Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/030101

Bimodal Class based on the Inverted Symmetrized Gamma Distribution with Applications

Ebtisam K. Abdulah[1,2,*] *and Hassan Elsalloukh*[3]

[1] Department of Applied Science, University of Arkansas at Little Rock, Arkansas 72204, USA
[2] University of Baghdad, Baghdad, Iraq
[3] Department of Mathematics and Statistics, University of Arkansas at Little Rock, Arkansas 72204, USA

Received: 2 May. 2013, Revised: 15 Jun. 2013, Accepted: 20 Jun. 2013
Published online: 1 Mar. 2014

Abstract: In this work, a new class of asymmetric probability densities, the Epsilon Skew Inverted Gamma (ESIΓ) distribution, which was first introduced by Abdulah [2], is applied to analyzing skewed and bimodality data. Basic properties of this distribution, such as the pdf, cdf, and moments are presented. In addition, computational forms of parameters estimation of MLE and MME are used. Finally, we illustrate the theory of ESIΓ distribution by modeling some real data.

Keywords: Inverted Gamma Distribution, Epsilon Skew Distributions, Maximum Likelihood, Fisher Information Matrix, Bimodal Distributions.

1 Introduction

The last three decades have witnessed a good attention with distributions that can capture skewness and leptokurtic properties for some kind of data exhibiting skewness and/or peakedness. Skewed technique is often utilized in an attempt to requite the problem of long tailed and skewed data sets. The normal curve with such data fails to give an adequate fitting. Therefore, many efforts have been motivated to build robust models, able to fitting the data even if it includes outliers, see with this respect Azzalini [5], Fernández and Steel [11], Mudholkar [14], Arellano [4], Elsalloukh in [10] and [9], Gómez [12], and Ali [3]. The inverse gamma (invΓ) distribution is one of most used distributions for analyzing skewed data and was introduced by Pearson [15] in 1901 and was known as Pearson type five. The invΓ distribution is commonly used in freqentist and Bayesian analysis. It plays an important role as proper prior distribution for the distributions with unknown scale parameter particularly, the normal distribution. Through our review of the literatures, we are needing to parametric models which can control asymmetry, peakedness, and bimodality features. In this research, we extend the reflected inverted Γ distribution, which is symmetric, to the Epsilon Skew Inverted Gamma ESIΓ Abdulah [2] by adding a new skewness parameter. In the process, we make these models more flexible and rich enough in their shapes by adding another shape parameter to the positive and negative orthants of the inverted reflected Γ distribution. The ESIΓ considers four parameters and provides great reliability in modeling skewed, tailed behavior, and bimodality features on the positive and negative real line. These features are eligible and suitable for modeling skewed data which come from different sources. Besides flexibility, we can fit a wide range of cases of this distribution, where invΓ, and inv-χ^2 distributions are two special cases of ESIΓ distribution and obtain reflected Γ Borghi [7] distribution via a transformation case. Typically, the distribution represents reciprocal of a variable distributed as the Epsilon Skew Gamma (ESΓ) distribution Abdulah [1].

The rest of this paper is organized as follows. In Section 2, we develop definitions and some basic properties. In Section 3, the estimation for the model parameters are estimated using the MLE and MME methods. In Section 4, Fisher information matrix is derived and finally, we present the ESIΓ model of the eruption times for the real data set of the Old Faithful Geyser in Section 5.

* Corresponding author e-mail: ekabdulah@ualr.edu

2 Definition and Basic Properties of the ESIΓ Distribution

Borghi [7] defined the pdf of the reflected Γ distribution as

$$f(x;\theta,\beta,k) = \frac{1}{2\Gamma(k)\beta^k}|x-\theta|^{k-1}e^{-\frac{|x-\theta|}{\beta}} \qquad x \in R,$$

where $\theta \in R$, $\beta > 0$, and $k > 0$ are the location, scale, and shape parameters, respectively. This distribution is symmetric about the location parameter θ and has a heavier or lighter tails than the normal distribution depending on the value of the shape parameter k. A formal definition of a pdf of an ESIΓ random variable comes by the following proposition

Proposition 1. *If $Y \sim$ standard reflected $\Gamma(0,1,k)$, then the random variable $X = 1/Y$ is said to have an Epsilon Skew Inverted Gamma distribution, denoted by ESI$\Gamma(\theta,\beta^{-1},k,\varepsilon)$, if it has the pdf*

$$f(x;\gamma) = \frac{\beta^k}{2\Gamma(k)}\begin{cases}\left(\frac{x-\theta}{1-\varepsilon}\right)^{-(k+1)}e^{-\frac{\beta(1-\varepsilon)}{x-\theta}} & if\, x \geq \theta \\ \left(\frac{\theta-x}{1+\varepsilon}\right)^{-(k+1)}e^{-\frac{\beta(1+\varepsilon)}{\theta-x}} & if\, x < \theta\,,\end{cases} \tag{1}$$

where $\gamma = (\theta,\beta,k,\varepsilon)$ and $\theta \in R$, β, $k > 0$, and $|\varepsilon| < 1$ are the location, scale, shape, and skewness parameters, respectively.

Proof. First we start with the pdf of the standard form of the reflected Γ distribution Borghi [7] with a random variable Y

$$f(y;k) = \frac{1}{2\Gamma(k)}|y|^{(k-1)}e^{-|y|}.$$

Let $x = \frac{1}{y}$, then we have $|\frac{d}{dx}f^{-1}(x)| = \frac{1}{x^2}$.
Thus, we have the pdf of a random variable X as

$$f(x;k) = \frac{1}{2\Gamma(k)}|\frac{1}{x}|^{(k-1)}e^{-|\frac{1}{x}|}\frac{1}{x^2}. \tag{2}$$

We can redefine the symmetrization procedure of (2) which leads to the standardized form of the reflected inverted Γ distribution, that is

$$f(x;k) = \frac{1}{2\Gamma(k)}\begin{cases}x^{-(k+1)}e^{-\frac{1}{x}} & \text{if } x \geq 0 \\ (-x)^{-(k+1)}e^{-\frac{1}{-x}} & \text{if } x < 0\,.\end{cases} \tag{3}$$

We now generalize (3) to become the pdf (1).

Note that (3) represents the pdf of the standard form of the ESIΓ distribution. To satisfy the validity of $f(x;\gamma)$ is a pdf, we can proof $f(x;\gamma) \geq 0$ and $\int_{-\infty}^{\infty} f(x;\gamma)\,dx = 1$. The ESI$\Gamma$ distribution can be obtained by standing two inverted gamma distributions back to back and adding a skewness parameter ε for more description of the tail shape. The distribution is sharply peaked near the location parameter, skewed to right when $\varepsilon > 0$, skewed to left when $\varepsilon < 0$, and resembles to the symmetric reflected inverted gamma distribution when ε goes to zero. Figure 1 shows the inverse reflected Γ distribution and Figure 2 shows a variety of the ESIΓ with different values for the skewness parameters ε.

Proposition 2.*If $X \sim$ ESI$\Gamma(\theta,\beta^{-1},k,\varepsilon)$, then the cumulative distribution function, $F(x)$, of X is Abdulah [2]*

$$F(x) = \begin{cases}1-\frac{(1-\varepsilon)}{2\Gamma(k)}\Gamma(k,g(x)) & for\, x \geq \theta \\ \frac{(1+\varepsilon)}{2\Gamma(k)}\Gamma(k,h(x)) & for\, x < \theta\,,\end{cases}$$

Figure 3 shows the cdf of ESIΓ distribution.

3 Central Moments and First Four Moments for the ESIΓ Distribution

In this section, we derive the central moments and first four moments of ESIΓ distribution by using the following proposition.

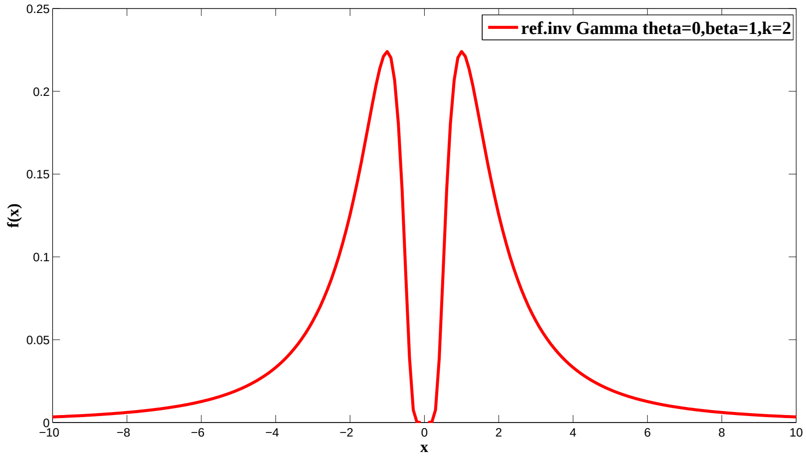

Fig. 1: Inverse Reflected Γ Density Function.

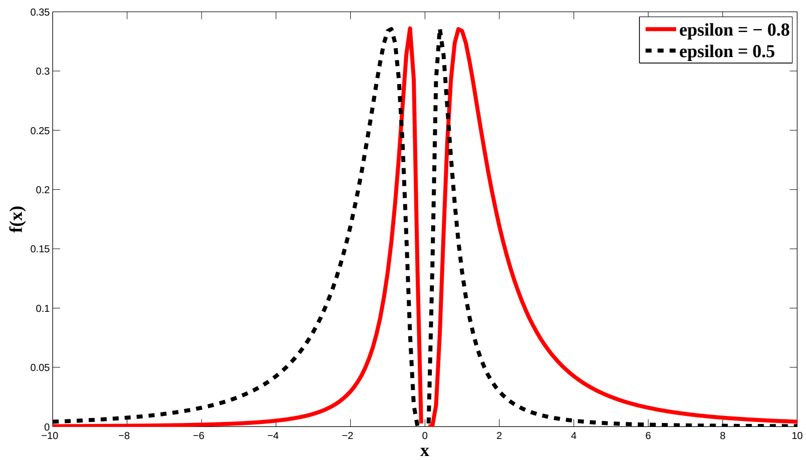

Fig. 2: ESIΓ Density Functions for Different Values of the Skewness Parameter

Proposition 3.*If $X \sim ESI\Gamma(\theta, \beta^{-1}, k, \varepsilon)$, then the central moments, mean, variance, skewness and kurtosis coefficients, and the coefficient of variation are, respectively*

$$E(X-\theta)^n = \frac{\beta^n \Gamma(k-n)}{2\Gamma(k)} \left[(-1)^n (1+\varepsilon)^{n+1} + (1-\varepsilon)^{n+1} \right], \tag{4}$$

$$E(X) = \theta - \frac{2\beta\varepsilon}{(k-1)}, \tag{5}$$

$$Var(X) = \frac{\beta^2}{(k-1)} \left[\frac{1}{(k-2)} + \left(\frac{3}{k-2} - \frac{4}{k-1}\right)\varepsilon^2 \right], \tag{6}$$

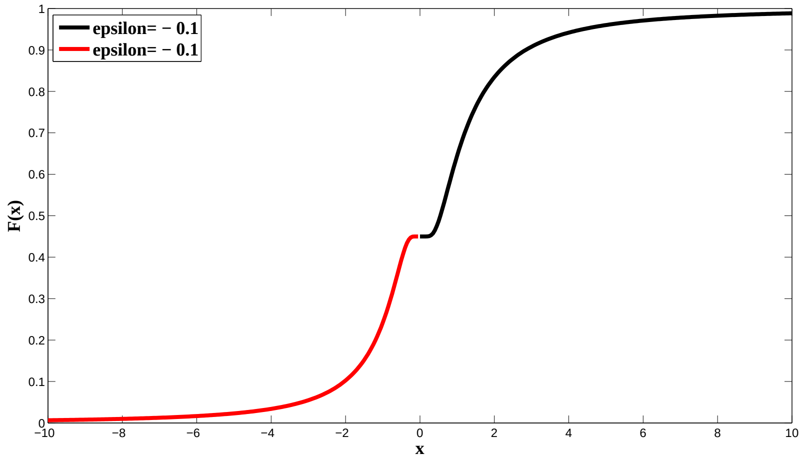

Fig. 3: CDF For ESIΓ Density Functions with ε= -0.1

$$\gamma_1 = \left\{ \frac{1}{(k-1)} \left[\frac{1}{(k-2)} + (\frac{3}{k-2} - \frac{4}{k-1})\varepsilon^2 \right] \right\}^{-3/2} \left[\frac{-2\varepsilon\Gamma(k-1) + (1+3\varepsilon^2)\Gamma(k-2) - 4\Gamma(k-3)\varepsilon(1+\varepsilon^2)}{\Gamma(k)} \right], \quad (7)$$

$$\gamma_2 = \left\{ \frac{1}{(k-1)} \left[\frac{1}{(k-2)} + (\frac{3}{k-2} - \frac{4}{k-1})\varepsilon^2 \right] \right\}^{-2} \left[\frac{-2\varepsilon\Gamma(k-1) + (1+3\varepsilon^2)\Gamma(k-2) - 4\Gamma(k-3)\varepsilon(1+\varepsilon^2) + 2\Gamma(k-4)(1+10\varepsilon^2+5\varepsilon^4)}{\Gamma(k)} \right], \quad (8)$$

$$\gamma_3 = \frac{\left\{ \frac{\beta^2}{(k-1)} \left[\frac{1}{(k-2)} + (\frac{3}{k-2} - \frac{4}{k-1})\varepsilon^2 \right] \right\}^{1/2}}{|\theta - \frac{2\beta\varepsilon}{(k-1)}|}. \quad (9)$$

The proof for all these moments are shown in Abdulah [2].

4 Maximum Likelihood Estimation for the ESIΓ Parameters

In this section, we discuss the estimation of the parameters of the ESIΓ distribution via the maximum likelihood method which provides estimators with asymptotic properties. We assume the location parameter $\theta = 0$, this means we standardize the distribution and treat the other parameters β, k, and ε as unknown.
Let $X \sim \text{ESI}\Gamma(0, \beta^{-1}, k, \varepsilon)$ with a pdf given in (1), then the log likelihood function is Abdulah [2]

$$\log L = nk\log(\beta) - n\log(2) - n\log\Gamma(k) - (k+1)\sum_{i=1}^{n}\log(\frac{x_i^+}{1-\varepsilon}) - \frac{\beta(1-\varepsilon)}{\sum_{i=1}^{n} x_i^+} - (k+1)\sum_{i=1}^{n}\log(\frac{x_i^-}{1+\varepsilon}) - \frac{\beta(1+\varepsilon)}{\sum_{i=1}^{n} x_i^-}, \quad (10)$$

where

$$x_i^+ = \begin{cases} x_i & \text{if } x_i \geq 0 \\ 0 & \text{o/w}, \end{cases}$$

and

$$x_i^- = \begin{cases} -x_i & \text{if } x_i \leq 0 \\ 0 & \text{o/w} . \end{cases}$$

Maximizing (10) leads to the MLE of β and ε, respectively

$$\hat{\beta} = \frac{n\hat{k}\sum_{i=1}^n x_i^+ \sum_{i=1}^n x_i^-}{(1-\hat{\varepsilon})\sum_{i=1}^n x_i^- + (1+\hat{\varepsilon})\sum_{i=1}^n x_i^+},$$

$$\hat{\varepsilon} = \frac{1 \pm \sqrt{1+f^2}}{f},$$

where

$$f = \frac{1}{n(\hat{k}+1)}\left[\frac{\hat{\beta}}{\sum_{i=1}^n x_i^-} - \frac{\hat{\beta}}{\sum_{i=1}^n x_i^+}\right]$$

and the MLE of k is solved numerically Abdulah [2]. The MLEs possess attractive properties, when the sample size increases, such that consistent, efficient and asymptotic normality with variance equal to the inverse of Fisher information matrix as we will derive it in the next section.

5 Method of Moments Estimation (MME)

The ESIΓ distribution consists of four parameters, so for estimating some of its parameters, we find the MME estimates by considering the following case

Using (5) and (6), and from first and second sample moments, we have the MME's of θ Abdulah [2]

$$\tilde{\theta} = \bar{x} + \frac{2\tilde{\beta}\varepsilon}{(k-1)} \tag{11}$$

and the estimation for the scale parameter

$$\begin{aligned} \tilde{\beta}^2 &= \frac{s^2(k-1)^2(k-2)}{(k-1)-\varepsilon^2(k-5)} \\ &= \frac{s(k-1)\sqrt{(k-2)}}{\sqrt{(k-1)-\varepsilon^2(k-5)}} \end{aligned} \tag{12}$$

substituting (12) in (11), we obtain the MME of location parameter as Abdulah [2]

$$\tilde{\theta} = \bar{x} + \frac{2s\varepsilon\sqrt{(k-2)}}{\sqrt{(k-1)-\varepsilon^2(k-5)}},$$

where $\bar{x}$ and s are the sample mean and standard deviation, respectively.

6 Fisher Information Matrix for the ESIΓ Distribution

In this section, we obtain a closed form expression of Fisher information matrix. This matrix equals the negative value of the expectation of second partial derivatives with respect to unknown parameters for the log of likelihood or for pdf

Proposition 4. *Abdulah [2] If $X \sim ESI\Gamma(\theta, \beta^{-1}, k, \varepsilon)$, then the Fisher information matrix for the random variable X is*

$$I = \begin{bmatrix} \frac{k}{\beta^2} & -\frac{1}{\beta} & 0 \\ -\frac{1}{\beta} & \psi'(k) & \frac{\varepsilon}{1-\varepsilon^2} \\ 0 & \frac{\varepsilon}{1-\varepsilon^2} & \frac{(k+1)(1+\varepsilon^2)}{(1-\varepsilon)^2(1+\varepsilon)^2} \end{bmatrix},$$

where $\psi'(k) = \frac{\Gamma''(k)}{\Gamma(k)}$ is the trigamma function.

Proof. We can obtain the elements of Fisher matrix taking straightforward integration as

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial\beta^2}\right] = \frac{k}{\beta^2},$$

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial k^2}\right] = \psi^{'}(k),$$

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial\varepsilon^2}\right] = \frac{(k+1)(1+\varepsilon^2)}{(1-\varepsilon)^2(1+\varepsilon)^2},$$

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial\beta\partial k}\right] = -\frac{1}{\beta},$$

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial\beta\partial\varepsilon}\right] = 0$$

and

$$-E\left[\frac{\partial^2 \log f(x;\gamma)}{\partial k\partial\varepsilon}\right] = \frac{\varepsilon}{(1-\varepsilon^2)}.$$

7 Example

In this section we present a case study of the ESIΓ model for a real data set of the eruption times and duration of the eruption for the Old Faithful Geyser in Yellowstone National Park, Wyoming state, USA. The data are taken from Azzalini [6] in framework of time series analysis and have been used by Härdle [13] in smoothing technique, by Olivero [8], and by Ali [3] in the context of modeling one of the bimodality distributions and skew inv-reflected Pareto distribution, respectively. The data consist of 299 observations of the waiting times (in minutes) for occurring the eruption, where the analysis includes the data that were calculated from August 1st to August 15th in 1985. The ESIΓ distribution interval extends over the real line, therefore we standardize the observations of the data. The estimated parameters, log likelihood scores, AIC, and BIC criteria of the distributions, regular Γ, reflected Γ, ESΓ, ESIΓ, and skew inv-reflected Pareto can be used to select the best fitting model for these data. The results in Table 1 reveals that the ESΓ is a good proper distribution for describing and fitting the waiting time of eruptions for the Old Faithful Geyser data set since it has the lower values of AIC 865.7, BIC 876.8, and highest value of $\log L = -429.8651$. Figure 4 reflects the density plots of the fitted distributions with the histogram of the waiting times of the eruption data. The plot results provide some evidence in favor the $ES\Gamma$ distribution since the histogram of our sample data is closer to ESΓ. This conclusion is also supported by the results in the Table 1.

Table 1: Results of the Parameter MLEs and Corresponding Values of $\log L$, AIC, and BIC for the Five Fitted Distributions for Geyser Data.

Distribution	scale	shape	skewness	$\log L$	*AIC*	*BIC*
regularΓ	2.8672	25.2211		-1217.7587	2439.5	2446.9
reflected Γ	0.7310	1.1318	0	-448.2238	902.4	913.5
ESΓ	0.4123	2.1076	-0.1179	-429.8651	865.7	876.8
ESIΓ	0.1638	0.7390	0.1704	-566.9228	1139.8	1150.9
skewed inv-reflected Pareto	0.2217	2.5878	-0.1973	-509.3197	1024.6	1024.6

8 Discussion

In this paper, we consider ESIΓ family of distributions which includes the invΓ, inverted reflected Γ, inv-χ^2, and ESΓ. It conducts the skewness, peakedness, and bimodality features by its four parameters. It is quite similar to the ESΓ

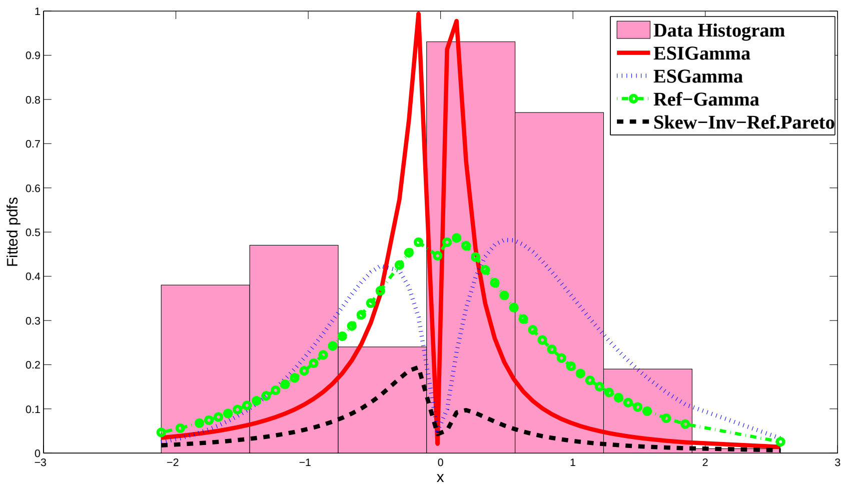

Fig. 4: Fitted Density Functions of the Distributions on the Histogram for Geyser Data.

distribution Abdulah [1] with most of its properties. The main motivation for regarding such model is that in the statistical literatures, there is a lack to fit skewed, peakedness of the top of a distribution curve, and bimodal data sets. The main properties and parameters estimation for this class with two methods are studied. Moreover, Fisher information matrix is derived and a case study has been applied to the Old Faithful Geyser for the models regular Γ, reflected Γ, ESΓ, ESIΓ, and skewed inv-reflected Pareto. We conclude that the ESΓ is better than the alternative distributions for describing the waiting time of eruptions.

References

[1] E. K. Abdulah and H. Elsalloukh, The Epsilon Skew Gamma Distribution, JSM Proceeding, Biometrics Section. San Diago California: American Statistical Association, (2012).

[2] E. K. Abdulah and H. Elsalloukh, The Epsilon Skew Inverted Gamma Distribution, JSM Proceeding, Biometrics Section. San Diago California: American Statistical Association, (2012).

[3] M. M. Ali and J. Woo and S. Nadarajah, Some Skew-Symmetric Inverse Reflected Distributions, Barazilian Journal of Probability and Statistics, **24**, 1-23 (2010).

[4] R. B. Arellano-Valle, H. W. Gómez and F. A. Quintana, Statistical Inference for A general Class of Asymmetric Distributions, Journal of Statistical Planning and Inference, **128**, 427-443 (2005).

[5] A. Azzalini, A Class of Distributions that Includes the Normal Ones, Scandinavian Journal of Statistics, **12**, 171-178 (1985).

[6] A. Azzalini and A. W. Bowman, A Look at Some Data on the Old Faithful Geyser, Journal of the Royal Statistical Society (Applied Statistics), **39**, 357-365 (1990).

[7] O. Borghi, Sobre una Distribution de Frecuenoies, Trabajos de Estadistica, **16**, 171-192 (1965).

[8] D. Elal-Olivero and H. W. Gómez and F. A Quintana, Bayesian Modeling Using a Class of Bimodal Skew-Elliptical Distributions, Journal of Statistical Planning and Inference, **139**, 1484-1492 (2009).

[9] H. Elsalloukh, The Epsilon-Skew Laplace Distribution, The 2008 Proceedings of the American Statistical Association, Biometrics Section, Denever Colorado (2008).

[10] H. Elsalloukh and H. J. Guardiola and M. D. Young, The Epsilon-Skew Exponential Power Distribution Family, Far East Journal of Theoretical Statistics, **16**, 97-112 (2005).

[11] C. Fernández and M. F. J. Steel, On Bayesian Modeling of Fat Tails and Skewness, Journal of American Statistical Association, **93**, 359-371 (1998) .

[12] H. W. Gómez and F.J. Torres and H. Bolfarire, Large Sample Inference for the Epsilon Skew t-Distribution, Communications in Statistics Theory and Methods, **36**, 73-81 (2007).

[13] W. Härdle, Smoothing Techinques, Springer Series in Statistics, New York, (1991).

[14] G. Mudholkar and A. D. Hutson, The Epsilon-Skew-Normal Distribution for Analyzing Near-Normal Data, Journal of Statistical Planning and Inference, **83**, 291-309 (2000).

[15] K. Pearson, Mathematical Contributions to the Theory of Evolution. x. Supplement to A Memoir on Skew Variation, Philosophical Transactions of the Royal Society of London. Containing Papers of Mathematical or Phisical Character, **197**, 443-459 (1901).