The most commonly used distribution in modeling bimodal data is the two-component normal mixture. Many authors have proposed the bimodal normal distribution, see for example, Sarma et al. (1990), Bhagavan et al. (1983), Rao et al. (1988), Prasada Rao (1987), and Rao et al. (1987) but these studies did not materialize in the real world of statistics. The bimodal normal is closely related to the Exponential Power Family (EPF) which is considered as one of the most important probability distributions in statistics. Different forms of this family have been studied in the literature. The symmetric exponential power distribution with normal, Laplace and rectangular as special cases was considered by Subbotin (1923), Box (1953), Turner (1960), Box and Tiao (1992) and others. Box and Tiao (1992) explain the use of such family of distributions in a Bayesian context. The skew exponential power distribution has been studied by Azzalini (1985), Azzalini and Dalla Valle (1996), Azzalini and Capitanio (1999), and Mudholkar and Hutson (2000). Most recently the generalization of skew normal was investigated by Gupta and Brown (2001), Gupta and Gupta (2004). DiCiccio and Monti (2004) investigated inferential aspects of the skew exponential power distribution. Despite all these studies on the unimodal symmetric exponential power distribution, the general form of the bimodal exponential power distribution is not yet considered in the literature except the special case of the bimodal normal. The two general cases of the exponential power distribution; the symmetric bimodal and the skew bimodal can be investigated separately. This paper considers the symmetric bimodal exponential power distribution which has the unimodal symmetric exponential power distribution as a special case. This distribution can be used as an alternative to the two component normal mixtures in modeling bimodal data, since it is more flexible, parsimonious and has the advantage of estimation simplicity.

Definition and properties of the proposed bimodal exponential power distribution are given in Section 2. The characterization of the distribution is studied in Section 3. The inferential aspects of the proposed distribution are discussed theoretically in Section 4 and numerically in Section 5. Two illustrative examples in which the bimodal exponential power distribution is used in modeling bimodal data are given in Section 6. General conclusions and possible extensions of the proposed distribution are given in Section 7.

2. DEFINITION AND SOME PROPERTIES

2.1 Definition

A random variable X is said to have a bimodal exponential power distribution $BEP(\mu,\psi,\delta,\alpha)$ if there exist parameters $-\infty<\mu<\infty$, $\delta\geq 0, \alpha\geq 1$, $\psi>0$ such that the density function of X has the form

$$f(x)=\frac{\alpha\left|\frac{x-\mu}{\psi}\right|^{\delta} e^{-\left|\frac{x-\mu}{\psi}\right|^{\alpha}}}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}, \quad -\infty<x<\infty \tag{2.1}$$

where μ is a location parameter and ψ is a scale parameters. Fig. (1a) shows the $BEP(0,1,1,\alpha)$ densities for $\alpha=1,2,3$ and Fig. (1b) shows the $BEP(0,1,\delta,2)$ densities for $\delta=0,1,5$.

Clearly α is a scale parameter, and it is inversely related to the variance of the distribution, where as the δ is the bimodality parameter. It is interesting to note that, if the shape parameter $\delta=0$, BEP coincides with the general unimodal symmetric exponential power distribution with the normal and the Laplace as special cases.

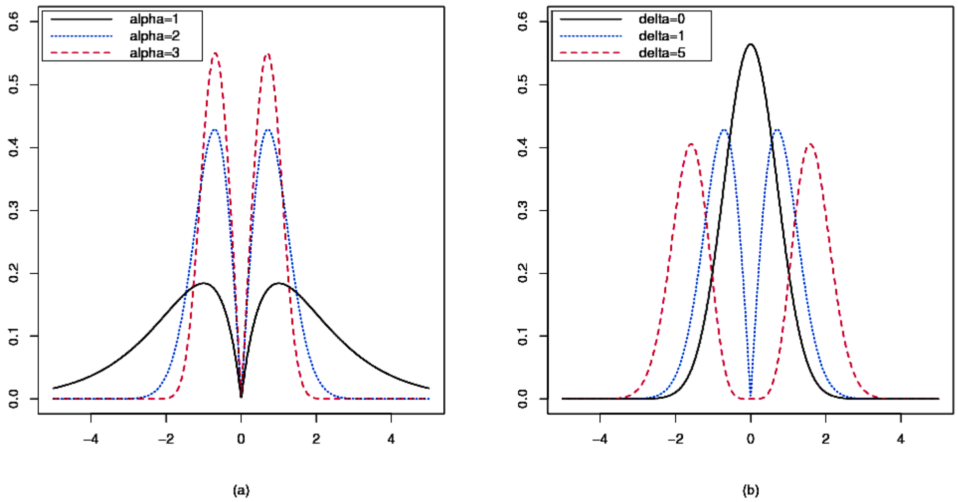

Fig. 1: Densities of the (a) $BEP(0,1,1,\alpha)$
(b) $BEP(0,1,\delta,2)$

2.2 Moments and Other Measures

Theorem 1: If $X \sim BEP(\mu,\psi,\delta,\alpha)$ then the central moments of X are given by

$$E(X-\mu)^k = \begin{cases} 0 & k = 2m+1 \\ \dfrac{\psi^{2m}\Gamma\left(\dfrac{2m+\delta+1}{\alpha}\right)}{\Gamma\left(\dfrac{\delta+1}{\alpha}\right)} & k = 2m \end{cases} \tag{2.2}$$

Proof:

$$\begin{aligned} E(X-\mu)^{2m} &= \int_{-\infty}^{\infty} \frac{\alpha}{2\Gamma\psi\left(\frac{\delta+1}{\alpha}\right)}(x-\mu)^{2m}\left|\frac{x-\mu}{\psi}\right|^{\delta} e^{-\left|\frac{x-\mu}{\psi}\right|^{\alpha}} dx \\ &= \int_{0}^{\infty} \frac{\alpha\psi^{2m}}{\Gamma\left(\frac{\delta+1}{\alpha}\right)} y^{2m+\delta} e^{-y^{\alpha}} dy \\ &= \int_{0}^{\infty} \frac{\psi^{2m}}{\Gamma\left(\frac{\delta+1}{\alpha}\right)} w^{\frac{2m+\delta+1}{\alpha}-1} e^{-w} dw \\ &= \frac{\psi^{2m}\Gamma\left(\dfrac{2m+\delta+1}{\alpha}\right)}{\Gamma\left(\dfrac{\delta+1}{\alpha}\right)} \end{aligned}$$

For $k = 2m+1$,

$$E(X-\mu)^{2m+1} = \int_{-\infty}^{\infty} \frac{\alpha}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}(x-\mu)^{2m+1}\left|\frac{x-\mu}{\psi}\right|^{\delta} e^{-\left|\frac{x-\mu}{\psi}\right|^{\alpha}} = 0 .$$

It can be easily seen that

$$Var(X) = E(X-\mu)^2 = \frac{\psi^2\Gamma\left(\frac{\delta+3}{\alpha}\right)}{\Gamma\left(\frac{\delta+1}{\alpha}\right)}, \tag{2.3}$$

and the kurtosis of X is

$$\kappa_4 = \frac{E(X-\mu)^4}{\left[E(X-\mu)^2\right]^2} - 3 = \frac{\Gamma\left(\frac{\delta+5}{\alpha}\right)\Gamma\left(\frac{\delta+1}{\alpha}\right)}{\left[\Gamma\left(\frac{\delta+3}{\alpha}\right)\right]^2} - 3 . \tag{2.4}$$

3. CHARACTERIZATION OF $BEP(\mu,\psi,\delta,\alpha)$

3.1 Maximum Entropy Property

Many families of univariate probability distributions are known to maximize the entropy among distributions that satisfy given constraints on the expectations of certain statistics. Examples of such families are the Beta, Gamma, and the normal distributions.

Theorem 2: Let W have an $BEP(\mu,\psi,\delta,\alpha)$ distribution with density f given by (2.1). Then the entropy of W is given by

$$H(W) = -\ln\left[\frac{\alpha}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\right] - \frac{\delta}{\alpha}\Psi\left(\frac{\delta+1}{\alpha}\right) + \frac{\Gamma\left(\frac{2\delta+1}{\alpha}\right)}{\Gamma\left(\frac{\delta+1}{\alpha}\right)} \tag{3.1}$$

where $\Psi\left(\frac{\delta+1}{\alpha}\right) = \frac{\Gamma_\delta\left(\frac{\delta+1}{\alpha}\right)}{\Gamma\left(\frac{\delta+1}{\alpha}\right)}$ is the digamma function and Γ_δ is the partial derivative of the gamma function with respect to δ.

Proof:

$$H(W) = \int_{-\infty}^{\infty} -\left(\ln\left[\frac{\alpha}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\right] + \ln\left|\frac{w-\mu}{\psi}\right|^{\delta} - \frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}\right)\frac{\alpha\left|\frac{w-\mu}{\psi}\right|^{\delta} e^{\frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}}}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\, dw$$

$$= -\int_{-\infty}^{\infty} \ln\left[\frac{\alpha}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\right] \frac{\alpha\left|\frac{w-\mu}{\psi}\right|^{\delta} e^{\frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}}}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\, dw$$

$$-\int_{-\infty}^{\infty} \left(\ln\left|\frac{w-\mu}{\psi}\right|^{\delta}\right) \frac{\alpha\left|\frac{w-\mu}{\psi}\right|^{\delta} e^{\frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}}}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\, dw$$

$$+\int_{-\infty}^{\infty} \left(\frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}\right) \frac{\alpha\left|\frac{w-\mu}{\psi}\right|^{\delta} e^{\frac{|w-\mu|^{\alpha}}{\psi^{\alpha}}}}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\, dw$$

$$= -\ln\left[\frac{\alpha}{2\psi\Gamma\left(\frac{\delta+1}{\alpha}\right)}\right] - \frac{\delta}{\alpha}\Psi\left(\frac{\delta+1}{\alpha}\right) + \frac{\Gamma\left(\frac{2\delta+1}{\alpha}\right)}{\Gamma\left(\frac{\delta+1}{\alpha}\right)}.$$

Lemma 1: Let $\beta_1 < 0$ and $\beta_1 > -1$ and let

$$p(w) = e^{\beta_0 + \beta_1 |w|^{\alpha} + \beta_2 \ln|w|} \qquad -\infty < w < \infty \tag{3.2}$$

be a probability density function, then

1. $\tau_1(\theta) = E_p \ln|Y|^{\alpha} = \dfrac{\beta_2 + 1}{\alpha|\beta_1|}$
2. $\tau_2(\theta) = E_p\left(\ln|Y|\right) = \dfrac{1}{\alpha}\Psi_{\beta_2}\left(\dfrac{\beta_2+1}{\alpha}\right) - \ln|\beta_1|^{1/\alpha}$ where $\theta = (\alpha, \beta_0, \beta_1, \beta_2)$.

Proof:

$$\int_{-\infty}^{\infty} e^{\beta_0} e^{-|\beta_1||w|^{\alpha}} |w|^{\beta_2}\, dw = 2e^{\beta_0}\int_{0}^{\infty} w^{\beta_2} e^{-|\beta_1| w^{\alpha}}\, dw = \frac{2e^{\beta_0}}{\alpha}\frac{\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}{|\beta_1|^{\frac{\beta_2+1}{\alpha}}},$$

So

$$e^{\beta_0} = \frac{\alpha|\beta_1|^{\frac{\beta_2+1}{\alpha}}}{2\Gamma\left(\frac{\beta_2+1}{\alpha}\right)},$$

In part (1), we have

$$E|W|^{\alpha} = \int_{-\infty}^{\infty} \frac{\alpha|\beta_1|^{\frac{\beta_2+1}{\alpha}}}{2\Gamma\left(\frac{\beta_2+1}{\alpha}\right)} |w|^{\alpha} e^{-|\beta_1||w|^{\alpha}} |w|^{\beta_2}\, dw = \frac{2}{\alpha}\frac{\alpha|\beta_1|^{\frac{\beta_2+1}{\alpha}}}{2\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}\frac{\Gamma\left(\frac{\beta_2+1}{\alpha}+1\right)}{|\beta_1|^{\frac{\beta_2+1}{\alpha}+1}} = \frac{\beta_2+1}{\alpha|\beta_1|}$$

Also

$$E\left(\ln|W|\right)=\int_{-\infty}^{\infty}\frac{\alpha\left|\beta_1\right|^{\frac{\beta_2+1}{\alpha}}}{2\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}\ln|w|\,e^{-|\beta_1||w|^{\alpha}}\,|w|^{\beta_2}\,dw=2\int_{0}^{\infty}\frac{\alpha\left|\beta_1\right|^{\frac{\beta_2+1}{\alpha}}}{2\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}\ln(w)e^{-|\beta_1|w^{\alpha}}\,w^{\beta_2}\,dw$$

Using integration by substitution, the above integration simplifies to

$$=\frac{1}{\alpha\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}\left[\int_0^{\infty}t^{\frac{\beta_2+1}{\alpha}-1}\ln(t)e^{-t}\,dt-\int_0^{\infty}t^{\frac{\beta_2+1}{\alpha}-1}\ln\left|\beta_1\right|e^{-t}\,dt\right]$$

Right side of the above equation simplifies to

$$\frac{1}{\alpha\Gamma\left(\frac{\beta_2+1}{\alpha}\right)}\left[\Gamma_{\beta_2}\left(\frac{\beta_2+1}{\alpha}\right)-\ln\left|\beta_1\right|\Gamma\left(\frac{\beta_2+1}{\alpha}\right)\right]=\frac{1}{\alpha}\left[\Psi_{\beta_2}\left(\frac{\beta_2+1}{\alpha}\right)-\ln\left|\beta_1\right|\right].$$

Theorem 3: The random variable W with a Bimodal Power Exponential Distribution with parameter θ has the maximum entropy among all positive, absolutely continuous random variables X with pdf $p(x)$ subject to the constraints

$$\tau_1(\theta)=E_p\left|X\right|^{\alpha} \tag{3.3}$$

$$\tau_2(\theta)=E_p\left(\ln|X|\right) \tag{3.4}$$

where $\tau_1(\theta)$ and $\tau_2(\theta)$ are defined in Lemma (1**Error! Reference source not found.**).

Proof:

Let P and W be probability measures, and let μ a Lebesgue measure in R, then by Jensen's inequality, we have

$$\int\frac{dP}{d\mu}\ln\frac{dP/dW}{dW/d\mu}d\mu\geq 0$$

Thus

$$\int\frac{dP}{d\mu}\ln\frac{dP}{d\mu}d\mu\geq\int\frac{dP}{d\mu}\ln\frac{dW}{d\mu}d\mu\,.$$

Consider the probability measure W, which satisfies the constraints in (3.3) and (3.4), and has the pdf of the form

$$dW(w)/d\mu=e^{\beta_0+\beta_1|w|^{\alpha}+\beta_2\ln(|w|)}\qquad -\infty<w<\infty \tag{3.5}$$

We have

$$
\begin{aligned}
-\int \frac{dP}{d\mu}\ln\frac{dP}{d\mu}\,d\mu &\le -\int \frac{dP}{d\mu}\ln\frac{dW}{d\mu}\,d\mu \\
&= -\int \frac{dP}{d\mu}\Big[\beta_0+\beta_1\,|w|^{\alpha}+\beta_2\,\ln(|w|)\Big]d\mu \\
&= -\beta_0+\beta_1\tau_1(\theta)+\beta_2\tau_2(\theta).
\end{aligned}
$$

The upper bound of the above entropy is achieved by the Bimodal Exponential Power distribution and the form of its pdf is (3.5).

3.2 Transformations

Theorem 4 If

$$U \sim Gamma\left(\frac{\delta+1}{\alpha},1\right),\ I=\begin{cases}1 & \textit{with prob. } \frac{1}{2}\\ -1 & \textit{with prob. } \frac{1}{2}\end{cases}.$$

I and U are independent and $X = IU^{\frac{1}{\alpha}}$, then $X \sim BEP(0,1,\delta,\alpha)$.

Proof: For $x > 0$,

$$
\begin{aligned}
P(X > x) &= P\left(IU^{\frac{1}{\alpha}} > x\right) \\
&= \tfrac{1}{2}P\left(U^{\frac{1}{\alpha}} > x\right) \\
&= \tfrac{1}{2}P\left(U > x^{\alpha}\right) \\
&= \tfrac{1}{2}\left[1-\int_0^{x^{\alpha}} f(u)du\right]
\end{aligned}
$$

where

$$f(u)=\frac{u^{\frac{\delta+1}{\alpha}-1}e^{-u}}{\Gamma\left(\frac{\delta+1}{\alpha}\right)},\quad 0<u<\infty .$$

So

$$F(x)=P(X<x)=\tfrac{1}{2}+\tfrac{1}{2}\int_0^{x^{\alpha}}\frac{u^{\frac{\delta+1}{\alpha}-1}e^{-u}}{\Gamma\left(\frac{\delta+1}{\alpha}\right)}du .$$

By the fundamental theorem of calculus, we have

$$f(x)=\frac{\alpha x^{\delta}e^{-x^{\alpha}}}{2\Gamma\left(\frac{\delta+1}{\alpha}\right)},\quad x>0 . \tag{3.6}$$

Similarly, for $x \le 0$,

$$F(x) = P(X \le x) = P\left(IU^{\frac{1}{\alpha}} \le x \right)$$
$$= \tfrac{1}{2} P\left(-U^{\frac{1}{\alpha}} \le x \right)$$
$$= \tfrac{1}{2} P(U \le \left(-x)^{\alpha}\right)$$
$$= \tfrac{1}{2} \int_{(-x)^{\alpha}}^{\infty} f(u)du$$

By the fundamental theorem of calculus, we have

$$f(x) = \frac{\alpha(-x)^{\delta} e^{-(-x)^{\alpha}}}{2\Gamma\left(\frac{\delta+1}{\alpha}\right)}, \quad x \le 0. \tag{3.7}$$

Finally, combine (3.6) and (3.7) to obtain

$$f(x) = \frac{\alpha \,|\, x \,|^{\delta} \; e^{-|x|^{\alpha}}}{2\Gamma\left(\frac{\delta+1}{\alpha}\right)}, \quad -\infty < x < \infty. \tag{3.8}$$

Corollary 1: If $V = \mu + \psi I U^{\frac{1}{\alpha}}$ then $V \sim BEP(\mu, \psi, \delta, \alpha)$.

Definition 1: A random variable Y is said to have a bimodal normal distribution with parameters μ and ψ if $Y \sim BEP(\mu, \psi, 2, 2)$ with the pdf

$$f(y) = \frac{2\left(\frac{y-\mu}{\psi}\right)^2 e^{-\left(\frac{y-\mu}{\psi}\right)^2}}{\psi\sqrt{\pi}}, \quad -\infty < y < \infty. \tag{3.9}$$

Letting $\psi^2 = 2\sigma^2$ in (3.9) gives $f(y) = \dfrac{\left(y-\mu\right)^2}{\sqrt{2\pi}\sigma^3} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \quad -\infty < y < \infty$

The notation $BN\left(\mu, \sigma^2\right)$ for bimodal normal will be used throughout the paper.

Theorem 5 (Standard Bimodal Normal): If X has a bimodal normal distribution and $Z = \dfrac{X-\mu}{\psi}$ then the pdf of Z is given by

$$f(z) = \frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty$$

with a moment generating function

$$M_Z(t) = \left(1 + t^2\right) e^{\frac{t^2}{2}}$$

It follows that the moment generating function of X is given by

$$M_X(t) = M_{\mu+\psi Z}(t) = e^{\mu t} M_Z(\psi t) = \left(1+\psi^2 t^2\right) e^{\mu t + \frac{\psi^2 t^2}{2}}.$$

From the moment generating function of the bimodal normal, we can see that, unlike the normal, the sum of independent bimodal normals is not a bimodal normal.

Theorem 6: If $Z \sim BN(0,1)$ and $W = Z^2$ then $W \sim \chi^2(3)$.

Proof:

$$P(W \le w) = P\left(-\sqrt{w} \le Z \le \sqrt{w}\right)$$

$$= \int_{-\sqrt{w}}^{\sqrt{w}} \frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$= 2\int_{0}^{\sqrt{w}} \frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

letting $t = z^2$ and using the fundamental theorem of calculus the pdf of W is as follows

$$g(w) = \frac{w^{\frac{3}{2}-1}}{2^{\frac{3}{2}}\Gamma\left(\frac{3}{2}\right)} e^{-\frac{w}{2}}, \quad 0 < w < \infty$$

which is the pdf of a Chi-square distribution with three degrees of freedom.

Theorem 7: If $Z_1 \sim BN(0,1)$ and $Z_2 \sim BN(0,1)$ are independent and $V = Z_1/Z_2$ then V is distributed as a bimodal Cauchy with the following pdf

$$f(v) = \frac{v^2}{2^3\pi\left(1+v^2\right)^3}, \quad -\infty < v < \infty. \tag{3.10}$$

Proof:

$$V = Z_1 / Z_2 \Rightarrow Z_1 = VZ_2$$

So

$$f(v) = \int_{-\infty}^{\infty} |z_2|\, f(vz_2, z_2)\, dz_2$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |z_2|\, z_2^4 v^2 e^{-\frac{z_2^2}{2}+\frac{vz_2^2}{2}} dz_2$$

$$= \frac{1}{\pi}\int_{0}^{\infty} z_2^5 v^2 e^{-\frac{z_2^2}{2}+\frac{vz_2^2}{2}} dz_2.$$

By using integration by substitution we obtain (3.10). It is important to note that, the bimodality is preserved, and the resulting distribution is a bimodal Cauchy.

Theorem 8: If $X \sim BN\left(\mu, \sigma^2\right)$ and $S = e^X$ then S is distributed as a log-normal of the second-type $SLog\left(\mu, \sigma^2\right)$ with the pdf

$$f(s) = \frac{(\log s - \mu)^2}{\sqrt{2\pi}\sigma^3 s} e^{-\frac{(\log s-\mu)^2}{2\sigma^2}}, \quad s > 0 .$$

It is interesting to see that the above log-normal is not bimodal but a unimodal. We call it a log-normal distribution of the second-type.

4. ESTIMATION

If $X \sim BEP(\mu, \psi, \delta, \alpha)$ and $z = \frac{x-\mu}{\psi}$, the loglikehood function can be written as

$$L = \log\alpha - \log 2 - \log\psi - \log\Gamma\left(\frac{\delta+1}{\alpha}\right) + \delta\log|z| - |z|^\alpha$$

and the elements of the score vector are

$$S_\mu = \frac{-\delta\, sign\left(z\right)}{\psi|z|} + \frac{\alpha}{\psi} sign\left(z\right)|z|^{\alpha-1}$$

$$S_\psi = \frac{\alpha}{\psi}|z|^\alpha - \frac{\delta+1}{\psi}$$

$$S_\alpha = \frac{\left(\delta+1\right)\psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha^2} - |z|^\alpha \log|z| + \frac{1}{\alpha}$$

$$S_\delta = \frac{\psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha} + \log|z| .$$

Thus, elements of the information matrix are given by

$$I_{\mu\mu} = \frac{\alpha\left(\alpha-1\right)}{\psi^2} E\left[|z|^{\alpha-1}\right] + \frac{\delta}{\psi^2} E\left(\frac{1}{|z|^2}\right)$$

$$I_{\psi\psi} = \frac{\delta+1}{\psi^2} - \frac{\alpha\left(\alpha+1\right)}{\psi^2} E\left(|z|^\alpha\right)$$

$$I_{\alpha\alpha} = \frac{-2\psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha^3} - \frac{\Psi'\left(\frac{\delta+1}{\alpha}\right)}{\alpha^4} - \frac{1}{\alpha^2} - E\left[|z|^\alpha \left(\log|z|\right)^2\right]$$

$$I_{\delta\delta} = \frac{\Psi'\left(\frac{\delta+1}{\alpha}\right)}{\alpha^2}$$

$$I_{\mu\psi} = -\frac{\alpha^2}{\psi^2} E\left[sign(z)|z|^{\alpha-1}\right]$$

$$I_{\mu\alpha} = \frac{E\left[sign(z)|z|^{\alpha-1}\right]}{\psi} + \frac{\alpha}{\psi} E\left[sign(z)|z|^{\alpha-1}\log|z| - E\right]\left[\frac{\delta\, sign(z)}{\alpha\psi|z|}\right]$$

$$I_{\mu\delta} = -E\left[\frac{sign(z)}{\psi|z|}\right]$$

$$I_{\psi\alpha} = -\frac{1}{\psi} E\left[|z|^\alpha\right] + \frac{\alpha}{\psi} E\left[|z|^\alpha \log|z|\right]$$

$$I_{\psi\delta} = \frac{-1}{\psi}$$

$$I_{\alpha\delta} = \frac{\Psi'\left(\frac{\delta+1}{\alpha}\right)}{\alpha^3} + \frac{1}{\alpha} - E\left[\log|z|\right].$$

So, the information matrix is given by

$$I(\theta) = \begin{pmatrix} \frac{\left(\alpha(\delta-1)+1\right)\Gamma\left(\frac{\delta-1}{\alpha}\right)}{\psi^2\Gamma\left(\frac{\delta+1}{\alpha}\right)} & 0 & 0 & 0 \\ 0 & \frac{\alpha(\delta+1)}{\psi^2} & -\frac{(\alpha+\delta+1)+(\delta+1)\Psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha\psi} & \frac{1}{\psi} \\ 0 & -\frac{(\alpha+\delta+1)+(\delta+1)\Psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha\psi} & I_{\alpha\alpha} & -\frac{(\delta+1)\left[\Psi'\left(\frac{\delta+1}{\alpha}\right)+\alpha\Psi\left(\frac{\delta+1}{\alpha}\right)\right]}{\alpha^3} \\ 0 & \frac{1}{\psi} & -\frac{(\delta+1)\left[\Psi'\left(\frac{\delta+1}{\alpha}\right)+\alpha\Psi\left(\frac{\delta+1}{\alpha}\right)\right]}{\alpha^3} & \frac{\Psi'\left(\frac{\delta+1}{\alpha}\right)}{\alpha^2} \end{pmatrix}$$

with

$$I_{\alpha\alpha} = \frac{1}{\alpha^2} + \frac{(\delta+1)(\delta+\alpha+1)\Psi'\left(\frac{\delta+1}{\alpha}\right)}{\alpha^4} + \frac{2(\delta+\alpha+1)\Psi\left(\frac{\delta+1}{\alpha}\right)}{\alpha^3} + \frac{(\delta+1)\Psi^2\left(\frac{\delta+1}{\alpha}\right)}{\alpha^3}$$

where Ψ and Ψ' are the digamma and the trigamma functions respectively. The information matrix does not exist for $\delta \le 1$.

5. SIMULATION RESULTS

A simulation study was conducted to assess the properties of the maximum likelihood estimates of the two parameters; α and δ using sample sizes of 50, 100, 200, 500 and 1000. For every case, 1000 samples from the bimodal exponential power with the specified parameters are drawn. By using Theorem (3.2) we can easily generate random samples from the $BEP(\mu,\psi,\delta,\alpha)$ through the following steps:

- Generate Y from $Gamma\left(\frac{\delta+1}{\alpha},1\right)$
- Generate U from Uniform (0,1)
- If $U \le 0.5$, set $I = 1$, else set $I = -1$
- Set $X = \mu + \psi I Y^{\frac{1}{\alpha}}$

The biases and the root mean squared errors (RMSEs) of the MLE estimators are shown in Tables 1-2. In Table (1), the samples were drawn from the standard bimodal exponential power distribution with $\mu = 0,\ \psi = 1,\ \alpha = (1,\ 2,\ 3,\ 4)$ and $\delta = 5$. In Table (2), the samples were drawn from the bimodal exponential power distribution with $\mu = 0$, $\psi = 1$, $\alpha = 2$ and $\delta =$ (1.5, 2, 2.5, 3). The biases of the maximum likelihood estimates; $\hat{\alpha}$ and $\hat{\delta}$ diminish for large samples while the RMSEs are quite large especially when $n \le 100$. The RMSEs increase with the value of □ but not with the value of δ. To get reliable estimates of α and δ, relatively large samples are required. To examine the accuracy of the standard errors obtained by inverting the observed information matrix and the coverage probability of the asymptotic confidence intervals, another simulation study was performed with 1000 samples using sample sizes of 50, 100, 200, 500 and 1000. The study focused on the two parameters; α and δ and samples were drawn from the standard bimodal exponential power distribution with $\alpha = \delta = 2$. The coverage probabilities of the obtained 95% confidence intervals are reported in Table (3) and are very close to the nominal level. The results suggested that the obtained standard errors and hence the asymptotic confidence intervals are reliable.

Table 1:
Biases and RMSEs of the maximum likelihood estimates from $BEP(0,1,5,\alpha)$

		$\hat{\alpha}$		$\hat{\delta}$	
α	n	Bias	RMSE	Bias	RMSE
1	50	0.0116	0.3649	0.2381	7.4726
	100	0.0070	0.2191	0.1535	4.8167
	200	0.0029	0.0911	0.0524	1.6452
	500	0.0009	0.0269	0.0205	0.6383
	1000	0.0007	0.0224	0.0122	0.3792
2	50	0.0320	0.9983	0.2463	7.6863
	100	0.0159	0.4961	0.1196	3.7369
	200	0.0075	0.2341	0.0616	1.9225
	500	0.0048	0.1494	0.0266	0.8220
	1000	-0.0003	0.0089	-0.0011	0.0333
3	50	0.0631	1.9758	0.2142	6.7139
	100	0.0273	0.8464	0.1014	3.1497
	200	0.0124	0.3878	0.0402	1.2528
	500	0.0055	0.1708	0.0096	0.2944
	1000	0.0039	0.1195	0.0143	0.4396
4	50	0.1092	3.4248	0.1910	5.9870
	100	0.0569	1.7769	0.1132	3.5355
	200	0.0140	0.4347	0.0354	1.0991
	500	0.0129	0.3979	0.0160	0.4928
	1000	0.0075	0.2282	0.0141	0.4316

Table 2:
Biases and RMSEs of the maximum likelihood estimates from $BEP(0,1,\delta,2)$

		$\hat{\alpha}$		$\hat{\delta}$	
δ	n	Bias	RMSE	Bias	RMSE
1.5	50	0.0766	2.4124	0.5957	3.0152
	100	0.0363	1.1453	0.5410	1.2916
	200	0.0137	0.4278	0.5204	0.6392
	500	0.0004	0.0130	0.5035	0.1098
	1000	0.0031	0.0955	0.5041	0.1263
2	50	0.0500	1.5710	0.0892	2.8041
	100	0.0186	0.5860	0.0523	1.6450
	200	0.0100	0.3154	0.0141	0.4429
	500	0.0046	0.1441	0.0077	0.2385
	1000	0.0032	0.0989	0.0034	0.1041
2.5	50	0.0439	1.3770	0.6260	3.9568
	100	0.0190	0.5986	0.5443	1.3924
	200	0.0081	0.2531	0.5287	0.8971
	500	0.0022	0.0690	0.5051	0.1588
	1000	0.0027	0.0824	0.5077	0.2382
3	50	0.0301	0.9438	0.0939	2.9384
	100	0.0196	0.6147	0.0748	2.3470
	200	0.0089	0.2765	0.0332	1.0357
	500	0.0038	0.1160	0.0203	0.6262
	1000	0.0022	0.0687	0.0094	0.2905

Table 3:
Coverage probability for the standard asymptotic 95% confidence intervals

Sample size	**Coverage Probability**	
	α	δ
50	0.9454	0.9464
100	0.9475	0.9556
200	0.9413	0.9342
500	0.9495	0.9475
1000	0.9592	0.9550

6. ILLUSTRATIVE EXAMPLES

6.1 Height Data

These data represent the height in inches of 126 students from University of Pennsylvania (Cruz-Medina 2001). Fig. (2a) shows that the heights distribution is roughly symmetric bimodal. Four models are fitted to the data; normal mixture with two components, bimodal exponential power, bimodal normal and bimodal Laplace $BEP(\mu, \psi, 1, 1)$. The maximum likelihood method is used to estimate the parameters in the fitted models. Table (4) displays the MLE estimates, the corresponding AIC (Akaike's Information Criterion) and Kolmogorov-Smirnov statistic for the fitted models. A histogram of the height data with the four fitted models is presented in Fig. (2a). The empirical CDF and the fitted CDFs of the four models are given in Fig. (2b). Fig. (2a) shows that the normal-mixture model failed to capture the bimodality of the data while the three models based on the bimodal exponential power distribution did capture the two modes. The bimodal exponential power model has the smallest K-S statistic and the second smallest AIC. The bimodal Laplace stands as a good competitor with less parameters and has the smallest AIC. The bimodal normal and the normal mixture did not perform well. Based on the graphical and the numerical results, the bimodal exponential power distribution is considered the best model for the heights data.

Table 4:
Estimates, AIC, and KS statistic (P-value) for Height data

Model	Estimated parameters	AIC	K-S statistic (P-value)
Normal Mixture	$\hat{p} = 0.8119$ $\hat{\mu}_1 = 67.56, \hat{\sigma}_1 = 3.75$ $\hat{\mu}_2 = 72.80, \hat{\sigma}_2 = 2.84$	367.668	0.0547 (0.8220)
BEP	$\hat{\mu} = 68.43, \hat{\psi} = 2.67$ $\hat{\alpha} = 1.22, \hat{\delta} = 0.73$	361.54	0.0503 (0.8882)
Bimodal Laplace	$\hat{\mu} = 68.42, \hat{\psi} = 1.70$	357.71	0.0550 (0.8170)
Bimodal Normal	$\hat{\mu} = 68.44, \hat{\sigma} = 2.39$	398.61	0.1189 (0.0520)

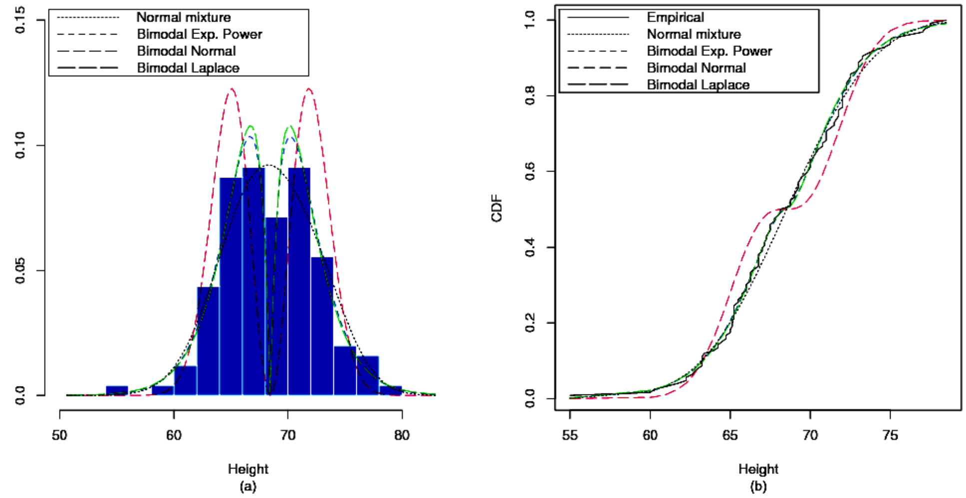

Fig. 2: (a) Histogram and density estimates
(b) Empirical CDF and the CDF's of the fitted models for the heights data using normal mixtures, Bimodal Exponential Power, Bimodal Normal and Bimodal Laplace, respectively

6.2 Egg Size Data

These data represent the logarithm of the egg diameters of 88 asteroid species (Sewell and Young (1997), Famoye et al. (2004)). The distribution of the data is roughly symmetric bimodal as shown in Fig. (3a). The normal mixture and two different cases of the bimodal exponential power models are fitted to the data. The bimodal exponential power models are $BEP(\mu,\psi,\delta,\alpha)$ and $BEP(\mu,\psi,2,1)$. Table (5) gives the estimates, the corresponding AIC and Kolmogorov-Smirnov test for the fitted models. Fig. (3a) presents the egg size data with the fitted models while the empirical CDF with the CDFs of these fitted models are given in Fig. (3b). Fig. (3a) shows that the normal mixture fitted well at the tails but did very badly in the middle. The bimodal exponential power models did fit the data adequately in middle but not in the tails. The normal mixture model has the smallest AIC and K-S statistic. The $BEP(\mu,\psi,1,2)$ model has the second lowest K-S value while the general BEP model tends to have the second lowest AIC.

Table 5:
Estimates, AIC, and KS statistic for asteroid data

Model	Estimated parameters	AIC	K-S statistic (P-value)
	$\hat{p}=0.3875$		
Normal Mixture	$\hat{\mu}_1=5.00, \hat{\sigma}_1=0.223$	111.31	0.0714 (0.7312)
	$\hat{\mu}_2=6.75, \hat{\sigma}_2=0.606$		
BEP	$\hat{\mu}=5.86, \hat{\psi}=1.08$	121.18	0.1368 (0.0667)
	$\hat{\alpha}=2.18, \hat{\delta}=0.95$		
$BEP(\mu,\psi,1,2)$	$\hat{\mu}=6.02, \hat{\psi}=0.98$	125.44	0.0952 (0.3805)

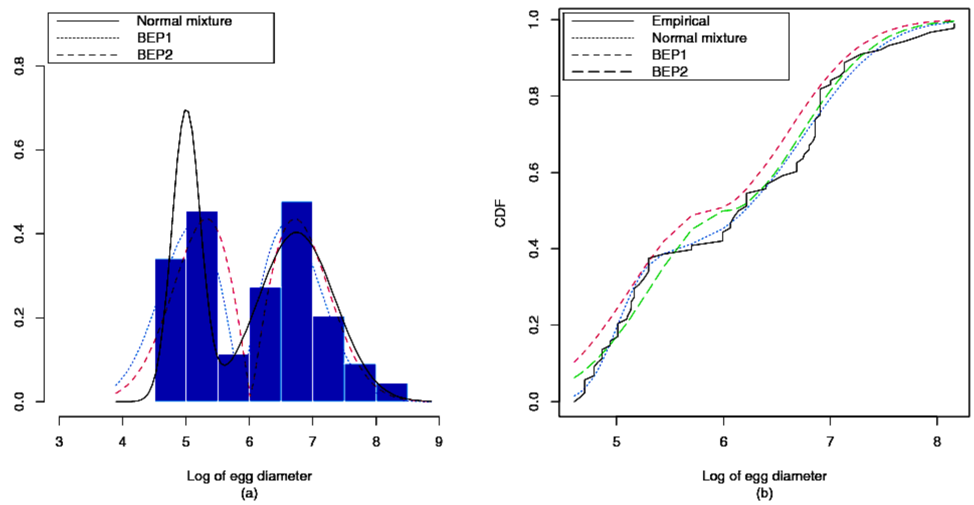

Fig. 3: (a) Histogram and density estimates
(b) Empirical CDF and the CDF's of the fitted models for the asteroid data using normal mixtures, BEP1 (Bimodal Exponential Power) and BEP2 (BEP with $\delta = 1$ and $\sigma = 2$)

7. CONCLUSION AND POSSIBLE EXTENSIONS OF BEP

In this paper we have proposed a family of flexible bimodal distributions that can be used as an alternative to the normal mixtures in modeling bimodal data. This family is more parsimonious than the normal mixtures and has the advantage of estimation simplicity over the mixtures. The family, which we call the bimodal exponential power distribution can be extended to a bimodal skew exponential power distribution, since if $q(x) \propto 2f(x)F(x\lambda)$, where $f(x)$ and $F(x)$ are the density and the distribution functions of the symmetric bimodal exponential distribution respectively then $q(x)$ is a skewed density function, and λ is a skewness parameter. The characterization of the bimodal exponential distribution is investigated. Simulation examples did show that such estimation procedures performed well. Real data examples demonstrated that bimodal exponential power distribution is flexible, parsimonious, and competitive model that deserves to be added to existing distributions in modeling bimodal data.

REFERENCES

1. Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.
2. Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distributions. *J. Roy. Statist. Soc., Series B*, 61, 579-602.
3. Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83, 715-726.
4. Bhagavan, C.S. et al. (1983). A generalization of normal distribution. Paper presented at *ISTPA Annual Conference*, Varanasi.

5. Box, G. (1953). A note on regions for tests of kurtosis. *Biometrika*, 40, 465-468.
6. Box, G. and Tiao, G. (1992). *Bayesian inference in statistical analysis*. John Wiley and Sons, New York.
7. Cruz-Medina, I.R. (2001). Almost nonparametric and nonparametric estimation in mixture models. *Ph.D. Dissertation*, Pennsylvania State University, University Park.
8. DiCiccio, T. and Monti, A. (2004). Inferential aspects of the skew exponential power distribution. *J. Amer. Statist. Assoc.*, 99, 439-450.
9. Dierickx, B., Vleugels, J. and Van der Biest, O. (2000). Statistical extreme value modeling of particle size distributions: experimental grain size distribution type estimation and parameterization of sintered zirconia D. *Materials Characterization*, 45, 61-70.
10. Ely, J.T.A., Fudenberg, H.H., Muirhead, R.J., LaMarche, M.G., Krone, C.A., Buscher, D., and Stern, E.A. (1999). Urine mercury in micromercurialism: bimodal distribution and diagnostic implications. *Bulletin of environmental contamination and toxicology*, 63(5), 553-59.
11. Famoye, F., Lee, C., and Eugene, N. (2004). Beta-Normal Distribution: Bimodality Properties and Application. *Journal of Modern Applied Statistical Methods,* 3(1), 85-103.
12. Gupta, R.C. and Brown, N. (2001). Reliability studies of the skew-normal distribution and its application to a strength-stress model. *Commun. in Statist.-Theory Meth.*, 30, 2427-2445.
13. Gupta, R.C. and Gupta, R.D. (2004). Generalized skew normal model. *Test*, 13, 501-525.
14. Mudholkar, G.S. and Hutson, A. (2000). The epsilon-skew-normal distribution for analyzing near-normal data. *J. Statist. Plann. and Infer.*, 83(2), 291-309.
15. Prasada Rao, R.M. (1987). Some generalizations of the normal distribution. *Unpublished Ph.D. Thesis* of Andhra University, Waltair, India.
16. Rao, K.S. et al. (1987). On a bivariate bimodal distribution. Paper presented at *ISTPA Annual Conference*, Delhi.
17. Rao, K.S., Narayana, J.L. and Papayya Sastry, V. (1988). A bimodal distribution. *Bulletin of the Calcutta Mathematical Society*, 80(4), 238-240.
18. Sarma, P.V.S., Rao, K.S. Srinivasa and Rao, R. Prabhakara (1990). On a family of bimodal distributions. *Sankhya, Series B*, 52, 287-292.
19. Sewell, M., and Young, C. (1997). Are echinoderm egg size distributions bimodal? *Biological Bulletin,* 193, 297-205.
20. Subbotin, M.T. (1923). On the law of frequency of error. *Matematicheskii Sbornik*, 31, 296-301.
21. Turner, M.E. (1960). On Heuristic Estimation Methods. *Biometrics*, 16(2), 299-302.
22. Zhang, C., Mapes, B.E., and Soden, B.J. (2003). Bimodality in tropical water vapor. *Quarterly Journal of the Royal Meteorological Society*, 129, 2847-2866.