Far East Journal of Theoretical Statistics
Volume 28, Number 2, 2009, Pages 201-212
Published Online: October 7, 2009
This paper is available online at http://www.pphmj.com

FURTHER RESULTS ON THE EPSILON-SKEW EXPONENTIAL POWER DISTRIBUTION

HASSAN ELSALLOUKH

Department of Mathematics and Statistics
University of Arkansas
Little Rock, AR 72204-1099, U. S. A.
e-mail: hxelsalloukh@ualr.edu

Abstract

The present paper is a natural continuation of a previous paper of the author [5], which introduced the Epsilon-Skew Exponential Power distribution family (ESEP). In the previous paper, we introduced a new distribution family that we named the Epsilon-Skew Exponential Power distribution (ESEP). We defined basic properties and highlighted special members of the ESEP distribution. We also derived general expressions for the mean, variance, skewness, kurtosis, general moments about zero and about a location parameter, and maximum entropy. In this paper, we derive a stochastic representation, Fisher information matrix, test score for symmetry, and maximum likelihood estimators to an ESEP random variable.

1. Introduction

In [5], we defined the ESEP distribution as

$$f(x) = \frac{\alpha}{2\sqrt{2}\sigma\Gamma(1/\alpha)} \begin{cases} \exp\left(\dfrac{-(x-\theta)^{\alpha}}{2^{\alpha/2}(1-\varepsilon)^{\alpha}\sigma^{\alpha}}\right), & \text{for } x \geq \theta, \\ \exp\left(\dfrac{-(\theta-x)^{\alpha}}{2^{\alpha/2}(1+\varepsilon)^{\alpha}\sigma^{\alpha}}\right), & \text{for } x < \theta, \end{cases} \tag{1}$$

2000 Mathematics Subject Classification: 62E15.

Keywords and phrases: skewness, kurtosis, incomplete gamma function, epsilon-skew Laplace distributions, normal distributions, epsilon-skew normal distributions, entropy.

Received February 25, 2009

where the parameters θ, σ, α and ε are the location, scale, shape and skewness parameters, respectively. The distribution function was defined as

$$F(x \mid \theta, \sigma, \alpha, \varepsilon) = \begin{cases} 1 - \dfrac{(1-\varepsilon)}{2\Gamma(1/\alpha)} \Gamma\left(\dfrac{1}{\alpha}, g(x)\right), & \text{for } x \geq \theta, \\ \dfrac{(1+\varepsilon)}{2\Gamma(1/\alpha)} \Gamma\left(\dfrac{1}{\alpha}, h(x)\right), & \text{for } x < \theta, \end{cases} \tag{2}$$

where $\Gamma\left(\frac{1}{\alpha}, g(x)\right)$ and $\Gamma\left(\frac{1}{\alpha}, h(x)\right)$ are incomplete gamma distributions with $g(x)$ $= \left[\frac{x-\theta}{2^{1/2}(1-\varepsilon)\sigma}\right]^{\alpha}$ and $h(x) = \left[\frac{\theta - x}{2^{1/2}(1+\varepsilon)\sigma}\right]^{\alpha}$.

We also derived in [5], general expressions for the mean, variance, skewness, kurtosis, general moments about zero and about a location parameter, and maximum entropy. Research, based on (1), has been developed, in particular, examples on (1) were applied in [3], special cases of (1) were studied in details in [2], and similar distributions to (1) were developed in [4]. The remainder of the paper is organized as follows. We present a stochastic representation for ESEP random variables in Section 2. In Section 3, we derive Fisher information matrix for the ESEP random variables. In Section 4, we derive a test score for symmetry within the ESEP. We next derive the maximum likelihood estimators to an ESEP random variable in Section 5 and in Section 6 we give a brief discussion of our results.

2. A Stochastic Representation for ESEP Random Variables

In this section, we present a representation of an arbitrary ESEP random variable. We make use of a representation of the symmetric case that is well known [6] and develop a stochastic representation that can be used for computer simulation of ESEP random variates.

Remark 2.1. Using (2), we can show that the cdf of $X \sim \text{ESEP}(0, 1, \alpha, \varepsilon)$ is

$$F(x \mid 0, 1, \alpha, \varepsilon) = \begin{cases} 1 - \dfrac{(1-\varepsilon)}{2\Gamma(1/\alpha)} \Gamma\left(\dfrac{1}{\alpha}, g(x)\right), & \text{for } x \geq 0, \\ \dfrac{(1+\varepsilon)}{2\Gamma(1/\alpha)} \Gamma\left(\dfrac{1}{\alpha}, h(x)\right), & \text{for } x < 0, \end{cases} \tag{3}$$

where $\Gamma\left(\frac{1}{\alpha}, g(x)\right)$ and $\Gamma\left(\frac{1}{\alpha}, h(x)\right)$ are incomplete gamma distributions with

$$g(x) = \left(\frac{x}{\sqrt{2}(1-\varepsilon)}\right)^{\alpha} \tag{4}$$

and

$$h(x) = \left(\frac{-x}{\sqrt{2}(1+\varepsilon)}\right)^{\alpha}. \tag{5}$$

Theorem 2.1. *If* $X \sim \text{ESEP}(0, 1, \alpha, \varepsilon)$, *then* X *admits the stochastic representation* $X \overset{d}{=} IW^{\frac{1}{\alpha}}$, *where W has a gamma distribution with density function*

$$g(x) = \frac{1}{\Gamma(1/\alpha)} x^{\frac{1}{\alpha}-1} \exp(-x), \quad x > 0, \tag{6}$$

and I has a uniform distribution with density function

$$I = \begin{cases} -\sqrt{2}(1+\varepsilon), & \textit{with prob.}\ \dfrac{1+\varepsilon}{2}, \\ \sqrt{2}(1-\varepsilon), & \textit{with prob.}\ \dfrac{1-\varepsilon}{2}. \end{cases} \tag{7}$$

Proof. For $x > 0$, we have

$$G(X) = \Pr\left(IW^{\frac{1}{\alpha}} < x\right) = \Pr\left(I = \sqrt{2}(1-\varepsilon) \text{ and } IW^{\frac{1}{\alpha}} < x\right)$$

$$= \frac{1-\varepsilon}{2} \Pr\left[W < \left(\frac{x}{\sqrt{2}(1-\varepsilon)}\right)^{\alpha}\right]$$

$$= 1 - \frac{1-\varepsilon}{2\Gamma(1/\alpha)} \int_{g(x)}^{\infty} t^{\frac{1}{\alpha}-1} \exp(-t)\, dt, \tag{8}$$

where $g(x)$ is as defined in (4). Similarly, for $x < 0$, we have

$$G(X) = \Pr\left(IW^{\frac{1}{\alpha}} < x\right) = \Pr\left(I = -\sqrt{2}(1+\varepsilon) \text{ and } IW^{\frac{1}{\alpha}} < x\right)$$

$$= \frac{1+\varepsilon}{2\Gamma(1/\alpha)} \int_{h(x)}^{\infty} t^{\frac{1}{\alpha}-1} \exp(-t)\, dt, \tag{9}$$

where $h(x)$ is as defined in (5). We now compare (8) and (9) to (3) in Remark 2.1 and therefore find that $G(X) \equiv F(X)$. We now obtain a stochastic representation for the general case where $X \sim \text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$.

Corollary 2.1. *If* $X \sim \text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$, *then* $X \overset{d}{=} \theta + \sigma I W^{\frac{1}{\alpha}}$, *where W and I are* (6) *and* (7), *respectively*.

We can use Corollary 2.1 to generate random variates from an ESEP distribution. We need to generate only a standard gamma random variate and a binary discrete random variable. The following algorithm is an $\text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$ random-variate generator based on the $\text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$ random variable representation given in Corollary 2.1.

- Generate a gamma $\Gamma\left(\frac{1}{\alpha}, 1\right)$ random variate *W*.
- Generate a standard uniform random variate *U*.
- If $U < \frac{1-\varepsilon}{2}$, set $I \leftarrow \sqrt{2}(1-\varepsilon)$, else set $I \leftarrow -\sqrt{2}(1+\varepsilon)$.
- Set $X \leftarrow \theta + \sigma I W^{\frac{1}{\alpha}}$.
- RETURN *X*.

3. Fisher Information Matrix

Before we consider the problem of testing distributional hypotheses, we first compute the Fisher information matrix $\boldsymbol{I}(\theta, \sigma, \alpha, \varepsilon)$ for an arbitrary ESEP random variable. Under regularity conditions, we have $\boldsymbol{I}(\theta, \sigma, \alpha, \varepsilon) = -E\left[\frac{\partial^2 \log f(x, \boldsymbol{\gamma})}{\partial \gamma_i \partial \gamma_j}\right]$ for $i, j = 1, 2, 3$ and 4, where $X \sim \text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$ with the parameter vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)' = (\theta, \sigma, \alpha, \varepsilon)'$. After lengthy calculations, the Fisher information matrix is

$\boldsymbol{I}(\theta, \sigma, \alpha, \varepsilon)$

$$
= \begin{bmatrix}
\dfrac{\alpha(\alpha-1)\Gamma(\upsilon)}{4\sigma^2(1-\varepsilon^2)\Gamma(1/\alpha)} & 0 & 0 & \dfrac{-\alpha^2}{\sqrt{2}\sigma\Gamma(1/\alpha)(1-\varepsilon^2)} \\
0 & \dfrac{\alpha}{\sigma^2} & \dfrac{-\alpha-(\alpha-1)\psi(1)}{\sigma\Gamma(1/\alpha)} & 0 \\
0 & \dfrac{-\alpha-(\alpha-1)\psi(1)}{\sigma\Gamma(1/\alpha)} & \begin{array}{l} \dfrac{1}{\alpha^2}+\psi'(1/\alpha) \\ +\dfrac{\psi(1)+(\alpha-1)[\psi^2(\upsilon)+\psi'(\upsilon)]}{\alpha\Gamma(1/\alpha)} \end{array} & 0 \\
\dfrac{-\alpha^2}{\sqrt{2}\sigma\Gamma(1/\alpha)(1-\varepsilon^2)} & 0 & 0 & \dfrac{\alpha+1}{1-\varepsilon^2}
\end{bmatrix}, \tag{10}
$$

where $\upsilon = \dfrac{\alpha - 1}{\alpha}$, $\psi(\bullet)$ is the digamma function and $\psi'(\bullet)$ is the trigamma function.

4. Score Test for Detecting Asymmetry within the ESEP

The score or Lagrange multiplier test is a general asymptotic parametric test. Score test statistics are often relatively simple to determine because the statistic requires one to estimate parameters, if necessary, under only the null hypothesis [9].

In the case of multi-dimensional parameters, such as with the case with the ESEP family of distributions, the null hypothesis can be composite because of the presence of nuisance parameters [1]. In this case, the determination of an appropriate score test statistic is more complex. We can develop a score test statistic that is based on the estimation of all parameters, even though we may be interested in testing only a subset of these parameters [7, 10].

In this section, we consider a statistical test for asymmetry within the ESEP distribution family. The score test accounts for nuisance parameters in an explicit fashion. That is, estimators of all nuisance parameters explicitly appear in the score test statistics.

For simplicity, let

$$(x-\delta)^{+} = x-\delta, \quad \text{for } I_{[\delta,\infty)}(x), \tag{11}$$

and

$$(x-\delta)^{-} = \delta - x, \quad \text{for } I_{(-\infty,\delta)}(x), \tag{12}$$

where $\delta \in \mathbb{R}$.

Theorem 4.1. *Let* $X_1, \ldots, X_n$ *be a random sample from an* $ESEP(\theta, \sigma, \alpha, \varepsilon)$ *distribution. Then, under the null hypothesis, we have* $H_0 : \varepsilon = 0,$ *which corresponds to symmetry, versus* $H_A : \varepsilon \neq 0,$ *which corresponds to asymmetry. A score statistic for detecting asymmetry within the* $ESEP(\theta, \sigma, \alpha, \varepsilon)$ *family is*

$$T = \frac{\hat{\alpha}^2 [\Gamma(3/\hat{\alpha})]^{\hat{\alpha}} \left\{ \sum_{i=1}^{n} [(x_i - \hat{\theta})^-]^{\hat{\alpha}} - \sum_{i=1}^{n} [(x_i - \hat{\theta})^+]^{\hat{\alpha}} \right\}^2}{n[\Gamma(1/\hat{\alpha})]^{\hat{\alpha}} 2^{\hat{\alpha}-1} \hat{\sigma}^{\hat{\alpha}} (\hat{\alpha} + 1)}, \tag{13}$$

where $\hat{\theta}$, $\hat{\sigma}$ *and* $\hat{\alpha}$ *are the MLEs of* θ, σ *and* α, *respectively, and* $(x - \hat{\theta})^+$, $(x - \hat{\theta})^-$ *are as defined in* (11) *and* (12), *respectively, with* $\delta = \hat{\theta}$. *Note that* $T \xrightarrow{d} \chi_2^2$ *as* $n \to \infty$, [8].

Proof. Let $\varphi = (\theta, \sigma, \alpha, \varepsilon)'$ in the $\text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$ family. Then we have

$$\frac{\partial L\varphi}{\partial \varepsilon} = \frac{\alpha}{2^{\alpha/2} \sigma^{\alpha}} \left\{ \frac{\sum_{i=1}^{n} [(x_i - \theta)^-]^{\alpha}}{(1 + \varepsilon)^{\alpha+1}} - \frac{\sum_{i=1}^{n} [(x_i - \theta)^+]^{\alpha}}{(1 - \varepsilon)^{\alpha+1}} \right\}$$

and

$$\left.\frac{\partial L\varphi}{\partial \varepsilon}\right|_{\varepsilon=0} = \frac{\hat{\alpha}}{2^{\hat{\alpha}/2} \hat{\sigma}^{\hat{\alpha}}} \left\{ \sum_{i=1}^{n} [(x_i - \hat{\theta})^-]^{\hat{\alpha}} - \sum_{i=1}^{n} [(x_i - \hat{\theta})^+]^{\hat{\alpha}} \right\}.$$

Also from (10), we have $I_n(\varepsilon_0) = -nE\left[\frac{\partial^2 L(\theta, \sigma, \alpha, \varepsilon)}{\partial \varepsilon^2}\right]_{\varepsilon=0} = n(\hat{\alpha} + 1).$ Therefore, a score test statistic for detecting asymmetry within the $\text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$ family is given in (13). Note that, $T \xrightarrow{d} \chi_2^2$ as $n \to \infty$, [8].

5. Maximum Likelihood Estimators for the ESEP Parameters

The likelihood of a set of data is the probability of obtaining that particular set of data given a chosen probability model. The parameter values that maximize the sample likelihood function, provided they exist, are known as the maximum likelihood estimators (MLEs). In many cases, computation of a MLE is usually straightforward when one knows the characteristics of the data generating distribution. However, in the case of multi-parameters, the mathematics needed to determine MLEs is more complex.

5.1. Maximum likelihood estimation of σ

In this section, our main focus is the estimation of σ when $X \sim \text{ESEP}(\theta, \sigma, \alpha, \varepsilon)$. We assume that the mode of the distribution is zero, that is, $\theta = 0$, then the log-likelihood function yields the estimating equation

$$L(\sigma) = -\log \sigma - \frac{x^{+}}{2^{\alpha/2}(1-\varepsilon)^{\alpha}\sigma^{\alpha}} - \frac{x^{-}}{2^{\alpha/2}(1+\varepsilon)^{\alpha}\sigma^{\alpha}},$$

where

$$x^{+} = \sum_{i=1}^{n} (x)^{\alpha}, \quad \text{for } I_{[0,\infty)}(x), \tag{14}$$

and

$$x^{-} = \sum_{i=1}^{n} (-x)^{\alpha}, \quad \text{for } I_{(-\infty,0)}(x). \tag{15}$$

Thus

$$\frac{\partial L(\sigma)}{\partial \sigma} = \frac{-1}{\sigma} + \frac{\alpha x^{+}}{2^{\alpha/2}(1-\varepsilon)^{\alpha}\sigma^{\alpha+1}} + \frac{\alpha x^{-}}{2^{\alpha/2}(1+\varepsilon)^{\alpha}\sigma^{\alpha+1}}, \tag{16}$$

and the MLE is

$$\hat{\sigma}_n = \left[\frac{\alpha x^{+}}{2^{\alpha/2}(1-\varepsilon)^{\alpha}} + \frac{\alpha x^{-}}{2^{\alpha/2}(1+\varepsilon)^{\alpha}}\right]^{1/\alpha}. \tag{17}$$

Note that $\dfrac{\partial^2 L(\hat{\sigma}_n)}{\partial \sigma^2} < 0$ and $L(\sigma)$ increases on $(0, \hat{\sigma}_n)$ and decreases on $(\hat{\sigma}_n, \infty)$ so that $\hat{\sigma}_n$ indeed maximizes (16).

Theorem 5.1.1. *Let* $X_1, X_2, \ldots, X_n$ *be iid* $ESEP(0, \sigma, \alpha, \varepsilon)$ *random variables with unknown dispersion parameter* σ*. Then the MLE of* σ*, given by* (17), *is*:

(1) *Consistent*;

(2) *Asymptotically normal*, *where* $n^{1/2}(\hat{\sigma}_n - \sigma) \xrightarrow{d} N\left(0, \frac{\sigma^2}{\alpha}\right)$;

(3) *Asymptotically efficient*.

Proof. We start with consistency. Let $h(y) \equiv \alpha^{1/\alpha} y^{1/\alpha}$ for $y \geq 0$ and $\alpha > 0$. Note that $h(y)$ is continuous and $\hat{\sigma}_n = h\left(\frac{1}{n}\sum_{i=1}^{n} V_i\right)$, where V_i are iid random variables defined as

$$V_i = \frac{(x_i^+)^\alpha}{2^{\alpha/2}(1-\varepsilon)^\alpha} + \frac{(x_i^-)^\alpha}{2^{\alpha/2}(1+\varepsilon)^\alpha} = \begin{cases} \dfrac{(x)^\alpha}{2^{\alpha/2}(1-\varepsilon)^\alpha}, & \text{for } x \geq 0, \\ \dfrac{(-x)^\alpha}{2^{\alpha/2}(1+\varepsilon)^\alpha}, & \text{for } x < 0, \end{cases}$$

we have

$$E(V_i) = \left(\frac{1}{2^{\alpha/2}(1+\varepsilon)^\alpha}\right)\left(\frac{\alpha}{2\sqrt{2}\sigma\Gamma(1/\alpha)}\right)\int_{-\infty}^{0} (-x)^\alpha \exp\left(\frac{-(-x)^\alpha}{2^{\alpha/2}(1+\varepsilon)^\alpha \sigma^\alpha}\right)dx$$

$$+ \left(\frac{1}{2^{\alpha/2}(1-\varepsilon)^\alpha}\right)\left(\frac{\alpha}{2\sqrt{2}\sigma\Gamma(1/\alpha)}\right)\int_{0}^{\infty} (x)^\alpha \exp\left(\frac{-(x)^\alpha}{2^{\alpha/2}(1-\varepsilon)^\alpha \sigma^\alpha}\right)dx$$

$$= \frac{\sigma^\alpha}{\alpha}.$$

Similarly, we have $E(V_i^2) = \frac{(\alpha+1)\sigma^{2\alpha}}{\alpha^2}$ and $\mathrm{Var}(V_i) = \frac{\sigma^{2\alpha}}{\alpha}$. By the weak law of large numbers, we can see that $\frac{1}{n}\sum_{i=1}^{n} V_i \xrightarrow{p} E(V_i) = \frac{\sigma^\alpha}{\alpha}$ and from the continuity of *h*, we have $\hat{\sigma}_n = h\left(\frac{1}{n}\sum_{i=1}^{n} V_i\right) \xrightarrow{p} h\left(\frac{\sigma^\alpha}{\alpha}\right) = \sigma$ and $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} V_i\right) = \frac{\sigma^{2\alpha}}{n\alpha} \to 0$, as $n \to \infty$. Therefore, $\hat{\sigma}_n$ is a consistent estimator of σ.

Next, by the central limit theorem, we have $n^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n} V_i - \frac{\sigma^\alpha}{\alpha}\right) \xrightarrow{d}$

$N\left(0, \frac{\sigma^{2\alpha}}{\alpha}\right)$, and by the continuity of h, we have

$$n^{1/2}\left(h\left(\frac{1}{n}\sum_{i=1}^{n} V_i\right) - h\left(\frac{\sigma^\alpha}{\alpha}\right)\right) \xrightarrow{d} N\left(0, \left[h'\left(\frac{\sigma^\alpha}{\alpha}\right)\right]^2 \frac{\sigma^{2\alpha}}{\alpha}\right)$$

or

$$n^{1/2}(\hat{\sigma}_n - \sigma) \xrightarrow{d} N\left(0, \frac{\sigma^2}{\alpha}\right).$$

Thus, $\hat{\sigma}_n$ is asymptotically normal.

From (10), we see that $\frac{\sigma^2}{\alpha} = [I(\sigma)]^{-1}$ and thus, $\hat{\sigma}_n$ is asymptotically efficient.

5.2. Maximum likelihood estimation of ε

In this section, our main focus is the estimation of the skew parameter ε. If we again assume the mode of the distribution is zero, that is, $\theta = 0$, then the log-likelihood yields the estimating equation

$$L(\varepsilon) = \frac{-x^+}{2^{\alpha/2}(1-\varepsilon)^\alpha \sigma^\alpha} + \frac{-x^-}{2^{\alpha/2}(1+\varepsilon)^\alpha \sigma^\alpha}, \tag{18}$$

where x^+ and x^- are as defined in (14) and (15), respectively. Then (18) yields the estimating equation

$$\frac{\partial L(\varepsilon)}{\partial \varepsilon} = \frac{\alpha x^-}{2^{\alpha/2}(1+\varepsilon)^{\alpha+1}\sigma^\alpha} - \frac{\alpha x^+}{2^{\alpha/2}(1-\varepsilon)^{\alpha+1}\sigma^\alpha} = 0,$$

and the MLE of ε is

$$\hat{\varepsilon}_n = \frac{(x^-)^{1/(\alpha+1)} - (x^+)^{1/(\alpha+1)}}{(x^-)^{1/(\alpha+1)} + (x^+)^{1/(\alpha+1)}}. \tag{19}$$

Note that $\frac{\partial^2 L(\hat{\varepsilon}_n)}{\partial \varepsilon^2} < 0$, $L(\varepsilon)$ decreases on $[-1, \hat{\varepsilon}_n)$, and increases on $(\hat{\varepsilon}_n, 1]$ so that $\hat{\varepsilon}_n$ indeed maximizes the likelihood function (18).

Theorem 5.2.1. *Let* $X_1, \ldots, X_n$ *be iid ESEP*$(0, \sigma, \alpha, \varepsilon)$ *random variables with an unknown value of* ε. *Then* $\hat{\varepsilon}_n$, *given by* (19), *is*:

(1) *Consistent*;

(2) *Asymptotically normal*, *where* $n^{1/2}(\hat{\varepsilon}_n - \varepsilon) \xrightarrow{d} N\left(0, \dfrac{1-\varepsilon^2}{\alpha+1}\right)$;

(3) *Asymptotically efficient*.

Proof. The proof of this theorem is similar to the proof of Theorem 5.1.1. We begin with consistency. Let $\mathbf{w} = \begin{bmatrix} x^+ \\ x^- \end{bmatrix}$ be the random vector, where x^+ and x^- are as defined in (14) and (15), respectively, and let $G(y_1, y_2) = \dfrac{y_2^{1/(\alpha+1)} - y_1^{1/(\alpha+1)}}{y_2^{1/(\alpha+1)} + y_1^{1/(\alpha+1)}}$.

Note that $\hat{\varepsilon}_n = G(\mathbf{w})$. By the weak law of large numbers, we have $\mathbf{w} \xrightarrow{p} E(\mathbf{w})$, and because of the continuity of G, we have $G(\mathbf{w}) \xrightarrow{p} G[E(\mathbf{w})]$. Now, $G[E(\mathbf{w})] = \varepsilon$ and $\text{Var}(\mathbf{w}) \to 0$ as $n \to \infty$ and, thus, $\hat{\varepsilon}_n \xrightarrow{p} \varepsilon$. Therefore, $\hat{\varepsilon}_n$ is a consistent estimator.

Next, by the central limit theorem, we have $n^{1/2}(\mathbf{w} - E(\mathbf{w})) \xrightarrow{d} N(0, \Sigma_{\mathbf{w}})$, where $E(\mathbf{w})$ and $\Sigma_{\mathbf{w}}$ are the mean and covariance matrices, respectively. From the continuity of G and large-sample theory, we have

$$n^{1/2}(G(\mathbf{w}) - G[E(\mathbf{w})]) \xrightarrow{d} N(0, \mathbf{\Omega}),$$

where $\mathbf{\Omega} = \mathbf{D}\mathbf{\Sigma}_{\mathbf{w}}\mathbf{D}'$ and $\mathbf{D} = \left[\dfrac{\partial G}{\partial y_i} \bigg|_{(y_1, y_2) = E(\mathbf{w})} \right]_{i=1,2}$ is the matrix of partial derivatives of G. We then facilitate the computations of the matrix $\mathbf{D}$ by logarithmic differentiation. The computations of $\mathbf{\Omega}$ are straightforward, thus the asymptotic variance of $\hat{\varepsilon}$ is therefore $\mathbf{\Omega} = \dfrac{1-\varepsilon^2}{\alpha+1}$. We therefore conclude that $\hat{\varepsilon}_n$ is asymptotically normal where

$$n^{1/2}(\hat{\varepsilon}_n - \varepsilon) \xrightarrow{d} N\left(0, \frac{1-\varepsilon^2}{\alpha+1}\right). \tag{20}$$

Note that in (20), the asymptotic variance is $\frac{1-\varepsilon^2}{\alpha+1} = [I(\varepsilon)]^{-1}$, see equation (10). Thus, $\hat{\varepsilon}_n$ is asymptotically efficient.

5.3. Maximum likelihood estimation of α

In this section, we assume that the mode is zero, that is, $\theta = 0$, and estimate α by maximum likelihood. The estimation of α is relatively complicated and requires a numerical estimation. Here we need to maximize the log-likelihood function

$$L(\sigma, \alpha, \varepsilon) = n\left[\log\alpha - \log\sigma - \log\Gamma(1/\alpha) - \frac{x^+}{2^{\alpha/2}(1-\varepsilon)^\alpha\sigma^\alpha} - \frac{x^-}{2^{\alpha/2}(1+\varepsilon)^\alpha\sigma^\alpha}\right], \tag{21}$$

where x^- and x^+ are as defined in (14) and (15), respectively. Substituting (17) and (19) in (21) for σ and ε, respectively, (21) becomes

$$L(\sigma, \alpha, \varepsilon) = \log\left(\frac{\alpha^{(\alpha-1)/\alpha}}{\Gamma(1/\alpha)[(x^-)^{1/(\alpha+1)} + (x^+)^{1/(\alpha+1)}]^{(\alpha+1)/\alpha}}\right) - \frac{1}{\alpha}. \tag{22}$$

We can optimize the function (22) numerically to approximate $\hat{\alpha}$, the maximum likelihood estimator of α. More precisely, to determine $\hat{\alpha}$, we solve

$$\frac{1}{\hat{\alpha}^2}\log\left[\frac{\hat{\alpha}}{(x^-)^{1/(\hat{\alpha}+1)} + (x^+)^{1/(\hat{\alpha}+1)}}\right]$$

$$+ \frac{1}{\hat{\alpha}}\left[\frac{(x^-)^{1/(\hat{\alpha}+1)}\ln(x^-) + (x^+)^{1/(\hat{\alpha}+1)}\ln(x^+)}{(x^-)^{1/(\hat{\alpha}+1)} + (x^+)^{1/(\hat{\alpha}+1)}} + 1\right] - \psi'(1/\hat{\alpha}) = 0.$$

Once we compute the value of α that maximizes the likelihood (22), we can graph (22) as well.

6. Discussion

In the last paper [5], we introduced a new distribution family that includes the normal and Laplace distributions and can be used for analyzing data from skewed or symmetric distributions. We have also derived in [5] expressions for the density function, distribution function, mean, variance, moments and entropy for the new distribution family. In this paper, we derived a stochastic representation, Fisher information matrix, test score for symmetry, and maximum likelihood estimators to an ESEP random variable. Much additional work can and should be done on this new family of random variables.

Acknowledgement

The author wishes to thank Dr. Dean M. Young for his enormous help and useful comments and inspiration.

References

[1] R. D. Cox and V. D. Hinkley, Theoretical Statistics, Chapman and Hall, London, 1974.

[2] H. Elsalloukh, The Epsilon-skew Laplace Distribution, JSM 2008 Proceedings, Biometrics Section, Denver, Colorado: American Statistical Association, 2008.

[3] H. Elsalloukh and J. H. Guardiola, Analyzing near-normal data using a new density function within Azzalini's class of skew distributions, J. Math. Stat. Allied Fields 2(1) (2008), 13 pp.

[4] H. Elsalloukh, J. H. Guardiola, J. W. Seaman and J. D. Stamey, The semi-circular normal distribution, Far East J. Theor. Stat. 20(2) (2006), 207-216.

[5] H. Elsalloukh, J. H. Guardiola and D. M. Young, The epsilon-skew exponential power distribution family, Far East J. Theor. Stat. 17(1) (2005), 97-112.

[6] N. L. Johnson, Systems of frequency curves derived from the first law of Laplace, Trabajos Estadist. 5 (1955), 283-291.

[7] C. Radhakrishna Rao, Two score and 10 years of score tests, Special issue on Rao's score test, J. Statist. Plann. Inference 97(1) (2001), 3-7.

[8] R. J. Serfling, Approximation Theorems of Mathematical Statistics, John Wiley & Sons, Inc., New York, 1980.

[9] S. D. Silvey, The Lagrangian multiplier test, Ann. Math. Statist. 30 (1959), 389-407.

[10] K. G. Smyth, Pearson's goodness of fit statistic as a score test statistic, Sci. Statist. 40 (2003), 115-126.