

## Research Article

# The Location-Scale Mixture Exponential Power Distribution: A Bayesian and Maximum Likelihood Approach

Z. Rahnamaei,<sup>1</sup> N. Nematollahi,<sup>2</sup> and R. Farnoosh<sup>3</sup>

<sup>1</sup> Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>2</sup> Department of Statistics, Allameh Tabataba'i University, Tehran, Iran

<sup>3</sup> School of Mathematics, Iran university of Science and Technology, Narmak, Tehran 16765-163, Iran

Correspondence should be addressed to Z. Rahnamaei, zohreh\_rahnamaei@yahoo.com

Received 12 March 2012; Revised 2 June 2012; Accepted 22 August 2012

Academic Editor: Huijun Gao

Copyright © 2012 Z. Rahnamaei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We introduce an alternative skew-slash distribution by using the scale mixture of the exponential power distribution. We derive the properties of this distribution and estimate its parameter by Maximum Likelihood and Bayesian methods. By a simulation study we compute the mentioned estimators and their mean square errors, and we provide an example on real data to demonstrate the modeling strength of the new distribution.

## 1. Introduction

The Exponential Power (EP) distribution can be considered as a general distribution for random errors. This distribution has the following density function

$$g_X(x) = \frac{1}{2\sigma p^{1/p}\Gamma(1+1/p)} \exp\left(-\frac{|x-\mu|^p}{p\sigma^p}\right), \quad -\infty < x < \infty, \quad (1.1)$$

where  $-\infty < \mu < \infty$ ,  $\sigma > 0$ , and  $p > 1$ . The normal distribution is obtained from this distribution when  $p = 2$ , whereas heavier (lighter) tail distributions are produced when  $p < 2$  ( $p > 2$ ). In particular, we obtain the double exponential distribution for  $p = 1$  and the uniform distribution for  $p \rightarrow \infty$ . This model and its extensions have been studied by [1–7] and others.

Hill and Dixon [8] have given evidence that, in real applications, the distribution of the data is often skew, while virtually all robust methods assume symmetry of the

error distribution. Moreover, the distribution of real data is seldom so heavily tailed as the ones employed in theoretical robustness studies. To handle both skewness and heavy tails simulate, Azzalini [9] proposed the skew exponential power  $\text{SEP}(\mu, \sigma, \lambda, p)$  distribution, which has probability density function pdf

$$f_X(x) = \frac{2}{\sigma} \Phi\left(\text{sign}\left(\frac{x-\mu}{\sigma}\right) \left|\frac{x-\mu}{\sigma}\right|^{p/2} \lambda \left(\frac{2}{p}\right)^{1/2}\right) g\left(\frac{x-\mu}{\sigma}\right), \quad -\infty < x < \infty, \quad (1.2)$$

where  $-\infty < \mu, \lambda < \infty$ ,  $\sigma > 0$ , and  $p > 1$ ,  $\Phi(\cdot)$  is the cumulative distribution function of the standard normal distribution and  $g(\cdot)$  is the density function of EP with  $\mu = 0, \sigma = 1$  in (1.1). The SEP distribution reduced to the EP distribution when  $\lambda = 0$ , to the Skew Normal distribution,  $\text{SN}(\mu, \sigma, \lambda)$  distribution (introduced by [10]) when  $p = 2$ , and to the normal distribution when  $(\lambda, p) = (0, 2)$ .

Another type of the skew-exponential power distribution proposed by Ferreira et al. [11], denoted by  $\text{SEP}(\mu, \sigma^2, \lambda, \nu)$  has pdf given by

$$f_X(x) = 2 \frac{\nu}{\sqrt{2^\nu} \sigma \Gamma(1/2\nu)} e^{-((x-\mu)^2)^\nu / 2\sigma^{2\nu}} \phi\left(\lambda \frac{x-\mu}{\sigma}\right), \quad 0.5 < \nu \leq 1, \quad (1.3)$$

which reduces to the skew-normal distribution when  $\nu = 1$ . They provide an EM type algorithm to estimate the parameters of this distribution. Since, the name of two distributions are the same, we use  $\text{NSEP}(\mu, \sigma^2, \lambda, \nu)$  for the second one.

In this paper, we introduce an EP location-scale mixture distribution. This distribution provides useful asymmetric and heavy-tailed extensions of its symmetric counterparts for robust statistical modeling of data sets involving distributions with heavy tails and skewness.

To this end, in Section 2, the location-scale mixture exponential power (LSMEP) distribution is introduced and some properties are given. A Maximum Likelihood and Bayesian methods are constructed to estimate its parameters. In Section 3, in order to investigate the performance of the proposed methods, we present some simulation studies and a real data application.

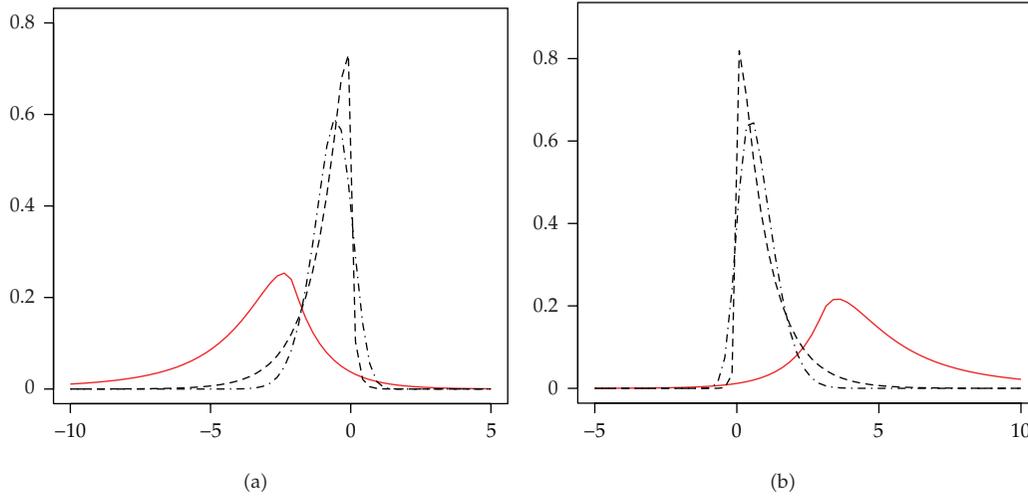
## 2. Location-Scale Mixture Exponential Power Distribution

In this section we introduce the location-scale mixture exponential power distribution and derive some distributional properties. We estimate its parameter by Maximum Likelihood (ML) and Bayesian methods.

*Definition 2.1.* A random variable  $X$  has LSMEP distribution with location parameter  $\mu$ , scale parameter  $\sigma$ , skew parameter  $\beta$ , and shape parameter  $p$ , denoted by  $X \sim \text{LSMEP}(\mu, \sigma, p, \beta, \alpha)$ , if

$$X = \mu + V^{-1}\beta + V^{-1/2}\sigma Y, \quad (2.1)$$

where  $-\infty < \mu, \beta < \infty$ ,  $\sigma > 0$ ,  $p > 1$ ,  $\alpha > 0$ , and  $Y \sim \text{EP}(0, 1, p)$  is independent of  $V \sim \text{beta}(\alpha, 1)$ .



**Figure 1:** (a) Density function of the LSMEP(0,1,1,-2,2) (solid line), SEP(0,1,-2,1) (dashed line), and SN(0,1,-2) (dashed-dotted line). (b) Density function of the LSMEP(0,1,1,3,2) (solid line), SEP(0,1,3,1) (dashed line), and SN(0,1,3) (dashed-dotted line).

Using (1.1) and independence of  $Y$  and  $V$ , the pdf of random variable  $X$  in (2.1) can be easily shown to be

$$f_X(x) = \begin{cases} \frac{\alpha}{2\sigma p^{1/p}\Gamma(1+1/p)} \int_0^1 v^{\alpha-1/2} \exp\left(-\frac{|x-\mu-v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv, & x \neq \mu, \beta \neq 0, \\ \frac{\alpha}{2\sigma p^{1/p}\Gamma(1+1/p)} \int_0^1 v^{\alpha-1/2} \exp\left(-\frac{|\beta|^p}{p\sigma^p v^{p/2}}\right) dv, & x = \mu, \beta \neq 0, \\ \frac{\alpha}{2\sigma p^{1/p}\Gamma(1+1/p)}, & x = \mu, \beta = 0, \end{cases} \quad (2.2)$$

where  $-\infty < x < \infty$ .

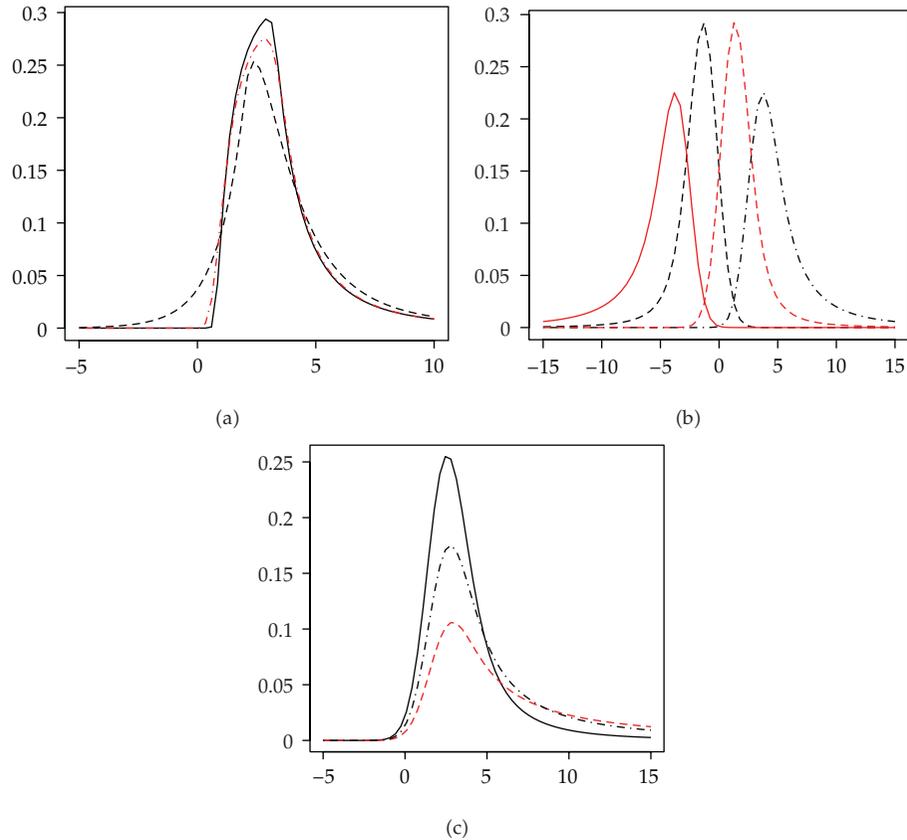
We draw the density curve of LSMEP, SEP, and SN distributions in Figure 1. We see that the LSMEP distribution is more skew and heavier than the other distributions.

We draw the density curve of LSMEP(0,1, $p$ , $\beta$ ,2) for different values of  $p$  and  $\beta$  in Figures 2(a) and 2(b), respectively. We can see that when  $p$  gets larger, the curves becomes more fatter, and when  $\beta$  gets larger, the curve becomes more skew. Also we draw the density curve of LSMEP(0,1,2,2, $\alpha$ ) for  $\alpha = 0.5, 1, 2$  in Figure 2(c). We see that when  $\alpha$  gets larger, the curve becomes more kurtosis.

### 2.1. Properties of the Distribution

Some properties of LSMEP distribution are given in the following theorems.

**Theorem 2.2.** *If  $X \sim \text{LSMEP}(\mu, \sigma, p, \beta, \alpha)$  and  $T = aX + b$ ,  $a, b \in R$ , then  $T \sim \text{LSMEP}(a\mu + b, |a|\sigma, p, a\beta, \alpha)$ .*



**Figure 2:** (a) Density function of the LSMEP(0,1,p,2,2) for  $p = 1$  (dashed line),  $p = 5$  (dashed-dotted line), and  $p = 10$  (solid line). (b) Density function of the LSMEP(0,1,2, $\beta$ ,2) for  $\beta = -3$  (solid line),  $\beta = -1$  (dashed line),  $\beta = 1$  (long-dashed line), and  $\beta = 3$  (dashed-dotted line). (c) Density function of the LSMEP(0,1,2,2, $\alpha$ ) for  $\alpha = 0.5$  (dashed line),  $\alpha = 1$  (dashed-dotted line), and  $\alpha = 2$  (solid line).

**Theorem 2.3.** If  $X \sim \text{LSMEP}(\mu, \sigma, p, \beta, \alpha)$  and  $V \sim \text{beta}(\alpha, 1)$ , then  $X | V = v \sim \text{EP}(\mu + v^{-1}\beta, \sigma/\sqrt{v}, p)$ .

**Theorem 2.4.** If  $X | V = v \sim \text{EP}(v^{-1}\beta, 1/\sqrt{v}, p)$  and  $V \sim \text{beta}(\alpha, 1)$ , then  $X \sim \text{LSMEP}(0, 1, p, \beta, \alpha)$ .

The proof of Theorems 2.2, 2.3, and 2.4, are easily derived from (1.1) and (2.2). From these theorems we can generate LSMEP deviates.

**Theorem 2.5.** The  $n$ th moment of standardized LSMEP distribution (i.e., when  $\mu = 0, \sigma = 1$ ) is

$$\mu_n = E(X^n) = \frac{\beta^n}{\Gamma(1/p)} \sum_{k=0}^{\lfloor n/2 \rfloor} (n2k) \beta^{-2k} \left( 1 + \frac{n-k}{2(\alpha - (n-k))} \right) p^{2k/p} \Gamma\left(\frac{2k+1}{p}\right), \quad \alpha > n-k. \quad (2.3)$$

*Proof.* Let  $Y \sim \text{EP}(0, 1, p)$  and independent of  $V \sim \text{beta}(\alpha, 1)$ . Then  $E(V^{-(k+n)/2}) = 2\alpha / (2\alpha - (k+n))$  and the  $n$ th moment of EP calculated by [6]. So the result is followed by a simple calculation.

From Theorem 2.5 we have

$$E(X) = \frac{\alpha\beta}{\alpha-1}, \quad \alpha > 1, \quad (2.4)$$

$$V(X) = \frac{\alpha}{\alpha-2}\beta^2 + \frac{\alpha}{\alpha-1}p^{2/p} \frac{\Gamma(3/p)}{\Gamma(1/p)} - \frac{\alpha^2}{(\alpha-1)^2}\beta^2, \quad \alpha > 2.$$

Also, the population skewness and kurtosis are easily derived.  $\square$

## 2.2. Maximum Likelihood Estimation

Let  $X_1, \dots, X_n$  be a random sample from  $\text{LSMEP}(\mu, \sigma, p, \beta, \alpha)$  with the observations  $x_1, \dots, x_n$ . We want to find the ML estimates of the parameters of this distribution. From (2.2) the log-likelihood function is given by

$$l(\theta) = c + n \ln \alpha - n \ln \sigma - n \ln \left( p^{1/p} \Gamma \left( 1 + \frac{1}{p} \right) \right) + \sum_{i=1}^n \ln \int_0^1 v^{\alpha-1/2} \exp \left( \frac{-v^{p/2} |x_i - \mu - v^{-1}\beta|^p}{p\sigma^p} \right) dv. \quad (2.5)$$

Suppose  $p$  and  $\beta$  are known. Differentiate (2.5) with respect to  $\mu, \sigma, \alpha$ , and equating the results to zero, we get the following system of equations,

$$\hat{\mu}^{(k+1)} = \frac{\sum_{i=1}^n \left( x_i \omega_0 \left( \hat{s}_{i,0,0}^{(k)} \right) - \beta \omega_1 \left( \hat{s}_{i,0,0}^{(k)} \right) \right)}{\sum_{i=1}^n \omega_0 \left( \hat{s}_{i,0,0}^{(k)} \right)},$$

$$\hat{\sigma}^{2(k+1)} = \frac{\sum_{i=1}^n \left( (x_i - \hat{\mu}^{(k+1)})^2 \omega_0 \left( \hat{s}_{i,1,0}^{(k)} \right) + \beta^2 \omega_2 \left( \hat{s}_{i,1,0}^{(k)} \right) - 2\beta (x_i - \hat{\mu}^{(k+1)}) \omega_1 \left( \hat{s}_{i,1,0}^{(k)} \right) \right)}{n}, \quad (2.6)$$

$$\hat{\alpha}^{(k+1)} = - \frac{n}{\sum_{i=1}^n \omega_3 \left( \hat{s}_{i,1,1}^{(k)} \right)}, \quad k = 0, 1, 2, \dots,$$

where

$$\begin{aligned}\widehat{S}_{i,j,m}^{(k)} &= \frac{|x_i - \widehat{\mu}^{(k+j)} - v^{-1}\beta|}{\widehat{\sigma}^{(k+m)}}, \quad j, m = 0, 1, \\ \omega_l(\widehat{S}_{i,j,m}^{(k)}) &= \frac{\int_0^1 v^{\alpha + ((p-1)/2) - l} (\widehat{S}_{i,j,m}^{(k)})^{p-2} \exp\left(-(\sqrt{v}\widehat{S}_{i,j,m}^{(k)})^p / p\right) dv}{\int_0^1 v^{\alpha-1/2} \exp\left(-(\sqrt{v}\widehat{S}_{i,j,m}^{(k)})^p / p\right) dv}, \quad l = 0, 1, 2, \quad j, m = 0, 1, \quad p > 1, \\ \omega_3(\widehat{S}_{i,j,m}^{(k)}) &= \frac{\int_0^1 v^{\alpha-1/2} \ln(v) \exp\left(-(\sqrt{v}\widehat{S}_{i,j,m}^{(k)})^p / p\right) dv}{\int_0^1 v^{\alpha-1/2} \exp\left(-(\sqrt{v}\widehat{S}_{i,j,m}^{(k)})^p / p\right) dv}, \quad l = 0, 1, 2, \quad j, m = 0, 1, \quad p > 1.\end{aligned}\tag{2.7}$$

By choosing the initial values  $\mu^{(0)}, \sigma^{(0)}, \alpha^{(0)}$ , and iterating (2.6) until convergence, we can find the ML estimates.

### 2.3. Bayesian Method

In this section, we implement the Bayesian methodology using Markov Chain Monte Carlo (MCMC) techniques for estimation of the parameters of the LSMEP distribution. Let  $x = (x_1, x_2, \dots, x_n)$ , then the likelihood function of  $\Theta = (\mu, \sigma, p, \beta, \alpha)$  is given by

$$L(\Theta | x) \propto \left( \frac{\alpha}{\sigma p^{1/p} \Gamma(1 + 1/p)} \right)^n \prod_{i=1}^n \int_0^1 v^{\alpha-1/2} \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv.\tag{2.8}$$

Now, to find the posterior distribution, we need to specify the prior distribution of the unknown parameters  $\Theta$ . By considering a normal prior  $N(\mu_{a_i}, \sigma_{a_i}^2)$   $i = 1, 2$  on both  $\mu$  and  $\beta$  and a truncated normal (on a  $(0, \infty)$ ) prior  $N(\mu_{b_i}, \sigma_{b_i}^2)I_{(0, \infty)}$   $i = 1, 2$  on both  $\sigma$  and  $\alpha$  and a truncated normal (on a  $(1, \infty)$ ) prior  $N(\mu_c, \sigma_c^2)I_{(1, \infty)}$  on  $p$ , and without loss of generality the independence of the parameters, that is,

$$\pi(\Theta) = \pi(\mu)\pi(\sigma)\pi(p)\pi(\beta)\pi(\alpha).\tag{2.9}$$

The posterior distribution of  $\Theta$  given  $x$  can be obtained from (2.8) and (2.9) as follows

$$\pi(\Theta | x) \propto \left( \frac{\alpha}{\sigma p^{1/p} \Gamma(1 + 1/p)} \right)^n \prod_{i=1}^n \left[ \int_0^1 v^{\alpha-1/2} \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv \right] \pi(\Theta).\tag{2.10}$$

Distribution (2.10) does not have a closed form. Hence for doing inference, MCMC algorithm such as the Metropolis-Hasting can be used to generate samples of the posterior distribution of the parameters. We present the following general scheme of sampling.

(1) Set  $t = 0$  and get starting values for the parameters  $\Theta^{(t)} = (\mu^{(t)}, \sigma^{(t)}, p^{(t)}, \beta^{(t)}, \alpha^{(t)})$ .

(2) For  $t = 1, 2, \dots$

(i) generate  $\tilde{\Theta}$  from  $q(\Theta | \Theta^{(t-1)})$ ,

(ii) compute

$$r = \frac{L(\tilde{\Theta} | x)\pi(\tilde{\Theta})}{L(\Theta^{(t-1)} | x)\pi(\Theta^{(t-1)})} \quad (2.11)$$

(we take  $q(\cdot)$  based on the symmetric function.)

(iii) generate  $u$  from  $U[0, 1]$ ,

if  $r < u$  then  $\Theta^{(k)} = \tilde{\Theta}$ , else  $\Theta^{(k)} = \Theta^{(t-1)}$ .

(3) Set  $t = t + 1$  and return 2 until convergence is achieved.

#### 2.4. The Observed Information Matrix

In this section we evaluate the observed information matrix of the LSMEP distribution, which is defined by

$$J_0(\Theta | y) = -\frac{\partial^2 l(\Theta | y)}{\partial \Theta \partial \Theta^T}. \quad (2.12)$$

Under some regularity conditions, the covariance matrix of the maximum likelihood estimates  $\hat{\Theta} = (\hat{\mu}, \hat{\sigma}, \hat{p}, \hat{\beta}, \hat{\alpha})$  can be approximated by the inverse of  $J_0(\Theta | y)$ . The observed information matrix can be obtained as follows

$$J_0(\hat{\Theta} | y) = \sum_{i=1}^n t_i t_i^T, \quad (2.13)$$

where

$$\hat{t}_i = \frac{\partial(\log f(y_i; \theta_j))}{\partial \theta_j}, \quad j = 1, 2, 3, 4, 5, \quad (2.14)$$

see [12, 13].

Now, we consider  $\hat{t}_i$  which is partitioned into components corresponding to all the parameters in  $\Theta$  as

$$\hat{t}_i = (\hat{t}_{i,\mu}, \hat{t}_{i,\sigma}, \hat{t}_{i,p}, \hat{t}_{i,\beta}, \hat{t}_{i,\alpha})^T, \quad (2.15)$$

where

$$\hat{t}_{i,\theta_j} = \frac{\partial \ln f(y_i; \Theta)}{\partial \theta_j}, \quad j = 1, 2, 3, 4, 5. \quad (2.16)$$

We define

$$\begin{aligned}
 I_1 &= \int_0^\infty v^{1/p} \ln v \exp(-v) dv, \\
 I_{2i}(u) &= \int_0^1 v^{\alpha-u/2} \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv, \\
 I_{2i}^v(u) &= \int_0^1 v^{\alpha-u/2} \ln v \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv, \\
 I_{3i}(u_1, u_2) &= \int_0^1 v^{\alpha+(p-u_1)/2} |x_i - \mu - v^{-1}\beta|^{p-u_2} \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv, \\
 I_{3i}^v(u_1, u_2) &= \int_0^1 v^{\alpha+(p-u_1)/2} \ln v |x_i - \mu - v^{-1}\beta|^{p-u_2} \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv, \\
 I_{3i}^{x,v}(u_1, u_2) &= \int_0^1 v^{\alpha+(p-u_1)/2} \ln |x_i - \mu - v^{-1}\beta| |x_i - \mu - v^{-1}\beta|^{p-u_2} \\
 &\quad \times \exp\left(-\frac{|x_i - \mu - v^{-1}\beta|^p v^{p/2}}{p\sigma^p}\right) dv.
 \end{aligned} \tag{2.17}$$

After some algebraic calculation, we obtain

$$\begin{aligned}
 \frac{\partial}{\partial \mu}(f(y_i; \Theta)) &= \frac{\alpha}{2\sigma^{p+1} p^{1/p} \Gamma(1+1/p)} [(x_i - \mu) I_{3i}(1, 2) - \beta I_{3i}(3, 2)], \\
 \frac{\partial}{\partial \sigma}(f(y_i; \Theta)) &= \frac{\alpha}{2\sigma^2 p^{1/p} \Gamma(1+1/p)} \left[ \frac{1}{\sigma^p} I_{3i}(1, 0) - I_{2i}(1) \right], \\
 \frac{\partial}{\partial p}(f(y_i; \Theta)) &= \frac{\alpha}{2\sigma p^{1+1/p} \Gamma(1+1/p)} \left[ \frac{1}{p \Gamma(1+1/p)} I_1 I_{2i}(1) - \frac{(1 - \ln p)}{p} I_{2i}(1) - \frac{1}{\sigma^p} I_{3i}^{x_i}(1, 0) \right. \\
 &\quad \left. - \frac{1}{2\sigma^p} I_{3i}^v(1, 0) + \frac{1}{\sigma^p p} I_{3i}(1, 0) + \frac{\ln \sigma}{\sigma^p} I_{3i}(1, 0) \right], \\
 \frac{\partial}{\partial \beta}(f(y_i; \Theta)) &= \frac{\alpha}{2\sigma^{p+1} p^{1/p} \Gamma(1+1/p)} [(x_i - \mu) I_{3i}(3, 2) - \beta I_{3i}(5, 2)], \\
 \frac{\partial}{\partial \alpha}(f(y_i; \Theta)) &= \frac{1}{2\sigma p^{1/p} \Gamma(1+1/p)} [I_{2i}(1) + \alpha I_{2i}^v(1)].
 \end{aligned} \tag{2.18}$$

In the next section we use the above techniques to estimate the parameters.

## 2.5. Sensitivity Analysis

In this section, we perform sensitivity analysis to detect observations that under small perturbation of the model exert great influence on the maximum likelihood estimators. The best known perturbation schemes are based on case deletion in which the effects are studied of completely removing cases from the analysis by metrics such as the likelihood distance and Cook's distance (see [14]). In this paper, we use the classical measures, namely, Cook distance and the likelihood displacement.

Let  $\hat{\Theta}_{(i)}$  be the ML estimate of  $\Theta$  without the  $i$ th observation in the sample. To assess the influence of the  $i$ th case on the ML estimate  $\hat{\Theta}$ , the basic idea is to compare the difference between  $\hat{\Theta}_{(i)}$  and  $\hat{\Theta}$ . If deletion of a case seriously influences the estimates, more attention should be paid to that case. Hence, if  $\hat{\Theta}_{(i)}$  is far from  $\hat{\Theta}$ , then the  $i$ th case is regarded as an influential observation. A first measure of global influence is defined as the standardized norm of  $\hat{\Theta}_{(i)} - \hat{\Theta}$ , namely, the generalized Cook distance

$$GD_i(\Theta) = \left(\hat{\Theta}_{(i)} - \hat{\Theta}\right)^T [-\check{L}(\Theta)] \left(\hat{\Theta}_{(i)} - \hat{\Theta}\right), \quad (2.19)$$

where  $-\check{L}(\Theta) = \partial^2 L(\Theta) / \partial \Theta \partial \Theta^T$  is the observed information matrix (in  $\Theta = \hat{\Theta}$  point) presented in Section 2.4. Another measure of the difference between  $\hat{\Theta}_{(i)} - \hat{\Theta}$  is the likelihood distance

$$LD_i(\Theta) = 2 \left( L(\hat{\Theta}) - L(\hat{\Theta}_{(i)}) \right). \quad (2.20)$$

In the next section we perform sensitivity analysis to illustrate the usefulness of the proposed methodology.

## 3. Applications

In this section, we present two examples of application of LSMEP distribution. The first one is the two small simulation studies and the other is two statistical analysis of real data sets.

### 3.1. Small Simulation Studies

We perform a small simulation study to investigate the ML estimators that were proposed in Section 2.2. We first generate 100 samples of different sizes from LSMEP distribution for fixed  $p$  and  $\beta$  parameters. We compute the ML estimates of  $\mu$ ,  $\sigma$ , and  $\alpha$  by the iteration method that was illustrated in Section 2.2, which we denote by  $\hat{\mu}_{(i)}$ ,  $\hat{\sigma}_{(i)}$ , and  $\hat{\alpha}_{(i)}$ ,  $i = 1, 2, \dots, 100$ . Then the mean and mean square error (MSE) of these values are reported as the estimates and MSE. For example,  $\hat{\mu} = (1/100) \sum_{i=1}^{100} \hat{\mu}_{(i)}$  and  $MSE(\hat{\mu}) = (1/100) \sum_{i=1}^{100} (\hat{\mu}_{(i)} - \mu)^2$ . The estimates and the MSEs are given in Table 1. This table shows that when the sample size increase, the MSE of estimates  $\hat{\mu}$ ,  $\hat{\sigma}$ , and  $\hat{\alpha}$  convergence to zero. A similar result was happened for the bias of these estimators.

The second small simulation study is the performance of the Bayesian method that was proposed in Section 2.3. We generate 100 samples of sample size 100 from LSMEP(0,1,1.5,1,3), LSMEP(-1,2,3.5,-5,5.5), and LSMEP(2,3,4.5,3,7). Then we compute the Bayes estimates of the

**Table 1:** Estimates of the parameters by choosing 100 samples of sizes  $n = 50, 100, 200,$  and  $500$  from the LSMEP distribution with  $\mu = 1, \sigma = 2, \alpha = 5,$  and  $(\beta, p) = (1.5, 1.5), (-3, 2.5), (5.5, 3.5).$ 

$n$	$\hat{\mu}$	MSE( $\hat{\mu}$ )	$\hat{\sigma}$	MSE( $\hat{\sigma}$ )	$\hat{\alpha}$	MSE( $\hat{\alpha}$ )
$(\beta, p) = (1.5, 1.5)$						
50	0.8889	0.1437	1.9073	0.0796	4.4925	2.0219
100	0.9095	0.0814	1.9181	0.0445	4.4072	1.6825
200	1.0248	0.0259	1.9685	0.0088	4.7590	0.5509
500	1.0422	0.0230	1.9694	0.0065	4.5369	0.5505
$(\beta, p) = (-3, 2.5)$						
50	1.0597	0.1671	1.9881	0.0497	5.6578	3.8411
100	1.0795	0.0496	1.9984	0.0137	4.8799	0.3877
200	1.0377	0.0433	2.0086	0.0192	5.3394	2.1051
500	1.0137	0.0309	2.0009	0.0121	5.1260	0.4680
$(\beta, p) = (5.5, 3.5)$						
50	0.9072	0.0247	1.9727	0.0073	4.8450	1.3445
100	0.9309	0.1365	1.9610	0.0402	4.9806	0.1360
200	0.9520	0.0989	1.9170	0.0624	5.0175	0.8248
500	0.9571	0.0712	1.9982	0.0274	4.8970	0.7012

**Table 2:** Bayesian estimation results for parameters of LSMEP distribution.

	$\mu$	$\sigma$	$p$	$\beta$	$\alpha$
Real parameter	0	1	1.5	1	3
Estimated parameter	0.0662	0.9295	1.7012	0.8850	2.8903
MSE	0.0043	0.0049	0.0405	0.0132	0.0120
Real parameter	-1	2	3.5	-5	5.5
Estimated parameter	-0.8894	1.8922	3.3992	-4.9062	5.4144
MSE	0.0122	0.0117	0.0101	0.0087	0.0073
Real parameter	2	3	4.5	3	7
Estimated parameter	2.0489	2.9102	4.7143	2.9061	6.9099
MSE	0.0024	0.0081	0.0459	0.0088	0.0081

parameters for each sample by MCMC method and derive the final estimates and their MSEs similar to the method that used in ML estimation. The estimated parameters and their MSEs are given in Table 2.

### 3.2. Real Data Application

We use the Australian athletes dataset analyzed in [15–19]. The dataset contains several variables measured on 202 athletes. We consider the variable red cell count (rcc). They note skewness on the left as well as heavy-tail behavior. We fit a LSMEP( $\mu, \sigma, p, \beta, \alpha$ ), an SEP( $\mu, \sigma, \lambda, p$ ) and an NSEP( $\mu, \sigma^2, \lambda, v$ ) distribution to this data set. In the first method we use the optim routine in R software to find the maximum likelihood estimates of the parameters for LSMEP and SEP distributions and we use the EM type algorithm for the estimated parameters for NSEP distribution. In the second method we use the Bayesian estimates of the parameters. The results are shown in Tables 3 and 4. These tables contain the estimates of the parameters of the LSMEP, SEP, and NSEP distributions, besides their corresponding

**Table 3:** Maximum likelihood estimates for fitting LSMEP( $\mu, \sigma, p, \beta, \alpha$ ), SEP( $\mu, \sigma, \lambda, p$ ) distribution, and NSEP( $\mu, \sigma^2, \lambda, v$ ) distribution to the rcc data.

Parameter	MLE of LSMEP	Se	MLE of SEP	Se	MLE of NSEP	Se
$\mu$	4.3800	0.2932	4.7689	0.1442	4.7514	0.3029
$\sigma$	0.5001	0.0424	0.4752	0.0346	0.4580	0.0326
$\lambda$	—	—	-0.1447	0.4268	-0.0898	0.8288
$v$	—	—	—	—	0.9996	0.0511
$p$	2.6900	0.8405	2.2041	0.2545	—	—
$\beta$	0.2811	0.2611	—	—	—	—
$\alpha$	7.4701	0.6649	—	—	—	—
AIC	260.1862		264.0664		264.9651	
EDC	264.3988		267.4365		268.3352	

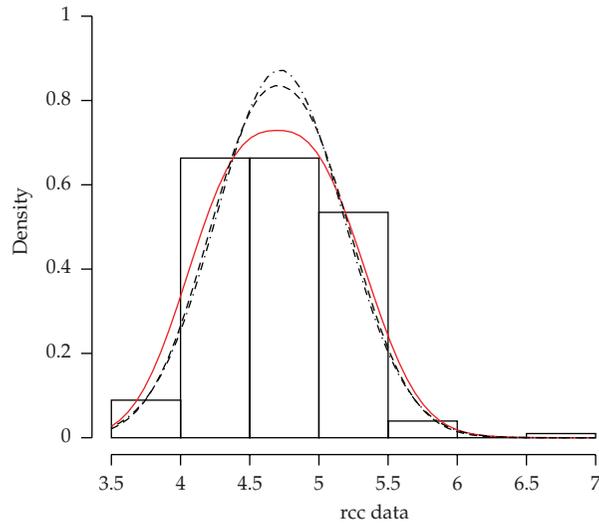
**Table 4:** Bayesian estimates for fitting LSMEP( $\mu, \sigma, p, \beta, \alpha$ ), SEP( $\mu, \sigma, \lambda, p$ ) distribution, and NSEP( $\mu, \sigma^2, \lambda, v$ ) distribution to the rcc data.

Parameter	Bayes estimates of LSMEP	Se	Bayes estimates of SEP	Se	Bayes estimates of NSEP	Se
$\mu$	4.3310	0.2211	4.7767	0.1446	4.7567	0.0814
$\sigma$	0.4594	0.0374	0.4788	0.0361	0.4588	0.0332
$\lambda$	—	—	-0.1224	0.4274	-0.0938	0.2028
$v$	—	—	—	—	0.9106	0.0455
$p$	2.7019	0.6869	2.3339	0.2539	—	—
$\beta$	0.3076	0.1910	—	—	—	—
$\alpha$	7.4884	0.6246	—	—	—	—
AIC	256.1897		264.3211		316.9519	
EDC	260.4024		267.6912		320.3221	

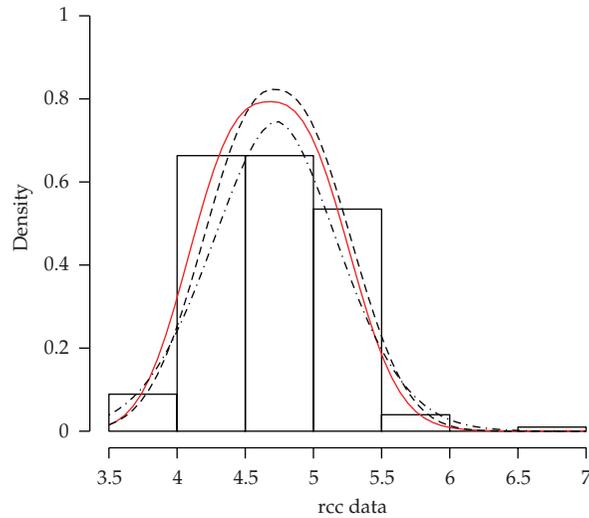
standard errors (Se), computed via the information-based method presented in Section 2.4 for the LSMEP distribution. We used the Hessian matrix in optim routine for computed Se of parameters of SEP distribution and the Se for NSEP distribution computed by [11]. For comparing the models, we also computed the AIC [20] and EDC [21] criteria. From these criteria and Figures 3 and 4, we see that LSMEP distribution has a better fit than the others to this dataset.

### 3.3. Sensitivity Analysis

We use the results obtained through simulated and real data set to illustrate the advantage of the proposed methodology.



**Figure 3:** Histogram of rcc data set with fitted LSMEP( $\mu, \sigma, p, \beta, \alpha$ ) distribution (solid line), SEP( $\mu, \sigma, \lambda, p$ ) distribution (dashed line), and NSEP( $\mu, \sigma^2, \lambda, v$ ) distribution (dashed-dotted line) by method of MLE.

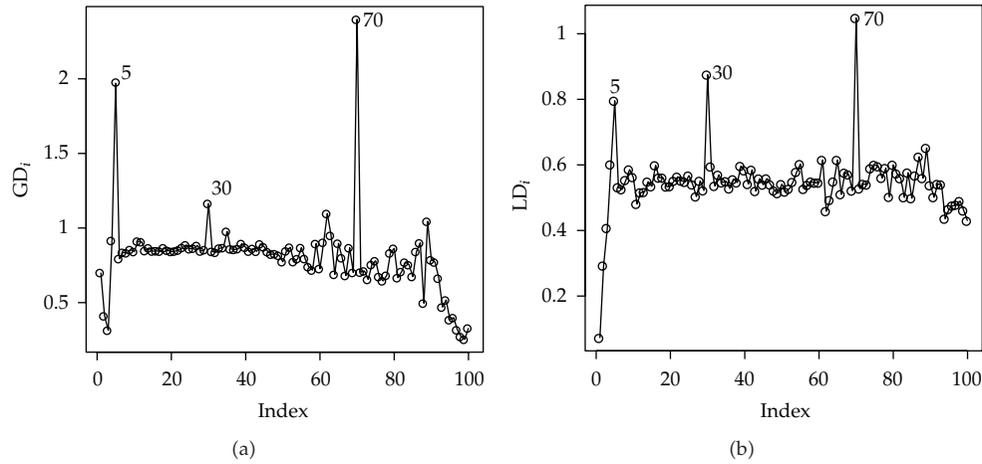


**Figure 4:** Histogram of rcc data set with fitted LSMEP( $\mu, \sigma, p, \beta, \alpha$ ) distribution (solid line), SEP( $\mu, \sigma, \lambda, p$ ) distribution (dashed line), and NSEP( $\mu, \sigma^2, \lambda, v$ ) distribution (dashed-dotted line) by method of Bayesian.

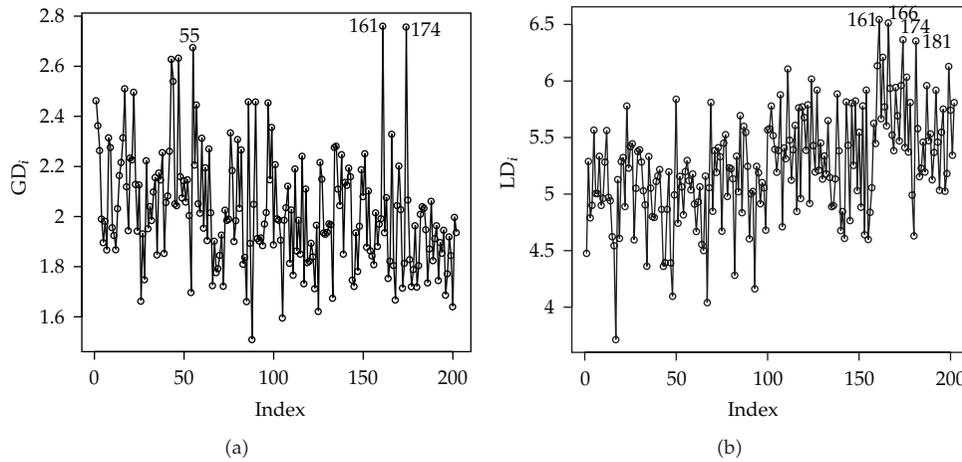
### 3.3.1. Simulated Data

We perform a simulation study to investigate the empirical performance of the proposed methods in Section 2.5. We generate 4 samples of 100 size from LSMEP distribution for fixed size  $p$  and  $\beta$  parameters. Now we consider the following atypical points

$$X_i^* = \begin{cases} X_i + a, & i = 5, 30, 70, \\ \text{otherwise,} & \text{otherwise,} \end{cases} \quad (3.1)$$



**Figure 5:** (a) Index plot of  $GD_i$  for case weights perturbation for simulated data. (b) Index plot of  $LD_i$  for case weights perturbation for simulated data.



**Figure 6:** (a) Index plot of  $GD_i$  for case weights perturbation for rcc data. (b) Index plot of  $LD_i$  for case weights perturbation for rcc data.

where  $a = 6$ . Then we compute  $\hat{\Theta} = (\hat{\mu}, \hat{\sigma}, \hat{\alpha})$  which is the ML estimate of  $\Theta$  with  $X^* = (X_1^*, \dots, X_n^*)$  sample and  $\hat{\Theta}_i$  which is the ML estimate of  $\Theta$  without the  $i$ th observation in  $X^*$ . We compute the generalized Cook distance and likelihood distance that was proposed in Section 2.5. Figure 5 depicts the index plot of  $GD_i$  and  $LD_i$  for case weights perturbation. In Figures 5(a) and 5(b) we see that for all the perturbation schemes considered, the atypical points 5, 30, and 70 were correctly picked up indicating that the methodology works very well when suspicious points are presented in the data set.

### 3.3.2. rcc Dataset

In this section, we use the real data set to find the points which are influential in parameters estimation. Let  $\hat{\Theta}$  be the ML estimate of  $\Theta$  in rcc data and let  $\hat{\Theta}_i$  be the ML estimate of  $\Theta$

without the  $i$ th observation, then we compute the  $GD_i$  and  $LD_i$  as diagnostics for global influence. For the case deletion diagnosis, the measures  $GD_i$  and  $LD_i$  presented in Figures 6(a) and 6(b), respectively, indicate individual 55, 161, 166, 174, and 181 as influential. Note that the individual 161 and 174 are very influential.

## References

- [1] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Mass, USA, 1973.
- [2] R. A. Haro-López and A. F. M. Smith, "On robust Bayesian analysis for location and scale parameters," *Journal of Multivariate Analysis*, vol. 70, no. 1, pp. 30–56, 1999.
- [3] G. Agrò, "Maximum likelihood estimation for the exponential power function parameters," *Communications in Statistics. Simulation and Computation*, vol. 24, no. 2, pp. 523–536, 1995.
- [4] T. J. DiCiccio and A. C. Monti, "Inferential aspects of the skew exponential power distribution," *The Journal of the American Statistical Association*, vol. 99, no. 466, pp. 439–450, 2004.
- [5] H. S. Salinas, R. B. Arellano-Valle, and H. W. Gómez, "The extended skew-exponential power distribution and its derivation," *Communications in Statistics. Theory and Methods*, vol. 36, no. 9–12, pp. 1673–1689, 2007.
- [6] A. İ. Genç, "A generalization of the univariate slash by a scale-mixed exponential power distribution," *Communications in Statistics. Simulation and Computation*, vol. 36, no. 4–6, pp. 937–947, 2007.
- [7] D. Zhu and V. Zinde-Walsh, "Properties and estimation of asymmetric exponential power distribution," *Journal of Econometrics*, vol. 148, no. 1, pp. 86–99, 2009.
- [8] M. A. Hill and W. J. Dixon, "Robustness in real life: a study of clinical laboratory data," *Biometrics*, vol. 38, no. 2, pp. 377–396, 1982.
- [9] A. Azzalini, "Further results on a class of distributions which includes the normal ones," *Statistica*, vol. 46, no. 2, pp. 199–208, 1986.
- [10] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1985.
- [11] C. da Silva Ferreira, H. Bolfarine, and V. H. Lachos, "Skew scale mixtures of normal distributions: properties and estimation," *Statistical Methodology*, vol. 8, no. 2, pp. 154–171, 2011.
- [12] K. E. Basford, D. R. Greenway, G. J. McLachlan, and D. Peel, "Standard errors of fitted component means of normal mixtures," *Computational Statistics*, vol. 12, pp. 1–17, 1997.
- [13] T. I. Lin, J. C. Lee, and S. Y. Yen, "Finite mixture modelling using the skew normal distribution," *Statistica Sinica*, vol. 17, no. 3, pp. 909–927, 2007.
- [14] R. D. Cook, "Detection of influential observation in linear regression," *Technometrics*, vol. 19, no. 1, pp. 15–18, 1977.
- [15] A. Azzalini and A. Dalla Valle, "The multivariate skew-normal distribution," *Biometrika*, vol. 83, no. 4, pp. 715–726, 1996.
- [16] A. Azzalini and A. Capitanio, "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew  $t$ -distribution," *Journal of the Royal Statistical Society B*, vol. 65, no. 2, pp. 367–389, 2003.
- [17] M.G. Genton, *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman and Hall/CRC, Boca Raton, Fla, USA, 2004.
- [18] J. Wang and M. G. Genton, "The multivariate skew-slash distribution," *Journal of Statistical Planning and Inference*, vol. 136, no. 1, pp. 209–220, 2006.
- [19] O. Arslan, "An alternative multivariate skew-slash distribution," *Statistics & Probability Letters*, vol. 78, no. 16, pp. 2756–2761, 2008.
- [20] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [21] Z. D. Bai, P. R. Krishnaiah, and L. C. Zhao, "On rates of convergence of efficient detection criteria in signal processing with white noise," *IEEE Transactions on Information Theory*, vol. 35, no. 2, pp. 380–388, 1989.