Article

Asymmetric Bimodal Exponential Power Distribution on Real Line

Mehmet Niyazi Çankaya [1]*

[1] Applied Sciences School, Department of International Trading & Faculty of Art and Sciences, Department of Statistics, Uşak University, Uşak, Turkey
* Correspondence: mehmet.cankaya@usak.edu.tr; Tel.: +90-276-221-21-21 (ext. 2585)

Abstract: The asymmetric bimodal exponential power (ABEP) distribution is an extension of the generalized gamma distribution to the real line via adding two parameters which fit the shape of peakedness in bimodality on real line. The special values of peakedness parameters of the distribution are combination of half Laplace and half normal distributions on real line. The distribution has two parameters fitting the height of bimodality, so capacity of bimodality is enhanced by using these parameters. Adding a skewness parameter is considered to model asymmetry in data. The location-scale form of this distribution is proposed. The Fisher information matrix of these parameters in ABEP is obtained explicitly. Properties of ABEP are examined. Real data examples are given to illustrate the modelling capacity of ABEP. The replicated artificial data from maximum likelihood estimates of parameters of ABEP and distributions having an algorithm for artificial data generation procedure are provided to test the similarity with real data.

Keywords: asymmetric bimodality; bimodal exponential power distribution; modelling; generalized Gaussian distribution.

---

1. Introduction

The different bimodal and skew distributions have been proposed over the last decade to construct flexible distributions. The proposed distributions are Refs. [1–26] and references therein via using different generating techniques [27] to get a probability density function (PDF). In these distributions, Refs. [5,6] proposed ε-skew form of gamma distribution on real line. The deficiency of these functions is that different height and shape of peakedness around location on real line cannot be modelled separately. The model proposed by [6] has a bimodality with the same height, which is not flexible enough to model bimodal data with different height and shape of peakedness. Ref. [24] proposed bimodal and alpha-skew Laplace distribution that does not model shape peakedness around location on real line. However, the best way is to find a function which can fit data around location separately. In other words, the left and right sides of location will be modelled with different parameters to have an efficient fitting for both sides of location. A bimodal exponential power (BEP) distribution is proposed by [28]. The properties of BEP distribution are few when BEP is compared with distribution proposed by [29], because BEP has same level of peaks around location on real line and it is also symmetric in both side of location. The shape of peakedness around location on real line is modelled by only one parameter, however two parameters are added to model different modes from distribution on real line [29]. Two parameters controlling to fit the shape of peakedness and two parameters controlling to fit the height of bimodality will be used together. Skewness parameter is also added to model asymmetry in data. Thus, modelling capacity of asymmetric bimodal exponential power (ABEP) distribution is better than current candidates proposed by [5,6,28,29], because ABEP distribution has parameters that control the fitting both sides of location separately.

The second aim is that we do not only propose ABEP distribution but also derive this distribution via constructing a normalizing constant (NC) which leads to produce a PDF. While deriving a PDF, producing NC can be a preferable approach. This approach can be taken care for deriving a PDF when one wants to add a new parameter to increase the modelling capacity of function if it is tractable to get

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 23; doi:10.3390/e20010023

NC from a function. The NC approach was examined by [30] to construct asymmetric distributions from symmetric distributions. Some techniques used to derive a PDF are reviewed by [27]. There are other techniques to produce PDFs derived from entropy functions via method of Lagrange multipliers as well [31,32] and references therein. The different goodness of fit tests (GOFTs) are applied on the ABEP. Thus, importance and advantage of GOFTs, such as Kolmogorov-Smirnov (KS), Cramér von Mises (CVM), Anderson-Darling (AD) via a cumulative distribution function (CDF) of a PDF will be expressed for ABEP distribution when the optimization problem of ABEP can arise.

Especially, the estimation of location parameter is important, such as the proteins in cancer cell are needed to determine, the image processing demands to get the quantitative value of colors at a prescribed range. A radar data, speech processing, etc. in many phenomena can be modelled via ABEP. The parametric models which can accommodate the shape of peakedness, bimodality and skewness are mostly preferred to be able to model the data set efficiently. In other words, the frequented data can be represented by the parameters which control to fit the shape of peakedness, the parameters which control to fit the bimodality and the skewness which controls to fit the asymmetry in data set. Due to this reason, ABEP distribution having these parameters is proposed. In addition to, since the generalized gamma distribution is a class for many distributions, it is chosen in order to reflect to the negative side of real line.

The paper is organized as follows. In Section 2, ABEP distribution is defined and mode, distributional properties, related distributions and tail behaviour of ABEP distribution are given. Maximum likelihood (ML) estimations of parameters are provided in Section 3. In Section 4, the real data examples are provided to make a comparison among candidate densities. The results are commented. Finally, in last section the conclusions are given and the remarks are considered.

2. Gamma Distribution: Reparametrization and ABEP Distribution on Real Line

The random variable Y will have a gamma distribution with PDF having parameters $\frac{\delta+1}{\alpha}$ and $\beta = 1$:

$$g(y) = \frac{1}{\Gamma(\frac{\delta+1}{\alpha})} y^{\frac{\delta+1}{\alpha}-1} \exp\{-\frac{y}{\beta}\}, \quad y > 0,\ \delta > 0,\ \alpha > 0. \tag{1}$$

Theorem 1. *Let Y be a continuous random variable defined on $[0,\infty)$, distributed as $G(\frac{\delta+1}{\alpha}, \beta = 1)$. Consider a discrete random variable T. It generates a function on real line and unequal probabilities at negative and positive sides of real line will be constructed. T is $1+\varepsilon$ for the probability $\frac{1+\varepsilon}{2}$ at positive side and T is $-(1-\varepsilon)$ for the probability $\frac{1-\varepsilon}{2}$ at negative side. A variable transformation $Z = Y^{1/\alpha}T$ is applied to get the α power of Gamma distribution. Here, the random variables Y and T are independent [5,29]. After applying this transformation on gamma distribution in equation (1), we will get the following PDF:*

$$f(z) = \begin{cases} f_1(z) = \frac{\alpha}{2(1-\varepsilon)^{\delta}\Gamma(\frac{\delta+1}{\alpha})}(-z)^{\delta} \exp\left\{-\left(\frac{-z}{1-\varepsilon}\right)^{\alpha}\right\}, & z < 0 \\ f_0(z) = \frac{\alpha}{2(1+\varepsilon)^{\delta}\Gamma(\frac{\delta+1}{\alpha})} z^{\delta} \exp\left\{-\left(\frac{z}{1+\varepsilon}\right)^{\alpha}\right\}, & z \geq 0, \end{cases} \tag{2}$$

with the parameters $\alpha > 0, \delta > 0$ and $\varepsilon \in (-1,1)$ [29]. The random variable T keeps to be PDF that will be generated, because the gamma distribution is a PDF on $[0,\infty)$. The probabilities of $(1+\varepsilon)$ and $-(1-\varepsilon)$ values of random variable T are $\frac{1+\varepsilon}{2}$ and $\frac{1-\varepsilon}{2}$ [30,33]. Thus, a function in equation (2) has the unequal probabilities at positive and negative sides of real line. The following PDF from function in equation (2) will be proposed:

$$f(z) = \begin{cases} f_1(z) = \frac{\alpha_1}{2[k_1(1-\varepsilon)]^{\delta_1+1}\Gamma(\frac{\delta_1+1}{\alpha_1})}(-z)^{\delta_1} \exp\left\{-\left(\frac{-z}{k_1(1-\varepsilon)}\right)^{\alpha_1}\right\}, & z < 0 \\ f_0(z) = \frac{\alpha_0}{2[k_0(1+\varepsilon)]^{\delta_0+1}\Gamma(\frac{\delta_0+1}{\alpha_0})} z^{\delta_0} \exp\left\{-\left(\frac{z}{k_0(1+\varepsilon)}\right)^{\alpha_0}\right\}, & z \geq 0, \end{cases} \tag{3}$$

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

with the parameters $\alpha_1 > 0, \alpha_0 > 0, \delta_1 > 0, \delta_0 > 0, k_1 > 0, k_0 > 0$ *and* $\varepsilon \in (-1, 1)$. *Without consulting the variable transformation technique, PDF can be obtained. This PDF is called as an asymmetric bimodal exponential power distribution (ABEP).* α_1 *and* α_0 *are for the shape of peakedness,* δ_1 *and* δ_0 *are for height of bimodality at negative and positive sides of real line.* k_1 *and* k_0 *are nuisance parameters to have same form of normal or Laplace distributions.* ε *is a skewness parameter that is responsible to have unequal probabilities at negative and positive sides of real line. Thus, a skewness on a function can be constructed. The details for function in equation (3) are given by the following proof.*

Proof. The preliminary tools for the calculation of integrals are required. The gamma function and the incomplete gamma functions are used to have the integral kernels which are appropriate to calculate the integrals. Thus, we can derive a PDF.

$$\Gamma(s) = \gamma(s, \alpha) + \Gamma(s, \alpha), \tag{4}$$

where $\Gamma(s) = \int_0^\infty x^{s-1} \exp\{-x\} \mathrm{d}x$, $\gamma(s, \alpha) = \int_0^\alpha x^{s-1} \exp\{-x\} \mathrm{d}x$, and $\Gamma(s, \alpha) = \int_\alpha^\infty x^{s-1} \exp\{-x\} \mathrm{d}x$. These are the gamma, the lower and upper incomplete gamma functions, respectively [34].

The reparametrization of gamma function is considered as:

$$\Gamma(s + 1/\alpha) = \int_0^\infty x^{s+1/\alpha-1} \exp\{-x\} \mathrm{d}x. \tag{5}$$

A variable transformation $x = (yp)^\alpha$ is applied to get the power version of gamma function:

$$\Gamma(s + 1/\alpha) = \alpha p^{\alpha s+1} \int_0^\infty y^{\alpha s} \exp\{-(yp)^\alpha\} \mathrm{d}y. \tag{6}$$

From equation (4), $\gamma(s^*, \alpha^*) = \Gamma(s^*) - \Gamma(s^*, \alpha^*)$. Now, let s^* be $s + 1/\alpha$ and $\alpha^* = (pk)^\alpha$. Then, $\gamma(s + 1/\alpha, (pk)^\alpha) = \int_0^{(pk)^\alpha} x^{s+1/\alpha-1} \exp\{-x\} \mathrm{d}x$. Now, the variable transformation $x = (yp)^\alpha$ is applied to the power version of the lower incomplete gamma function:

$$\gamma(s + 1/\alpha, (pk)^\alpha) = \alpha p^{\alpha s+1} \int_0^k y^{\alpha s} \exp\{-(yp)^\alpha\} \mathrm{d}y. \tag{7}$$

From equation (4), $\Gamma(s^*, \alpha^*) = \Gamma(s^*) - \gamma(s^*, \alpha^*)$. Now, let s^* be $s + 1/\alpha$ and $\alpha^* = (pk)^\alpha$. Then, $\Gamma(s + 1/\alpha, (pk)^\alpha) = \int_{(pk)^\alpha}^\infty x^{s+1/\alpha-1} \exp\{-x\} \mathrm{d}x$. Now, the variable transformation $x = (yp)^\alpha$ is applied to the power version of the upper incomplete gamma function:

$$\Gamma(s + 1/\alpha, (pk)^\alpha) = \alpha p^{\alpha s+1} \int_k^\infty y^{\alpha s} \exp\{-(yp)^\alpha\} \mathrm{d}y. \tag{8}$$

The equations (6)-(8) are power versions of gamma functions defined on the positive axis. These three functions can be transferred to the negative axis via the variable transformation $y = -u$. For equation (6),

$$\Gamma(s + 1/\alpha) = \alpha p^{\alpha s+1} \int_{-\infty}^0 (-u)^{\alpha s} \exp\{-(-up)^\alpha\} \mathrm{d}u. \tag{9}$$

For equation (7),

$$\gamma(s + 1/\alpha, (pk)^\alpha) = \alpha p^{\alpha s+1} \int_{-k}^0 (-u)^{\alpha s} \exp\{-(-up)^\alpha\} \mathrm{d}u. \tag{10}$$

For equation (8),

$$\Gamma(s + 1/\alpha, (pk)^\alpha) = \alpha p^{\alpha s+1} \int_{-\infty}^{-k} (-u)^{\alpha s} \exp\{-(-up)^\alpha\} \mathrm{d}u. \tag{11}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

For two cases of $x < 0$ and $x \geq 0$, we have the integrals of equation (3). So, equation (6) and equation (9) can be used to calculate these integrals. One can easily show that the integrated value of negative and positive sides of equation (3) are $1/2$, respectively. Due to the fact that we must have a PDF defined on the real line, the summation of these two results is 1. Here, the variable transformation technique is not used. Thus, we can guarantee that the function gotten is on the interval $[0, 1]$. It is well known that if a function is defined on the interval $[0, 1]$, this function will be a PDF. □

The location-scale form of this distribution is given by the following form. Suppose that Z is distributed as ABEP$(\alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$. Then, the random variable $X = \mu + \sigma Z$, $\mu \in \mathbb{R}$ and $\sigma > 0$ will have ABEP distribution with the following density function:

$$g(x) = \begin{cases} g_1(x) = \frac{\alpha_1}{2\sigma[k_1(1-\varepsilon)]^{\delta_1+1}\Gamma(\frac{\delta_1+1}{\alpha_1})}(-\frac{x-\mu}{\sigma})^{\delta_1}\exp\{-\left[\frac{-(x-\mu)}{\sigma k_1(1-\varepsilon)}\right]^{\alpha_1}\}, & x < \mu \\ g_0(x) = \frac{\alpha_0}{2\sigma[k_0(1+\varepsilon)]^{\delta_0+1}\Gamma(\frac{\delta_0+1}{\alpha_0})}(\frac{x-\mu}{\sigma})^{\delta_0}\exp\{-\left[\frac{x-\mu}{\sigma k_0(1+\varepsilon)}\right]^{\alpha_0}\}, & x \geq \mu, \end{cases} \tag{12}$$

where μ and σ are the location and the scale parameters, respectively. Here, the random variable X is distributed as ABEP$(\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$, that is, $X \sim$ ABEP$(\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$.

2.1. Properties of ABEP Distribution

2.1.1. Mode of a kernel function in ABEP

The mode of function in equation (12) is examined. It is obvious that this function is a reflected function in equation (3) from the reparameterized gamma function in equation (1). Thus, examining the mode of positive side of equation (3) means that the negative side of equation (3) is also examined. Now, it is examined whether or not there is one root of the following function:

$$h(t) = t^{\delta_0}\exp\{-t^{\alpha_0}\}, \quad t > 0,\ \delta_0 > 0,\ \alpha_0 > 0. \tag{13}$$

Here, we will give comments about getting root of this function: NC can be ignored, because NC produces a function at interval $[0, 1]$. It does not affect the modes of function. At the same way, the location parameter μ can be ignored, because the location shows where the function in equation (12) is located. The scale σ and its variants k_0 or k_1 and ε parameters change the rescaling of function in equation (12).

The root of derivative of function in equation (13) with respect to t is $\exp\{\alpha_0^{-1}\log(\alpha_0^{-1}\delta_0)\}$. For $t = 0$, $h(t) = 0$, which is obvious root that does not lead to modality. Thus, there is only one root of function in equation (13), that is, there is one mode of function of generalized gamma at positive side. Since it is reflected to negative side of real line, the function has a mode at negative side of real line. Totally, this function in equation (12) has two modes at real line. Note that it is not necessary to use second derivative test, because maximization of a function is equivalent to minus minimization of that function. Detecting the root is enough for having modality.

2.1.2. Cumulative distribution function of ABEP distribution

Let $X \sim$ ABEP$(\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$. Let G be CDF of PDF g. Then, CDF of the random variable X is:

$$G(x) = \begin{cases} G_1(x) = \frac{1}{2\Gamma\left(\frac{\delta_1+1}{\alpha_1}\right)}\Gamma\left(\frac{\delta_1+1}{\alpha_1}, (\frac{-(x-\mu)}{k_1\sigma(1-\varepsilon)})^{\alpha_1}\right), & x < 0 \\ G_0(x) = \frac{1+\varepsilon}{2} + \frac{1}{2\Gamma\left(\frac{\delta_0+1}{\alpha_0}\right)}\gamma\left(\frac{\delta_0+1}{\alpha_0}, (\frac{x-\mu}{k_0\sigma(1+\varepsilon)})^{\alpha_0}\right), & x \geq 0, \end{cases} \tag{14}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

where γ and Γ are the lower and upper incomplete gamma functions, respectively.

2.1.3. rth moment of random variable X distributed as ABEP

Let $X \sim \text{ABEP}(\mu = 0, \sigma = 1, \alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$. The rth, $r \geq 0$, non-central moment is given by

$$\mathbb{E}(X^r) = \frac{[k_1(1-\varepsilon)]^r \Gamma\left(\frac{\delta_1+r+1}{\alpha_1}\right)}{2\Gamma\left(\frac{\delta_1+1}{\alpha_1}\right)} + \frac{[k_0(1+\varepsilon)]^r \Gamma\left(\frac{\delta_0+r+1}{\alpha_0}\right)}{2\Gamma\left(\frac{\delta_0+1}{\alpha_0}\right)}. \tag{15}$$

One can get the results via equations (6) and (9). Since $\mathbb{E}(X^r)$ is finite for finite values of parameters $\alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0$ and when the extremely big values of parameters $\alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0$ and r are not taken, the ABEP distribution can produce finite values for the estimates of parameters, because finiteness of moments gurantees to have a finite value of function [35]. Note that the domain of skewness parameter ε is the interval $(-1, 1)$.

2.1.4. Moment generating function for random variable X distributed as ABEP

Let $X \sim \text{ABEP}(\mu = 0, \sigma = 1, \alpha_1, \alpha_0, \delta_1, \delta_0, k_1, k_0, \varepsilon)$. The moment generating function of the random variable X is:

$$\mathbb{E}[\exp(tX)] = \sum_{m=0}^{\infty} \left[\frac{t^m [k_1(1-\varepsilon)]^m \Gamma\left(\frac{\delta_1+m+1}{\alpha_1}\right)}{2\Gamma\left(\frac{\delta_1+1}{\alpha_1}\right) m!} + \frac{t^m [k_0(1+\varepsilon)]^m \Gamma\left(\frac{\delta_0+m+1}{\alpha_0}\right)}{2\Gamma\left(\frac{\delta_0+1}{\alpha_0}\right) m!} \right], \tag{16}$$

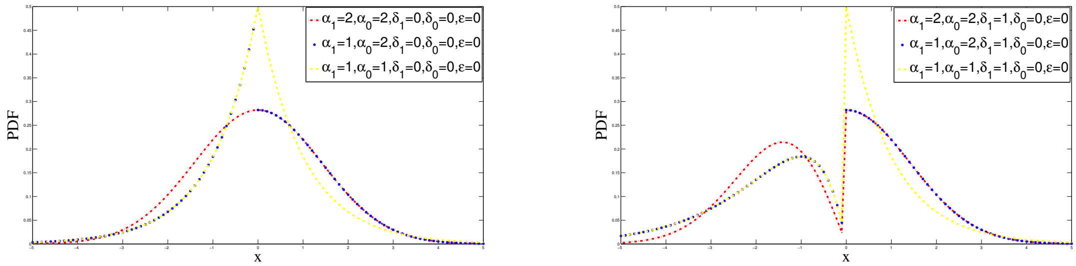
where $t \in \mathbb{R}$ and $m \in \mathbb{N}$. In order to calculate the integral $\mathbb{E}[\exp(tX)]$, the Taylor expansion at $x = 0$ of the function $\exp(tx) = \sum_{m=0}^{\infty} \frac{(tx)^m}{m!}$ must be gotten. After some straightforward calculation for the integral $\mathbb{E}[\exp(tX)]$ via using equations (6) and (9), the result of integral can be obtained.

2.1.5. PDFs for different values of parameters in ABEP

Figures 1 and 2 illustrate the examples of PDF of ABEP distribution for some values of parameters that give all possible shapes of function. It is seen from these figures, the shape of peakedness, bimodality and asymmetry can be controlled at the same time via parameters in ABEP. When the different values of parameters α_1, α_0 and δ_0, δ_1 are chosen, the different shape of peakedness and the bimodality with different height around location parameter μ are obtained, respectively. The skewness parameter ε makes an asymmetry around parameter μ.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

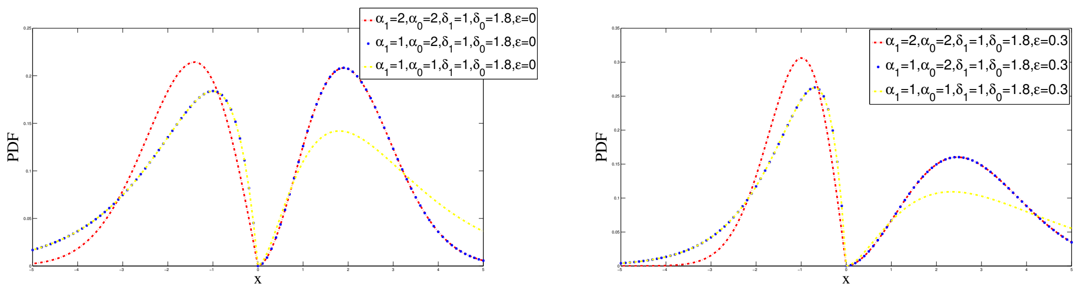
Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023



(a) Unimodal densities due to $\delta_1 = \delta_0 = 0$, examples for normal and Laplace and their half forms due to $\alpha_1 > 0$ and $\alpha_0 > 0$.

(b) Bimodal densities due to $\delta_1 > 0$, right of density is normal and Laplace due to $\delta_0 = 0$ and α_0.

Figure 1. Examples of PDFs of the ABEP distribution for the different values of parameters ($\mu = 0, \sigma = 1$): Unimodality, bimodality, half of Laplace and half of normal.



(a) Bimodal densities constructed via $\delta_1 > 0$, $\delta_0 > 0$, $\alpha_1 > 0$ and $\alpha_0 > 0$.

(b) Bimodal densities via $\delta_1 > 0$, $\delta_0 > 0$, $\alpha_1 > 0$ and $\alpha_0 > 0$ with skewed form: The left and right sides of location have unequal probabilities due to ε.

Figure 2. Examples of PDFs of the ABEP distribution for the different values of parameters ($\mu = 0, \sigma = 1$): Bimodality.

2.1.6. Tail behaviour property of ABEP

Tail behaviour or heavy tailedness of a distribution is examined by means of definitions given below [36]:

Definition 1. *Let $\bar{G}(x)$ be $1 - G(x)$. If $\lim_{x \to +\infty} \exp(\lambda x)\bar{G}(x) = \infty$ for all $\lambda > 0$, then $G(x)$ is a heavy-tailed.*

From equation (14), the positive part of CDF includes the lower incomplete gamma function γ. The function $\gamma(a, b)$ is examined to get the limit in Definition 1. For $b > a$, this function goes to zero. Then, $\lim_{x \to +\infty} \exp(\lambda x)G(x)$ can go to zero when b is more bigger than a. Otherwise, this limit is infinite. If $\lim_{x \to +\infty} \exp(\lambda x)G(x) \to 0$, then $\lim_{x \to +\infty} \exp(\lambda x)\bar{G}(x) \to \infty$ for $b > a$ in γ function.

$\lim_{x \to +\infty} \exp(\lambda x)\bar{G}(x)$ is undefined for a case $a \geq b$. It is seen that when b as a variable x of the function γ has big values, that is, an outlier is included by data, the heavy-tailedness property of ABEP can be obtained. For $a \geq b$, there is already a tendency to get small values of variable x in γ function in equation (14), which does not correspond an outlier in data set when it is compared with case $b > a$ in γ function. Thus, having an undefined value for $\lim_{x \to +\infty} \exp(\lambda x)\bar{G}(x)$ is not problem in order to test the heavy-tailedness property of function G via Definition 1.

Definition 2. *Suppose that random variable X has a PDF g defined on $[0, \infty)$. If $\mathbb{E}[\exp(tX)] = \infty$, for all t, then g is a heavy-tailed.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 23; doi:10.3390/e20010023

Note that the generalized gamma distribution is reflected to negative axis or $x < \mu$. The tail behaviour at $x > \mu$ or $x < \mu$ has a same role. Then, Definition 2 can be used for ABEP.

From equation (16), $\mathbb{E}[\exp(tX)] = \infty$ is satisfied due to m in summation in equation (16) of ABEP distribution, because m goes to infinity and Γ function gives infinity for big values of m. Then, ABEP is a heavy-tailed distribution.

A comment for heavy-tailedness from the results of Definitions 1 and 2 is given: The skewness parameter ε and also shape parameters $\alpha_1, \alpha_0, \delta_1, \delta_0$ work together in order to get a heavy-tailed function, because they are responsible to change the shape of function.

2.2. Special Cases, Related Distributions and Flexibility of ABEP

When we want to make a comparison among them from lowest to highest for capacity on modelling frequency, ordered form is Refs. [28,29] and ABEP distribution. For this aim, ABEP distribution is defined by using the generalized gamma distribution. The resulting distribution has five parameters. Thus, ABEP distribution will have some properties: when $\alpha_1 = 1$ and $\alpha_0 = 2$, left side of location is half of Laplace distribution and right side of location is half of normal distribution for $\varepsilon = 0$ and $\delta_1 = \delta_0 = 0$. For values of $\alpha_1 = 2$ and $\alpha_0 = 1$, the resulting function will be vice versa of previous case. For these situations, when $\varepsilon \neq 0$, ABEP will be ε-skew form of half from Laplace and normal distributions. It is easily seen that ABEP distribution can be a combination of Laplace and normal distributions for values of peakedness parameters α_1 and α_0 of distribution in ε-skew form. The nuisance parameters k_1 and k_0 are added to have same form of normal and Laplace distributions. The location-scale form is also provided. The parameters α_1, δ_1 and α_0, δ_0 also determine the overall shape of function for $x < \mu$ and $x \geq \mu$, respectively. Tails at negative and positive sides of real line can be platykurtic ($\alpha_1, \alpha_0 \to \infty$) and leptokurtic ($\alpha_1, \alpha_0 \to 0$). The special cases, related distributions and flexibility of ABEP distribution are given in the following items:

1. When $\alpha_1 = \alpha_0 = \alpha > 0$, ABEP distribution drops to the kernel of distribution in [29] for $\beta = 1$.
2. If $\delta_0 = \delta_1 = \delta > 0$, the density function has two modes (bimodal case) with the same height. If $\delta_0 = \delta_1 = 0$, the distribution is a unimodal.
3. When $\varepsilon = 0$, the distribution is the symmetric with two different modes.
4. When $\alpha_1 = \alpha_0 = 2$, $\delta_1 = \delta_0 = 0$, $k_1 = k_0 = 2$ and $\varepsilon = 0$, the distribution is a standard normal distribution. Location $\mu \in \mathbb{R}$, scale $\sigma > 0$ and $k_1 = k_0 = 2$ case of ABEP distribution is defined in equation (12).
5. When $\alpha_1 = \alpha_0 = 1$, $\delta_1 = \delta_0 = 0$, and $\varepsilon = 0$, the distribution is the Laplace distribution with the parameters location $\mu \in \mathbb{R}$, scale $\sigma > 0$ and $k_1 = k_0 = 1$ in equation (12).
6. When $\alpha_1 = \alpha_0 = \alpha > 0, \delta_1 = \delta_0 = \delta > 0$ and $\varepsilon = 0$, the distribution is BEP in [28].
7. When $\alpha_1 = \alpha_0 = 2$ and $\delta_1 = \delta_0 = \delta > 0$, ABEP distribution is used to model bimodality with ε-skew asymmetry in its modes at left and right sides of location $\mu \in \mathbb{R}$, which is a similar manner with [9].
8. When $\delta_1 = \delta_0 = k - 1$, $\alpha_1 = \alpha_0 = 1$, the ABEP distribution becomes ε-skew gamma distribution in [5].
9. When $\alpha_1 = \alpha_0 = 2, \delta_1 = \delta_0 = 0$ and $k_1 = k_0 = 2$, the distribution becomes the ε-skew normal distribution in [33].
10. When $\alpha_1 = \alpha_0 = \alpha > 0$, $\delta_1 = \delta_0 = 0$, $k_1 = k_0 = 1$ and $\varepsilon = 0$, ABEP is a generalized normal or Gaussian (exponential power, abbreviated as EP) distribution in [37].
11. When $\delta_1 = \delta_0 = 0$, $\varepsilon = 0$, $\alpha_1 = \alpha_0 = 2/b$, $b \in (0, 2]$ in [38], $\delta_1 = \delta_0 = 0$, $\alpha_1 = \alpha_0 = \alpha > 0$, $\kappa_1 = 1 - \varepsilon, \kappa_0 = 1 + \varepsilon, \varepsilon \in (-1, 1)$ in [39], $\delta_1 = \delta_0 = 0$, a rescaling via convex combination in [40], $\delta_1 = \delta_0 = 0$, a skewed form via a rescaling in [41,43] and $\delta_1 = \delta_0 = 0$, ε-skew form in [42], the skewed EP and the symmetric EP distributions are equivalent to distributions from Refs. [38–43]. The Refs. [39–41,43] are asymmetric EP distributions based on different sense of skewed form of symmetric EP distribution. The special functions in equations (6) and (9) can be used to get a same kernel of EP with recalculated NC in [38–43].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

12. The ε-skew EP distribution in [44] is a special case of this family for $\delta_0 = \delta_1 = 0$ and $k_1 = k_0 = 2$.
13. The kernel of EP distribution without bimodality in [45,46] is a special case of ABEP when $k_1 = k_0 = k > 0$, $\delta_1 = \delta_0 = 0$ and $\alpha_1 = \alpha_0 = \alpha > 0$.
14. When the variable transformation $z = y^{1/\alpha}$ on function in equation (1) is done,

$$f(z) = \frac{\alpha}{\Gamma(\frac{\delta+1}{\alpha})} z^{\delta} \exp\{-z^{\alpha}\}, \quad z > 0,\ \delta > 0,\ \alpha > 0 \tag{17}$$

is obtained. This is also called as a generalized gamma (GG) distribution. The Pearson type III and V, Erlang, exponential, Weibull, Pareto, Levy, Rayleigh, Nakagami, Frechet, Helmert, Maxwell-Boltzmann and four-parameter exponential gamma as algebraic and exponential functions are members of a function in equation (17) [31,47–49] and references therein.

The first developer of EP is Ref. [45] via solving the differential equation as a different sense from GG in equation (17). The Ref. [46] proposed EP as a generalized error distribution. In ABEP distribution, there are parameters for modelling $x < \mu$ and $x \geq \mu$. Thus, the bimodality can be produced (see also section 2.1.1) and the role of parameters that creates bimodality due to reflection approach in equation (2) of GG function can be observed easily.

3. Maximum Likelihood Estimations for Parameters of ABEP Distribution

Let $x_1, x_2, ..., x_n$ be a random sample of size n from an ABEP distributed population. The unknown parameters $\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0$ and ε will be estimated by ML estimation method [35]. Here, the parameters k_1 and k_0 are nuisance parameters. The log-likelihood $\log(L)$ function is:

$$\begin{aligned}
\log[L(x;\boldsymbol{\theta})] \quad = \quad & n_1[\log(\alpha_1) - \log(2\sigma[k_1(1-\varepsilon)]^{\delta_1+1}) - \log(\Gamma(\frac{\delta_1+1}{\alpha_1}))] \\
& +\delta_1 \sum_{i=1}^{n_1} \log\left(\frac{-(x_i-\mu)}{\sigma}\right) - \sum_{i=1}^{n_1}\left(\frac{-(x_i-\mu)}{\sigma[k_1(1-\varepsilon)]}\right)^{\alpha_1} \\
& +n_0[\log(\alpha_0) - \log(2\sigma[k_0(1+\varepsilon)]^{\delta_0+1}) - \log(\Gamma(\frac{\delta_0+1}{\alpha_0}))] \\
& +\delta_0 \sum_{i=1}^{n_0} \log\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{n_0}\left(\frac{x_i-\mu}{\sigma[k_0(1+\varepsilon)]}\right)^{\alpha_0},
\end{aligned} \tag{18}$$

where n_0 is the number of non-negative observations and n_1 is the number of negative observations. $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}, \hat{\alpha}_1, \hat{\alpha}_0, \hat{\delta}_1, \hat{\delta}_0, \hat{\varepsilon})$ are ML estimators of parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0, \varepsilon)$.

The second derivative test can be used whether or not the $\log(L)$ function in equation (18) has the maximum value, however since PDF has seven parameters $\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0$ and ε, using the Hessian matrix cannot be possible. There can be a solution to overcome this problem if we focus on improving the modelling capacity of PDF having more parameters which help us to increase flexibility of the function and so the efficiency for ML estimators of the parameters μ and σ, especially. A solution in indirect way for this problem is that one can use GOFT statistics, such as KS, CVM and AD to see the distances between expected and empirical cumulative distributions. It is well known that the more small values of the GOFT statistics mean the more fitting performance is accomplished by the function. In the computation process, optimization of nonlinear function in equation (18) is conducted via hybrid genetic algorithm (HGA) in MATLAB 2016a. In HGA, intervals for parameters that will optimize the $\log(L)$ function in equation (18) are used. The intervals for $\mu, \sigma, \alpha_1, \alpha_0, \delta_1, \delta_0$ and ε are $[-5,5]$, $[0,5]$, $[0,10]$, $[0,10]$, $[0,10]$, $[0,10]$ and $(-1,1)$ that is domain of skewness parameter ε. k_1 and k_0 as nuisance parameters are taken to be α_1 and α_0. This form is an appropriate to have same form of normal and Laplace. Let us remind that ABEP is a generalized normal or Laplace distribution. Thus, k_1 and k_0 are nuisance parameters.

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

The Fisher information matrix for parameters μ and σ from ABEP is given by matrix I in the following form:

$$I(\boldsymbol{\theta}) = \begin{bmatrix} \mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu^2}\right] + \mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu^2}\right] & \mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] + \mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] \\ \mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] + \mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] & \mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\sigma^2}\right] + \mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\sigma^2}\right] \end{bmatrix}. \tag{19}$$

The equations (6) and (9) are used to calculate the integrals in matrix I. Due to the analytical expression of PDF in equation (12), undiagonal elements of matrix I are non-zero. Here, shape α_1, α_0, bimodality δ_1, δ_0, skewness ε and nuisance k_1, k_0 parameters make a covariance structure between location μ and scale σ parameters. From this result, covariance structure on ML estimators of other parameters can be seen. Since it is possible to obtain the covariance among ML estimators, Fisher information matrix is obtained only ML estimators of two parameters μ and σ. If there can be a covariance among ML estimators, the inverse of matrix I cannot be obtained except the generalized inverse. Note that getting matrix I for μ and σ from ABEP is tractable for calculation of integration of Fisher information. Using the generalized inverse cannot be preferable due to loss of information in an inverse of a matrix. The loss of information occurs, because the multiplication of inverse of matrix I and I does not give an identity matrix [50]. When $\alpha_1 = \alpha_0 = \alpha$, $\delta_1 = \delta_0 = \delta$, $\varepsilon = 0$ and $k_1 = k_0 = k$, $\mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] + \mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] = 0$, that is, the covariance between ML estimators of μ and σ from ABEP is zero.

$$\mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu^2}\right] = \frac{\delta_1\Gamma(\frac{\delta_1-1}{\alpha_1}) + \alpha_1(\alpha_1-1)\Gamma(1-\frac{1-\delta_1)}{\alpha_1}}{2[\sigma k_1(1-\varepsilon)]^2\Gamma(\frac{\delta_1+1}{\alpha_1})}, \tag{20}$$

$$\mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu^2}\right] = \frac{\delta_0\Gamma(\frac{\delta_0-1}{\alpha_0}) + \alpha_0(\alpha_0-1)\Gamma(1-\frac{1-\delta_0)}{\alpha_0}}{2[\sigma k_0(1+\varepsilon)]^2\Gamma(\frac{\delta_0+1}{\alpha_0})}, \tag{21}$$

$$\mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] = \frac{-\alpha_1^2\Gamma(1+\delta_1/\alpha_1)}{2k_1(1-\varepsilon)\sigma^2\Gamma(\frac{\delta_1+1}{\alpha_1})}, \tag{22}$$

$$\mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\mu\partial\sigma}\right] = \frac{\alpha_0^2\Gamma(1+\delta_0/\alpha_0)}{2k_0(1+\varepsilon)\sigma^2\Gamma(\frac{\delta_0+1}{\alpha_0})}, \tag{23}$$

$$\mathbb{E}_1\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\sigma^2}\right] = \frac{1}{2\sigma^2}\left[-1-\delta_1+\frac{\alpha_1(\alpha_1+1)\Gamma(1+\frac{\delta_1+1}{\alpha_1})}{\Gamma(\frac{\delta_1+1}{\alpha_1})}\right], \tag{24}$$

$$\mathbb{E}_0\left[\frac{\partial^2 \log[f(x;\mu,\sigma)]}{\partial\sigma^2}\right] = \frac{1}{2\sigma^2}\left[-1-\delta_0+\frac{\alpha_0(\alpha_0+1)\Gamma(1+\frac{\delta_0+1}{\alpha_0})}{\Gamma(\frac{\delta_0+1}{\alpha_0})}\right]. \tag{25}$$

Some of regularity conditions [35] are as follows:

1. $\det[I(\mu,\sigma)] < \infty$ and
2. $|\frac{\partial^3}{\partial\boldsymbol{\theta}^3}\log f(x;\boldsymbol{\theta})| \leq M(x)$. Then, $\mathbb{E}[M(X)] < \infty$.

One can verify that the conditions can be satisfied by using Maple or Mathematica. Here, it is possible to get $M(X)$ as X^r in equation (15). Then, the condition 2 is satisfied. The other regularity conditions are already satisfied obviously. Since the ABEP distribution satisfies these two conditions,

$$\sqrt{n}\left(\begin{bmatrix}\hat{\mu}\\ \hat{\sigma}\end{bmatrix} - \begin{bmatrix}\mu\\ \sigma\end{bmatrix}\right) \xrightarrow{D} N(0,[I(\mu,\sigma)]^{-1}), \tag{26}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

that is, $\sqrt{n}\left(\begin{bmatrix}\hat{\mu}\\ \hat{\sigma}\end{bmatrix}-\begin{bmatrix}\mu\\ \sigma\end{bmatrix}\right)$ is asymptotically normal with mean zero vector and covariance matrix $[I(\mu,\sigma)]^{-1}$ and $\hat{\mu}$, $\hat{\sigma}$ are asymptotically efficient and asymptotic normally distributed [35].

4. Real Data Examples

In this section, the modelling capability of ABEP is shown by applying it on two data sets from microarray (http://discover.nci.nih.gov/nature2000/data/selected-data/at-matrix.txt). The analysing of proteins in cancer cell is important. The efficient estimates of location and scale parameters for these proteins are a crucial role in medical care. For this reason, we prefer to focus on these data sets that have the different shapes of peakedness, bimodality and asymmetry.

In the second step, the distributions are considered to model these data sets. In the estimation process, we use the maximum likelihood method together with GOFT statistics, mostly prominent ones that are KS, CVM and AD (robust one) distances to test the fitting capability of distributions [51]. When the estimates of parameters are computed, we can examine via GOFT statistics which of the five PDFs is the best fit on data.

The bimodal extended generalized gamma (BEGG) [29], the Rathie–Swamee (RS) (RS is also known to be a modified version of generalized logistic) [11–13], the exponentiated sinh Cauchy (ESC) [10] and the alpha-skew Laplace (ASL) [24] distributions are used to fit the data and make a comparison between them and ABEP. There are many different distributions which have been proposed, however using explicit expression for CDF should be preferred to fit the data. For this reason, the distributions having explicit expression for their CDFs are used. Thus, GOFTs can be used without including the numerical integration methods having the computational errors.

Modelling data (or Riemann integration in randomly putting the bin of histograms on real line) is an equivalent to an integration. So, the discontinuity at $x = \mu$ is not problem for estimations of parameters. For computation, the HGA is used. HGA also includes the derivative free approach [52] for optimization. Then, the discontinuity point $x = \mu$ is not problem for optimization of $\log(L)$ function in equation (18) according to parameters. At the same time, GOFT statistics are used while performing the computation process.

The Rao-Cramér lower bounds (RCLBs) for ML estimators of parameters are given. The Monte Carlo numerical integration is used to compute the integrals in Fisher information in equation (19) for RS, ESC and ASL distributions.

Since the data generation procedure in Appendix A for ABEP is provided, the performance of fitting can be checked via the counted data at the prescribed ranges of domain as well. However, this procedure is rough when it is compared with GOFTs. It is also beneficial to observe the performance of the random number generation procedure.

The number of replicated sample size n is 100 000. Data generated from ABEP, BEGG and ESC distributions are sorted from small to big values for each sample size n. After sorting, arithmetic mean of 100 000 artificial data is obtained for $n = 118$. After artificial data are generated from their corresponding PDFs, it is also possible to check the fitting performance of these functions via the artificial data (see Tables 3 and 6). Since ABEP, BEGG and ESC are competitive distributions for fitting data and they have a random number generation procedure, they are preferred to check their similarities with real data.

4.1. Example 1: Modelling shape of peakedness, bimodality and asymmetry

The data set labelled as "Homo sapiens Pig7 (PIG7) mRNA, complete cds Chr.16 [381663, (EW), 5′:AA059047, 3′:AA059031]" from microarray is modelled by ABEP, BEGG, RS, ESC and ASL distributions.

 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 23; doi:10.3390/e20010023

Table 1. ML estimates of parameters and GOFT statistics of fitted densities for microarray data.

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\delta}_1$	$\hat{\delta}_0$	$\hat{\varepsilon}$	KS	CVM	AD
ABEP	0.0395	0.1060	1.7322	1.4499	1.2434	0.0505	0.3864	0.0510	0.0662	0.7150
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}_1 = \hat{\alpha}$	$\hat{\alpha}_0 = \hat{\alpha}$	$\hat{\delta}_1$	$\hat{\delta}_0$	$\hat{\varepsilon}$	KS	CVM	AD
BEGG	0.0389	0.0926	1.4880	1.4880	1.0673	0.2657	0.2261	0.0574	0.0850	0.9568
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$	$\hat{b}$	$\hat{p}$			KS	CVM	AD
RS	0.0468	0.2049	1.6278	0.7525	1.1703			0.0865	0.1229	0.8152
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\beta}$				KS	CVM	AD
ESC	0.0226	0.0725	0.4091	1.1730				0.0737	0.1052	0.7086
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$					KS	CVM	AD
ASL	-0.0700	0.1052	-0.5039					0.1318	0.4449	2.3821

Table 2. Asymptotic variances and covariances of ML estimators $\hat{\mu}$ and $\hat{\sigma}$ (10^{-3}).

ABEP		BEGG		RS		ESC		ASL	
$\widehat{Var}(\hat{\mu})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$
$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Var}(\hat{\sigma})$
0.0215	0.0082	0.0073	0.0014	0.6481	0.0375	0.5739	-0.0174	4.365	0.4549
	0.0383		0.0296		0.0615		0.0756		0.0419

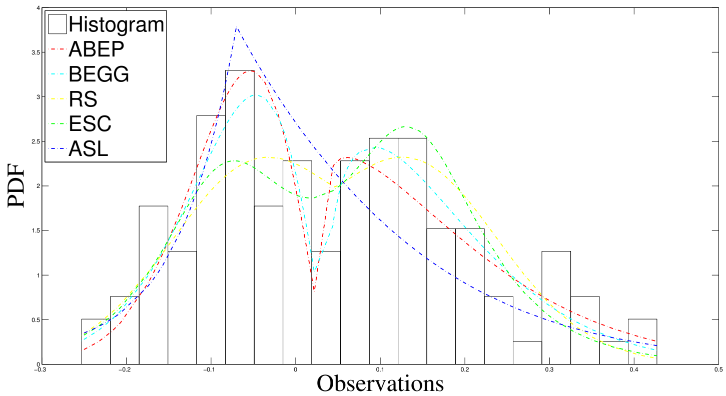


Figure 3. PDF of ABEP, BEGG, RS, ESC and ASL distributions for the estimates of their parameters.

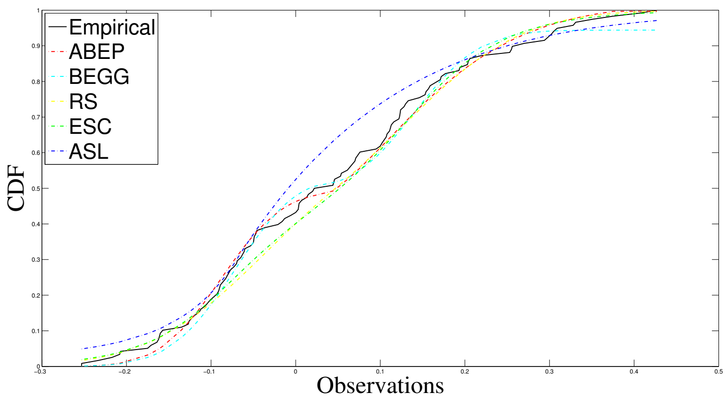


Figure 4. CDF of ABEP, BEGG, RS, ESC and ASL distributions for the estimates of their parameters.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

Table 3. Counted data at ranges [-10, -0.3, -0.1, 0, 0.1, 0.3, 10].

Real data	0	22	28	22	37	9	0
ABEP	0	17	29	22	38	12	0
BEGG	0	18	28	22	42	8	0
ESC	1	21	25	25	41	5	0

4.2. *Example 2: Modelling shape of peakedness, bimodality and asymmetry*

The data set from microarray labelled as "SID 377353, ESTs [5':, 3':AA055048]" is modelled by ABEP, BEGG, RS, ESC and ASL distributions.

Table 4. ML estimates of parameters and GOFT statistics of fitted densities for microarray data.

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\delta}_1$	$\hat{\delta}_0$	$\hat{\varepsilon}$	KS	CVM	AD
ABEP	0.0070	0.0810	2.1174	1.3610	0.4937	0.0031	-0.0380	0.0392	0.0203	0.2773
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}_1 = \hat{\alpha}$	$\hat{\alpha}_0 = \hat{\alpha}$	$\hat{\delta}_1$	$\hat{\delta}_0$	$\hat{\varepsilon}$	KS	CVM	AD
BEGG	-0.0113	0.0516	1.0770	1.0770	1.7593	0.8923	-0.0048	0.0763	0.0936	0.7397
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$	$\hat{b}$	$\hat{p}$			KS	CVM	AD
RS	-0.0201	0.3848	2.7876	3.9241	0.6641			0.0996	0.1083	0.5158
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\beta}$				KS	CVM	AD
ESC	-0.0361	0.0561	0.3143	1.1959				0.0630	0.0396	0.2502
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{a}$					KS	CVM	AD
ASL	0.0340	0.0988	0.2357					0.1099	0.2491	1.5098

Table 5. Asymptotic variances and covariances of ML estimators $\hat{\mu}$ and $\hat{\sigma}$ (10^{-4}).

ABEP		BEGG		RS		ESC		ASL	
$\widehat{Var}(\hat{\mu})$ $\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$ $\widehat{Var}(\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$ $\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$ $\widehat{Var}(\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$ $\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$ $\widehat{Var}(\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$ $\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$ $\widehat{Var}(\hat{\sigma})$	$\widehat{Var}(\hat{\mu})$ $\widehat{Cov}(\hat{\mu},\hat{\sigma})$	$\widehat{Cov}(\hat{\mu},\hat{\sigma})$ $\widehat{Var}(\hat{\sigma})$
1.3731	0.0919 0.2517	0.0602	$3.2295 \cdot 10^{-4}$ 0.0901	0.0317	-0.1177 0.0085	3.1921	1032 3747	344.4	7.592 0.6688

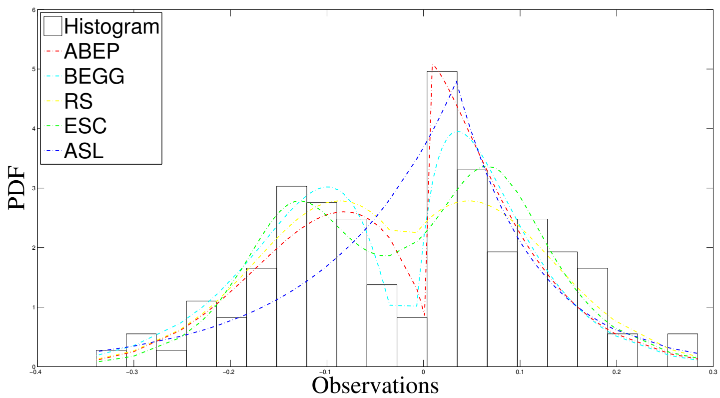


Figure 5. PDF of ABEP, BEGG, RS, ESC and ASL distributions for the estimates of their parameters.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

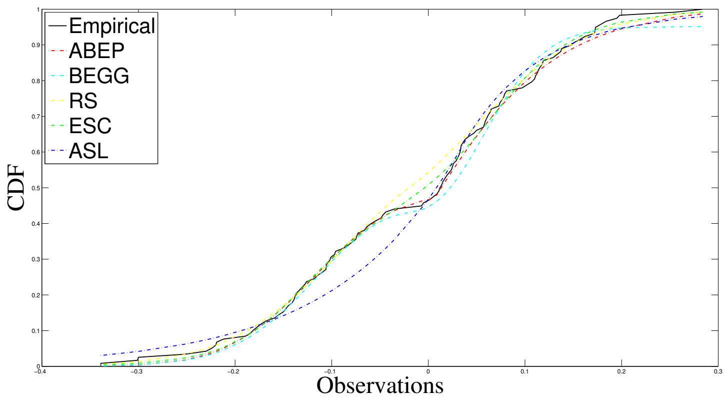


Figure 6. CDF of ABEP, BEGG, RS, ESC and ASL distributions for the estimates of their parameters.

Table 6. Counted data at ranges [-10, -0.4, -0.2, 0, 0.2, 0.4, 10].

Real data	0	9	45	62	2	0	0
ABEP	0	8	46	60	4	0	0
BEGG	0	11	44	59	4	0	0
ESC	0	8	52	54	4	0	0

4.3. Comments on the Results of Examples 1 and 2

For both of two examples, Figures 3 and 5 show that ABEP fits better than the other distributions. Especially, the modalities around location have been modelled as the different modes of heights and the shape of peakedness can be modelled as well. Especially, the right of location is modelled very well by ABEP at example 2. The asymmetry illustrating from example 1 has been modelled. The histograms of data at example 2 do not show an asymmetry and ML estimate of skewness parameter is very near to zero, because as it is seen from Figure 5, the histograms do not have an asymmetry when they are compared with histograms in Figure 3. The unequally distributed histograms around location in Figure 3 can show that there is an asymmetry in data set.

For both of two examples, Tables 1 and 4 represent the ML estimates of parameters of distributions and GOFT statistics of fitted densities. ABEP distribution fits the best data set when we consider on the values of KS and CVM statistics. When we look at the fitting performance for all distributions from Figures 3 and 5, it is seen that ABEP, BEGG and ESC have better fitting performance than RS and ASL. However, when ABEP and ESC are compared, it is observed that two parameters λ and β of ESC are not enough to get the precise fitting on data, because these parameters work together around location. In BEGG, there is only one parameter α to control the fitting shape of function on real line. In ABEP, the role of parameters $\alpha_1, \alpha_0, \delta_1$ and δ_0 around location is constructed definitely. Thus, these parameters affect to get the more precise estimates for parameters μ and σ, which is important if the data are from many phenomena.

It is well known that the probability value (p-value) of a test statistic depends on the fitted density. For this reason, the more efficient density must be preferred before getting the p-value of a test statistic from corresponding density. Then, the potential problem that can occur in future from phenomena can be refrained. The estimates of μ from fitted densities of ABEP, BEGG, RS and ESC can be near to each other, but the estimates of μ of ABEP are more precise one, because ABEP is the best one for fitting on data. At the same way, the estimates of σ of ABEP from both of two examples are the best one.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

The random number generation procedure can be conducted at a more precise way for ABEP, BEGG and ESC distributions, because ABEP and BEGG have an algorithm of random number generation in Appendix A. The inverse of CDF of ESC distribution [10] can be taken to get the random numbers from ESC. The artificial data generated from ABEP distribution also show that the counted artificial data at ranges can be similar with the counted real data at ranges (see Tables 3 and 6). It is noted that the mostly counted data (the numbers 37 and 62 in Example 1 and 2, respectively) at an interval for real data are constructed by the artificial data generated from ABEP distribution for the prescribed ranges at real line. The counted artificial data from ABEP represent the counted real data when they are compared with that from BEGG and ESC. Thus, we can infer that the data generation procedure is also successful after we get the precise estimates of parameters in ABEP via collaboration with GOFT statistics.

GOFT statistics in Tables 1 and 4 show that there can a numerical error in the computation of special function from CDF of ABEP. The AD for ABEP can have a numerical error from the computation of CDF, because CDF of ABEP is a special function. Even if CDF of ABEP depends on special functions that are incomplete gamma functions, the fitting performance of ABEP is the best one due to fact that all possible parameters (shape, bimodality and skewness) that can fit data are added into ABEP.

5. Conclusions and Discussions

A family for bimodal distribution with two parameters fitting the shape of peakedness (α_1 and α_0), two parameters fitting the height of bimodality (δ_1 and δ_0) and a parameter fitting the asymmetry (ε) in data set has been proposed. The unimodal case of this family is obtained when $\delta_1 = \delta_0 = 0$. The skewness parameter in this family is from ε-skew approach that can produce the asymmetry around location. The importance of having these parameters in ABEP for modelling around location separately has been observed when we make a comparison among ABEP, BEGG, ESC and RS distributions that have explicit expression for CDF. As a result, ABEP can model efficiently the shape of peakedness, the bimodality and the asymmetry at the same time, because ABEP has parameters which are responsible to fit the shape of peakedness, the bimodality and the asymmetry in data when it is compared with BEGG, RS, ESC and ASL distributions.

The well known approach which derives PDF without consulting the variable transformation technique is applied for the tractable functions in equations (6)-(11) to propose a new distribution. It is clear that this approach can be applied for other kind of distributions which are on the negative, positive or real line. The disadvantage of this approach is that the analytical expression of a function must be tractable to derive a PDF. The equations (6)-(8) are the power version of gamma, lower and upper incomplete gamma functions. The functions in equations (9)-(11) are transferred to the negative side of real line via using functions in equations (6)-(8). They are new kind of the special functions to calculate the integrals having the kernel of gamma function. One can get distributions via these functions. For example, alpha-skew Laplace [24], alpha–beta skew normal [3], alpha-skew generalized t with variable transformation [53,54], symmetric and asymmetric EP [38–43] distributions with the recalculated NC can also be gotten by these special functions. The special cases, related distributions and flexibility of ABEP are given in relevant section.

The algorithm for generating artificial data from ABEP is provided. Thus, the similarity between artificial and real data sets has been observed as a rough approach and the performance of optimization for the $\log(L)$ function and GOFTs can be supported by this similarity as well. The benefit of GOFTs is depicted when a PDF has more parameters, because the nearness to data, that is, the best performance on optimization for $\log(L)$ function when the competitive PDFs are used, can be checked by GOFT statistics. Thus, if CDF of a PDF exists, using GOFTs as an indirect way to check the potential optimization problem(s) is provided when the second derivative test is a problem for getting the Hessian matrix with respect to parameters of $\log(L)$ function. HGA is also used to overcome the problems that can occur while performing optimization of $\log(L)$ function according to the parameters in ABEP. As a result, performing a cross check between the optimization tool HGA and the GOFT

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at Entropy 2018, 20, 23; doi:10.3390/e20010023

statistics is a beneficial approach to overcome the potential problem(s) from the computation process. Thus, the more precise ML estimates for parameters can be gained. When it is considered on overall results from illustrating of PDF and CDF and also artificial data, the GOFT statistics and these results support each others to show the fitting performance of ABEP.

RCLBs for ML estimators of parameters μ and σ are obtained. The properties of ABEP are provided and so the heavy-tailedness property of ABEP distribution has been examined. The heavy-tailedness of ABEP from Definitions 1 and 2 are guaranteed when $b > a$ in γ function. Definitions 1 and 2 imply that ABEP can be a heavy-tailed distribution together with that comment in there.

The entropy-based parameter estimation for ABEP is on going issue from Refs. [31,32] to study via the proposed special functions in equations (6)-(11). In future, a package in a statistical software R from open access will be prepared for ABEP distribution with the different estimation methods added into this package.

Acknowledgments: I would like to thank sincerely Editor in Chief and Associate Editor and finally referees for their supports and efforts in the process. I would like to thank the partial financial support from Turkish government. I also wish to thank my great parents. The memory of my nice brother...

Author Contributions: The author has finished, read and approved the final form of manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used throughout the text:

PDF	Probability density function
CDF	Cumulative density function
ML	Maximum likelihood
NC	Normalizing constant
GOFT	Goodness of fit test
RCLBs	Rao-Cramér lower bounds
ABEP	Asymmetric bimodal exponential power
BEP	Bimodal exponential power
EP	Exponential power
BEGG	Bimodal extended generalized gamma
RS	Rathie–Swamee
ESC	Exponentiated sinh Cauchy
ASL	Alpha–skew–Laplace
HGA	Hybrid genetic algorithm
KS	Kolmogorov-Smirnov
AD	Anderson-Darling
CVM	Cramér–von Mises
GG	Generalized gamma

Appendix Random Number Generation Procedure from ABEP Distribution

FOR i FROM TO the number of sample size n_1 from α_1,

——— $v_1 = k_1(1-\varepsilon)$,

——— Generate y from Gamma distribution with parameters $\frac{\delta_1+1}{\alpha_1}$ and 1,

——— $x_1 = \mu + \sigma v_1 y^{1/\alpha_1}$.

END FOR

FOR i FROM TO the number of sample size n_0 from α_0,

——— $v_0 = k_0(1+\varepsilon)$,

——— Generate y from Gamma distribution with parameters $\frac{\delta_0+1}{\alpha_0}$ and 1,

——— $x_0 = \mu + \sigma v_0 y^{1/\alpha_0}$.

END FOR

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 December 2017 doi:10.20944/preprints201712.0080.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 23; doi:10.3390/e20010023

Let x be a row vector with $n = n_1 + n_0$ elements of two vectors x_1 with n_1 for negative data and x_0 with n_0 for positive data, that is, $x_n = (x_1, x_0)$.

References

1. Arnold, B.C.; Gómez, H.W.; Salinas, H.S. A doubly skewed normal distribution. Statistics **2015**, 49, 842–858.
2. Bolfarine, H.; Martínez-Flórez, G.; Hugo, Salinas, S. Bimodal symmetric-asymmetric power-normal families Commun. statist.—theory meth. **2013**, dx.doi.org/10.1080/03610926.2013.765475.
3. Shafiei, S.; Doostparast M.; Jamalizadeh, A. The alpha–beta skew normal distribution: properties and applications. Statistics **2016**, 50, 338–349.
4. Sharafi, M.; Sajjadnia, Z.; Behboodian, J. A new generalization of alpha-skew-normal distribution. Commun. statist.—theory meth. **2017**, 46, 6098-6111.
5. Abdulah, E.; Elsalloukh, H. Analyzing skewed data with the epsilon skew gamma distribution. Journal of Statistics Applications & Probability **2013**, 2, 195-202.
6. Abdulah, E.; Elsalloukh, H. Bimodal class based on the inverted symmetrized gamma distribution with Applications. Journal of Statistics Applications & Probability **2014**, 3, 1-7.
7. Ahmed, S.E.; Goria, M.N.; Hussein, A. Gamma mixture: bimodality, inflexions and L-moments. Commun. statist.—theory meth. **2008**, 8, 1147-1161.
8. Dexter, O.C. Some skew-symmetric distributions which include the bimodal ones. Commun. statist.—theory meth. **2015**, 44, 554-563.
9. Arellano-Valle, R.B.; Cortés, M.A.; Gómez, H.W. An extension of the epsilon-skew-normal distribution. Communications in Statistics - Theory and Methods **2010**, 39, 912-922.
10. Cooray, K. Exponentiated sinh Cauchy distribution with applications. Commun. statist.—theory meth. **2013**, 42, 3838-3852.
11. Andrade, B.B.; Rathie, P.N. Fitting asymmetric bimodal data with selected distributions. Journal of Statistical Computation and Simulation **2016**, 86, 3205-3224.
12. Rathie, P.N.; Swamee, P.K. On a new invertible generalized logistic distribution approximation to normal distribution; Technical Research Report No. 07/2006; Department of Statistics, University of Brasilia: Brasilia, Brazil, 2006.
13. Rathie, P.N.; Silva, P.; Olinto, G. Applications of skew models using generalized logistic distribution. Axioms **2016**, 5, 10.
14. Elal-Olivero, D. Alpha-skew-normal distribution. Proyecciones(Antofagasta) **2010**, 29, 224-240.
15. Osvaldo Venegas, O.; Salinas, H.S.; Gallardo, D.I.; Bolfarine, H.; Gómez, H.W. Bimodality based on the generalized skew-normal distribution. Journal of Statistical Computation and Simulation **2017**, 1-26, 10.1080/00949655.2017.1381698.
16. Eugene, N.; Lee, C.; Famoye, F. Beta-normal distribution and its applications. Commun. statist.—theory meth. **2002**, 31, 497-512.
17. Famoye, F.; Lee, C.; Eugene, N. Beta-normal distribution: Bimodality properties and application. Journal of Modern Applied Statistical Methods **2004**, 3, 85-103.
18. Gui, W. A generalization of the slashed distribution via alpha skew normal distribution. Statistical Methods & Applications **2014**, 23, 1-17.
19. Gómez, H.W.; Elal-Olivero, D.; Salinas, H.S.; Bolfarine, H. Bimodal extension based on the skew-normal distribution with application to pollen. Environmetrics **2011**, 22, 50-62.
20. Jamalizadeh, A.; Arabpour, A.R.; Balakrishnan, N.A. Generalized skew two-piece skew-normal distribution. Statistical Papers **2011**, 52, 431–446.
21. Rêgo, L.C.; Cintra, R.J.; Cordeiro, G.M. On some properties of the beta normal distribution. Commun. statist.—theory meth. **2012**, 41, 3722–3738.
22. Rocha, G.H.M.A.; Loschi, R.H.; Arellano-Valle, R.B. Inference in flexible families of distributions with normal kernel. Statistics **2013**, 47, 1184-1206.
23. Sanhueza, A.; Leiva, V.; Balakrishnan, N. The generalized Birnbaum–Saunders distribution and its theory, methodology, and application. Commun. statist.—theory meth. **2008**, 37, 645-670.
24. Shams, H.S.; Alamatsaz, M.H. Alpha-Skew-Laplace distribution. Statistics & Probability Letters **2013**, 83, 774-782.

25. Cordeiro, G.M.; Lemonte, A.J. The beta Laplace distribution. Statistics and Probability Letters **2011**, 81, 973-982.
26. Hassan, M.Y.; El-Bassiouni, M.Y. Bimodal skew-symmetric normal distribution. Commun. statist.—theory meth. **2016**, 45, 1527-1541.
27. Lee, C.; Famoye, F.; Alzaatre A.Y. Methods for generating families of univariate continuous distributions in the recent decades. Comput Stat. **2013**, 5, 219-238.
28. Hassan, Y.M.; Hijazi, R.H. A Bimodal exponential power distribution. Pakistan Journal of Statistics **2010**, 26, 379-396.
29. Çankaya, M.N.; Bulut, Y.M.; Doğru, F.Z.; Arslan, O. A bimodal extension of the generalized gamma distribution. Revista Colombiana de Estadística **2015** 38, 371-384.
30. Arellano-Valle, R.B.; Gomez H.W.; Quintana, F.A. Statistical inference for a general class of asymmetric distributions. Journal of Statistical Planning and Inference **2005**, 128, 427–443.
31. Singh, V. *Entropy-Based Parameter Estimation in Hydrology*. Water Science and Technology Library Vol. 30, Springer Science+Business Media Dordrecht, 1998.
32. Çankaya, M.N.; Korbel, J. On statistical properties of Jizba–Arimitsu hybrid entropy. Physica A **2017**, 475, 1–10.
33. Mudholkar, G.S.; Hutson, A.D. The epsilon-skew-normal distribution for analyzing near-normal data. Journal of Statistical Planning and Inference **2000**, 83, 291-309.
34. Gradshteyn, I.S.; Ryzhik, I.M. *Tables of integrals, sums, series, and products*. Nauka, Moscow, 2014.
35. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*. Wadsworth & Brooks/Cole. Pacific Grove, CA, 589. USA, 1998.
36. Embrechts, P.; Kluppelberg, C.; Mikosch, T. *Modelling Extremal Events for Insurance and Finance*. Springer, Heidelberg, 1997.
37. Nadarajah, S. A generalized normal distribution. Journal of Applied Statistics **2005**, 32, 685-694, DOI: 10.1080/02664760500079464.
38. Choy, S.T.B.; Walker, S.G. The extended exponential power distribution and Bayesian robustness. Statistics & Probability Letters **2003**, 65, 227–232.
39. Delicado, P.; Goria, M.N. A small sample comparison of maximum likelihood, moments and L-moments methods for the asymmetric exponential power distribution. Computational Statistics & Data Analysis **2008**, 52, 1661–1673.
40. Zhu, D.; Victoria Zinde-Walsh, V. Properties and estimation of asymmetric exponential power distribution. Journal of Econometrics **2009**, 148, 86-99.
41. Fernandez, C.; Osiewalski, J.; Steel, M.F.J. Modeling and inference with ν-spherical distributions. J. Amer. Statist. Assoc. **1995** 90, 1331-1340.
42. Theodossiou, P. Skewed generalized error distribution of financial assets and option pricing. SSRN Working Paper, **2000**.
43. Komunjer, I. Asymmetric power distribution: Theory and applications to risk measurement. J. Appl. Econom. **2007**, 22, 891-921.
44. Elsalloukh, H.; Guardiola, J.H.; Young, M. The epsilon-skew exponential power distribution family. Far East Journal of Theoretical Statistics **2005**, 17, 97-107.
45. Subbotin, M.T. On the law of frequency of errors. Matematicheskii Sbornik **1923**, 31, 296–301.
46. Box, G.E.P.; Tiao, G.C. *Bayesian Inference in Statistical Analysis*, Reading, MA: Addison-Wesley, 1973.
47. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions* - Volume 1. Second edition, John Wiley & Sons, New York, 1994.
48. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions* - Volume 2. Second edition, John Wiley & Sons, New York, 1995.
49. Mathai, A.M.; Haubold, H.J. *Special Functions for Applied Scientists*. Springer Science+Business Media, LLC, 2008.
50. Horn, R.A.; Johnson, C.R. *Matrix Analysis*. Second Edition. Cambridge, New York, 2013.
51. Tiku, M.L.; Akkaya, A.D. *Robust Estimation and Hypothesis Testing*. New Age International (P) Limited, Publishers, 2004.
52. Conn, A.R.; Scheinberg, K.; Vicente, L.N. *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics and the Mathematical Programming Society, USA, 2008.

53. Arslan, O.; Genç, A.I. The skew generalized t distribution as the scale mixture of a skew exponential power distribution and its applications in robust estimation. Statistics **2009**, 43, 481-498.
54. Acıtaş, S.; Şenoğlu, B.; Arslan, O. Alpha-skew generalized t distribution. Revista Colombiana de Estadística **2015**, 38, 353-370.