JIRSS (2013)
Vol. 12, No. 2, pp 271-291

A Family of Skew-Slash Distributions and Estimation of its Parameters via an EM Algorithm

R. Farnoosh[1], N. Nematollahi[2], Z. Rahnamaei[3], A. Hajrajabi[1]

[1]School of Mathematics, Iran University of Science and Technology, Tehran, Iran.
[2]Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.
[3]Department of Mathematics, Firoozkooh Branch, Islamic Azad University, Firoozkooh, Iran.

Abstract. In this paper, a family of skew-slash distributions is defined and investigated. We define the new family by the scale mixture of a skew-elliptically distributed random variable with the power of a uniform random variable. This family of distributions contains slash-elliptical and skew-slash distributions. We obtain the moments and some distributional properties of the new family of distributions. In the special case of slash skew-t distribution, an EM-type algorithm is presented to estimate the parameters. Some applications are provided for illustrations.

Keywords. Heavy tailed distributions, scale mixture distributions, skew-elliptical distributions, slash distribution.

MSC: 62E15, 62F10.

R. Farnoosh(rfarnoosh@iust.ac.ir), N. Nematollahi(✉)(nematollahi@atu.ac.ir), Z. Rahnamaei(rahnamaei@iaufb.ac.ir), A. Hajrajabi(hajirajabi@iust.ac.ir)
Received: October 2012; Accepted: June 2013

1 Introduction

In analyzing real data sets, there is a general tendency to find more flexible models to presents features of the data. When the data sets have skewness and heavy tails, one should model the data with a skew and/or heavy tailed distribution. In the past three decades, there was a substantial growth in studying and using a number of distributional families for modeling such data sets. See for example Genton (2004) and references cited therein. Some families of distributions which allow for skewness and contain the normal distribution as a proper member have played an important role in these developments. Among them are the skew-normal distribution (Azzalini, 1985, 1986), the multivariate skew-normal distribution (Azzalini and Dalla Valle, 1996), the skew-t distribution (Jones and Faddy, 2003) and the skew-elliptical distribution (Branco and Day, 2001; Sahu et al., 2003).

The class of elliptical distributions, introduced by Kelker (1970) and developed by Fang et al. (1990), includes a vast set of known symmetric distributions, for example, normal, student-t and Pearson type II distributions. Azzalini and Capitanio (1999), Branco and Day (2001) and Sahu et al. (2003) construct skew-elliptical distributions based on elliptical distributions which are skew and contain elliptical distributions as a proper member.

Some authors construct flexible distributions that can be alternatives to the normal distributions for modeling of data sets with heavier tails than normal. First heavy-tailed distributions alternative to the normal, are t and slash distributions. The random variable

$$X = \mu + \sigma \frac{Z}{U^{\frac{1}{q}}} \tag{1.1}$$

is said to have slash-distribution with parameter μ, σ and q (denoted by $X \sim SL(\mu, \sigma, q)$) where $Z \sim N(0,1)$ and $U \sim U(0,1)$ are independent random variables. General properties of this distribution are studied in Rogers and Tukey (1972) and Mosteller and Tukey (1977). For a review of the literature and extensions of slash distributions, see Kafadar (1982, 2001), Wang and Genton (2006), Genc (2007), Gomez et al. (2007), Gomez and Venegas (2008), Arslan (2008, 2009) and Arslan and Genc (2009). By replacing Z in (1.1) with the multivariate normal, multivariate skew-normal and elliptical distributions, Wang and Genton (2006) and Gomez et al. (2007) construct the Multivariate Slash (MSL), the Multivariate Skew-Slash (MSSL) and the Slash-Elliptical (SLEL) distributions, respectively.

In this paper, by replacing Z in (1.1) with skew-elliptical distributions, we define a new family of distributions. These distributions are appropriate for fitting skewed and heavy-tailed data sets . To this end, in Section 2 we state some preliminary definitions. In Section 3, the Slash Skew-Elliptical (SLSEL) distribution is introduced and some properties are given. In Section 4, in the special case of slash skew-t distribution, an EM-type algorithm is constructed to estimate its parameters. A small simulation study and an application to real data are given is Section 5. A conclusion is given in Section 6.

2 Preliminaries

In this section the definitions of elliptical and skew-elliptical distributions' that are needed for subsequent sections' are given.

Definition 2.1. A random variable W has elliptical distribution with location parameter μ and scale parameter σ, denoted by $W \sim EL_1(\mu, \sigma; g)$, if W has probability density function (p.d.f.)

$$f_W(w) = \frac{1}{\sigma} g((\frac{w-\mu}{\sigma})^2), \tag{2.1}$$

for some non-negative function $g(u)$, $u \geq 0$ satisfying $\int_0^\infty u^{-\frac{1}{2}} g(u)du = 1$.

For example, Normal, Cauchy, Student-t, Pearson and Kotz-type distributions belong to this family of distributions. The elliptical distribution was originally defined by Kelker (1970) and a comprehensive review of its properties can be found in Fang et al. (1990).

Definition 2.2. A random variable X has skew-elliptical distribution with location parameter μ, scale parameter σ and skew parameter λ, denoted by $X \sim SEL_1(\mu, \sigma, \lambda; g)$, if X has p.d.f.

$$f_X(x) = \frac{2}{\sigma} f_g(\frac{x-\mu}{\sigma}) F_g(\lambda \frac{(x-\mu)}{\sigma}) \tag{2.2}$$

where $f_g(.)$ and $F_g(.)$ are the p.d.f. and cumulative distribution function (c.d.f.) of $EL_1(0, 1; g)$, respectively.

This form of skew-elliptical distribution was defined by Azzalini and Capitanio (1999), Branco and Dey (2001) and Sahu et al. (2003).

3 Slash Skew-Elliptical Distributions

In this section we define SLSEL distribution. The p.d.f. and some distributional properties are derived.

Definition 3.1. A random variable Y has SLSEL distribution with location parameter μ, scale parameter σ, skew parameter λ and tail parameter $q > 0$, denoted by $Y \sim SLSEL_1(\mu, \sigma, \lambda, q; g)$, if

$$Y = \mu + \sigma \frac{X}{U^{\frac{1}{q}}}, \tag{3.1}$$

where $X \sim SEL_1(0, 1, \lambda; g)$ is independent of $U \sim U(0, 1)$.

Using (2.2) and independence of X and U, the p.d.f. of random variable Y in (3.1) can be easily shown to be

$$f_Y(y) = \begin{cases} \frac{q\sigma^q}{|y-\mu|^{q+1}} \int_0^{(\frac{y-\mu}{\sigma})^2} u^{\frac{q-1}{2}} g(u) F_g(\lambda\sqrt{u}h(y-\mu))du & y \neq \mu \\ \frac{q}{\sigma(q+1)} g(0) & y = \mu, \end{cases} \tag{3.2}$$

where $h(t) = \frac{t}{|t|}$ and g(.) is the density generator function of $EL_1(0, 1; g)$ distribution defined in (2.1). In the special canonical case, i.e., when q=1, (3.2) reduces to

$$f_Y(y) = \begin{cases} \frac{\sigma}{(y-\mu)^2} \int_0^{(\frac{y-\mu}{\sigma})^2} g(u) F_g(\lambda\sqrt{u}h(y-\mu))du & y \neq \mu \\ \frac{1}{2\sigma} g(0) & y = \mu. \end{cases} \tag{3.3}$$

When $q \to \infty$ we obtain $SEL_1(\mu, \sigma, \lambda; g)$ distribution. When $\lambda = 0$, (3.2) reduce to SLEL density defined by Gomez et al. (2007). If g(.) is the generator of normal distribution, i.e.,

$$g(t) = (2\pi)^{-\frac{1}{2}} e^{-\frac{t}{2}}, \tag{3.4}$$

then (3.2) reduces to the density of univariate skew-slash distribution defined by Wang and Genton (2006) and we denoted (3.1) by $Y \sim SLSN_1(\mu, \sigma, \lambda; q)$ (X has skew-normal distribution, $X \sim SN(0, 1, \lambda)$). When g(t) is given by (3.4) and $\lambda = 0$, (3.2) reduces to the density of general slash distribution defined in (1.1).

3.1 Some Properties of SLSEL Distributions

Let $Y \sim SLSEL_1(0, 1, \lambda, q; g)$ and $T \sim SLEL_1(0, 1, \lambda; g)$ (SLEL distribution defined by Gomez et al., 2007). Then it is easy to see that the

p.d.f.s of $W = |Y|$ and $V = Y^2$ are given by

$$\begin{aligned} f_W(w;q) &= \frac{q}{w^{q+1}} \int_0^{w^2} u^{\frac{q-1}{2}} g(u)du, \\ f_V(v;q) &= \frac{q}{2v^{\frac{q+2}{2}}} \int_0^{v} u^{\frac{q-1}{2}} g(u)du, \end{aligned} \tag{3.5}$$

which are the p.d.f.s of $|T|$ and T^2 derived by Gomez et al. (2007), respectively. Therefore $Y^2 = T^2$, and this property is similar to the relation between normal and skew-normal distributions.

Theorem 3.1. *If $Y \sim SLSEL_1(\mu, \sigma, \lambda, q; g)$ and $T = aY + b$, $a, b \in R$, then $T \sim SLSEL_1(a\mu + b, |a|\sigma, q, \lambda h(a); g)$.*

Theorem 3.2. *If $Y|U = u \sim SEL_1(0, u^{-\frac{1}{q}}, \lambda; g)$ and $U \sim U(0,1)$, then $Y \sim SLSEL_1(0, 1, \lambda, q; g)$.*

The proof of Theorems 3.1 and 3.2 are easily derived from (2.2) and (3.2).

Remark 3.1. Theorems 3.1 and 3.2 show that the SLSEL distribution is invariant under linear transformations and can be represented as a particular scale-mixture of SEL and $U(0,1)$ distributions. The results are also useful for generating $SLSEL_1(\mu, \sigma, \lambda, q; g)$ deviates.

Theorem 3.3. *If $Z \sim SLSEL_1(0, 1, \lambda, q; g)$ and $Y \sim SLSEL_1(\mu, \sigma, \lambda, q; g)$ then*

$$\mu_k = E(Z^k) = \frac{q}{q-k} a_k, \qquad k = 0, 1, 2, ... \qquad (q > k),$$

where $a_k = 2\int_R x^k g(x^2) F_g(\lambda x)dx$ is the r-th moment of $SEL_1(0, 1, \lambda; g)$-distribution given in (2.2), and

$$\mu'_k = E(Y^k) = \sum_{l=0}^{k} \left(\begin{array}{c} k \\ l \end{array} \right) \sigma^l \mu^{k-l} \mu_l, \qquad k = 0, 1, 2, ... \qquad (q > k).$$

Proof. Let $X \sim SEL_1(0, 1, \lambda; g)$ and independent of $U \sim U(0,1)$. Then $E(U^{-\frac{k}{q}}) = \frac{q}{q-k}$, $Z = \frac{X}{U^{\frac{1}{q}}}$ and $Y = \mu + \sigma Z$. So the results are followed by a simple calculation.

As a special case, let $Y \sim SLSN_1(\mu, \sigma, \lambda, q)$, which has the generator function g(.) given by (3.4). Then a_k is the k-th moment of skew-normal distribution and for $k = 1, 2$ is given by (Azzalini, 1985) $a_1 = \sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$, $a_2 = 1$. So, from Theorem 3.3 we have

$$\begin{aligned}
\mu_1' &= E(Y) = \mu + \frac{q}{q-1}\sqrt{\frac{2}{\pi}}\frac{\sigma\lambda}{\sqrt{1+\lambda^2}}, && q > 1,\\
\mu_2' &= E(Y^2) = \mu^2 + 2\mu\frac{q}{q-1}\sqrt{\frac{2}{\pi}}\frac{\sigma\lambda}{\sqrt{1+\lambda^2}} + \frac{q}{q-2}\sigma^2, && q > 2,\\
Var(Y) &= \frac{q}{q-2}\sigma^2 - \frac{2}{\pi}(\frac{q}{q-1})^2\frac{\sigma^2\lambda^2}{1+\lambda^2}, && q > 2
\end{aligned}$$

which are given in Proposition 4 of Wang and Genton (2006) in the univariate case (see also Basso et al., 2010).

Note that the measure of skewness $r_1 = \frac{E(X-E(X))^3}{[Var(X)]^{3/2}}$ and kurtosis $r_2 = \frac{E(X-E(X))^4}{[Var(X)]^2} - 3$ depends on $a_k, k = 2, 3, 4$, the moments of $SEL_1(0, 1, \lambda, g)$-distribution. These measures are obtained in the following special case.

3.2 A Special Case

Let Y be given by (3.1), where X has a standard skew-t distribution with r degrees of freedom, $ST_1(0, 1, \lambda, r)$, which has the p.d.f. given by (2.2) with $\mu = 0$, $\sigma = 1$ and the generator function

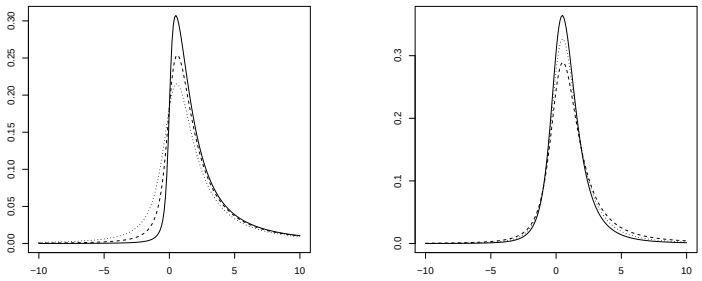
$$g(t) = \frac{\Gamma(\frac{(1+r)}{2})}{\Gamma(\frac{r}{2})\sqrt{\pi r}}(1 + \frac{t}{r})^{-\frac{(1+r)}{2}}, \quad t\epsilon R. \tag{3.6}$$

We will call this distribution Slash Skew-t (SLST) distribution and denoted it by $Y \sim SLST_1(\mu, \sigma, \lambda, q, r)$. To illustrate the tail behavior and skewness of the SLSEL distribution, we draw the density curve of $SLST_1(0, 1, \lambda, 1, 3)$ and $SLST_1(0, 1, 1, q, 3)$ distributions for some values of λ and q in Figure 1. We can see that when λ gets larger, the curve becomes more skewed, and when q gets smaller, the curve becomes heavier. Also we draw the density curve of t_3, $ST_1(0, 1, 5, 3)$, $SLST_1(0, 1, 5, 1, 3)$ and $SLSN_1(0, 1, 5, 1)$ distributions in Figure 2. We see that the $SLST_1(0, 1, 5, 1, 3)$ is more skewed and heavier than the other distributions.

From Azzalini and Capitano (2003), the moments of skew-t distribution, $ST_1(0, 1, \lambda, r)$, are given by

$$a_1 = c\delta, \quad a_2 = \frac{r}{r-2},\ r > 2,$$

$$a_3 = \frac{r}{r-3} c\delta(3 - \delta^2),\ r > 3, \quad a_4 = \frac{3r^2}{(r-2)(r-4)},\ r > 4,$$



(a) **(b)**

FIGURE.1 a Density function of the $SLST_1(0, 1, \lambda, 1, 3)$ for $\lambda = 1$ (dotted line), $\lambda = 2$ (dashed line) and $\lambda = 5$ (solid line)
b Density function of the $SLST_1(0, 1, 1, q, 3)$ for q=2 (dashed line), q=3 (dotted line) and q=5 (solid line)

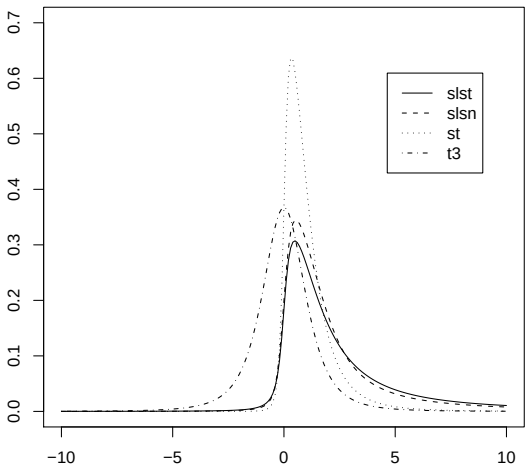


FIGURE. 2 Density function of the $SLST_1(0, 1, 5, 1, 3)$, $SLSN_1(0, 1, 5, 1)$, $ST_1(0, 1, 5, 3)$ and t_3

where $c = \sqrt{\frac{r}{\pi}}\frac{\Gamma(\frac{r-1}{2})}{\Gamma(\frac{r}{2})}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. So, from Theorem 3.3, the measures of skewness and kurtosis of $SLST_1(\mu, \sigma, \lambda, q, r)$ can be computed and are given by

$$r_1 = \frac{\frac{q}{q-3}\frac{r}{r-3}c\delta(3-\delta^2) + 2(\frac{q}{q-1})^3c^3\delta^3 - 3\frac{q^2}{(q-1)(q-2)}\frac{r}{r-2}c\delta}{(\frac{q}{q-2}\frac{r}{r-2} - (\frac{q}{q-1})^2c^2\delta^2)^{3/2}}, \quad q > 3, \quad r > 3,$$

and

$$r_2 = \frac{3\frac{q}{q-4}\frac{r^2}{(r-2)(r-4)} + 6\frac{q}{q-2}\frac{r}{r-2}(\frac{q}{q-1})^2c^2\delta^2 - 4\frac{q}{q-3}\frac{r}{r-3}\frac{q}{q-1}c^2\delta^2(3-\delta^2) - 3(\frac{q}{q-1})^4c^4\delta^4}{(\frac{q}{q-2}\frac{r}{r-2} - (\frac{q}{q-1})^2c^2\delta^2)^2} - 3,$$
$$q > 4, \quad r > 4,$$

respectively. Note that r_1 and r_2 are complicated functions of δ, q and r. We use the Mathematica software to find the behavior of these functions. For fixed $q > 3$ and $r > 3$, r_1 is an increasing odd function of δ. For $\delta > 0$ ($\lambda > 0$) we have $0 < r_1 < \infty$ with $\lim_{q\to 3^+} r_1 = \infty$, for $\delta < 0$ ($\lambda < 0$) we have $-\infty < r_1 < 0$ with $\lim_{q\to 3^+} r_1 = -\infty$, and $r_1 = 0$ for $\delta = 0$ ($\lambda = 0$). Also, for fixed $q > 4$ and $r > 4$, r_2 is a convex and even function of δ with minimum $r_2^* = \frac{3(q-2)^2(r-2)}{q(q-4)(r-4)} - 3$ at $\delta = 0$. So,

$$r_2 \geq r_2^* \geq \frac{3(r-2)}{r-4} - 3 \geq 0,$$

and $\lim_{q\to 4^+} r_2 = +\infty$. Therefore $-\infty < \delta < +\infty$ and $\frac{3(r-2)}{r-4} - 3 \leq r_2 < +\infty$, and the ranges of skewness and kurtosis depend on the parameters δ, q and r.

Note that if $X \sim ST_1(0, 1, \lambda, r)$, then X can be represented by $X = \frac{Z}{\sqrt{V}}$, where $Z \sim SN_1(0, 1, \lambda)$ and $V \sim \frac{1}{r}\chi_r^2$ (chi-square distribution with r degrees of freedom), see Azzalini and Capitanio (2003). Also if $Z \sim SN_1(0, 1, \lambda)$ then from Henze (1986) a stochastic representation of Z is given by

$$Z = \delta|T_0| + (1-\delta^2)^{\frac{1}{2}}T_1, \qquad \delta = \frac{\lambda}{\sqrt{1+\lambda^2}} \tag{3.7}$$

where T_0 and T_1 are independent standard normal random variables and $|.|$ denotes absolute value.

Now, if $Y \sim SLST_1(\mu, \sigma, \lambda, q, r)$ then from (3.1), $Y = \mu + \sigma\frac{X}{U^{\frac{1}{q}}}$ where

$X \sim ST_1(0,1,\lambda,r)$. So Y has the following stochastic representation

$$\begin{aligned} Y &= \mu + \sigma\delta U^{-\frac{1}{q}} V^{-\frac{1}{2}} |T_0| + \sigma(1-\delta^2)^{\frac{1}{2}} U^{-\frac{1}{q}} V^{-\frac{1}{2}} T_1 \\ &= \mu + \alpha T + \beta^{\frac{1}{2}} U^{-\frac{1}{q}} V^{-\frac{1}{2}} T_1 \end{aligned} \tag{3.8}$$

where $\alpha = \sigma\delta$, $\beta = \sigma^2(1-\delta^2)$ and $T = U^{-\frac{1}{q}} V^{-\frac{1}{2}} |T_0|$. This representation can be used to simulate realizations of Y and to implement the EM-type algorithm.

4 Maximum Likelihood Estimation via an EM-Type Algorithm

In this section we estimate the parameters of the new skew-slash distribution. We consider the special case of SLST distribution and estimate its parameter by maximum likelihood method. For maximum likelihood estimation of the parameters of $SLST_1(\mu,\sigma,\lambda,q,r)$ we used an EM-type algorithm (Dempster et al., 1977) which is used similarly by Basso et al. (2010) for estimation of the parameters of scale mixture of skew-normal distributions. Let $\mathbf{Y} = (Y_1, ..., Y_n)^T$ be a random sample of size n from $SLST_1(\mu,\sigma,\lambda,q,r)$ distribution. For the simplicity in computation, we set $q = 2$. Consider the stochastic representation (3.8) for $Y_i,\ i = 1,2,...,n$. Following the EM algorithm, let $(Y_i, U_i, V_i),\ i = 1,2,...,n$ be the complete data, where Y_i is called observed data and U_i, V_i are considered as missing data. Let $\theta = (\mu,\sigma,\lambda,r)$. From (3.8), for $i = 1,2,...,n$ we have

$$\begin{aligned} Y_i|u_i,v_i,t_i &\sim N(\mu + \alpha t_i, u_i^{-1}v_i^{-1}\beta), \qquad U_i \sim U(0,1) \\ T_i|u_i,v_i &\sim N(0, u_i^{-1}v_i^{-1})I(0,\infty), \qquad V_i \sim \frac{1}{r}\chi_r^2 \\ T_i|y_i,u_i,v_i &\sim N(\mu_{T_i}, u_i^{-1}v_i^{-1}M_T^2)I(0,\infty) \end{aligned}$$

where $\mu_{T_i} = \frac{\alpha}{\beta+\alpha^2}(y_i - \mu)$, $M_T^2 = \frac{\beta}{\beta+\alpha^2}$ and $N(a,b^2)I_{(0,\infty)}$ denote the truncated normal distribution $N(a,b^2)$ on $(0,\infty)$. Since

$$f(y_i,u_i,v_i,t_i) = f(y_i|u_i,v_i,t_i)f(t_i|u_i,v_i)f(u_i)f(v_i),$$

so, the log-likelihood function for the complete data is given by

$$L_c(\theta) = c - \frac{n}{2}\ln(\beta) - \frac{nr}{2}ln(\frac{r}{2}) - n\ln(\Gamma(\frac{r}{2})) - \frac{1}{2\beta}\sum_{i=1}^{n} u_i v_i (y_i - \mu - \alpha t_i)^2$$
$$+(\frac{r}{2} - 1)\sum_{i=1}^{n}\ln(v_i) - \frac{r}{2}\sum_{i=1}^{n} v_i$$

where c is a constant that is independent of θ.
Now the conditional expectation of $L_c(\theta)$ given the observed data y_i and the current estimates of the parameters, say $\hat{\theta}$, is given by

$$\begin{aligned} Q(\theta|\hat{\theta}) &= E(L_c(\theta)|y_i,\hat{\theta}) \qquad (4.1)\\ &= c-\frac{n}{2}\ln(\beta)-\frac{nr}{2}\ln(\frac{r}{2})-n\ln(\Gamma(\frac{r}{2}))\\ &\quad -\frac{1}{2\beta}\sum_{i=1}^{n}(E(V_iU_i|y_i,\hat{\theta})(y_i-\mu)^2\\ &\quad +\alpha^2E(V_iU_iT_i^2|y_i,\hat{\theta})-2\alpha(y_i-\mu)E(V_iU_iT_i|y_i,\hat{\theta}))\\ &\quad +(\frac{r}{2}-1)\sum_{i=1}^{n}E(\ln(V_i)|y_i,\hat{\theta})-\frac{r}{2}\sum_{i=1}^{n}E(V_i|y_i,\hat{\theta}). \qquad (4.2)\end{aligned}$$

Let

$$\begin{aligned} \hat{k}_i &= E(V_iU_i|y_i,\hat{\theta}), & \hat{S}_{4i} &= E(\ln(V_i)|y_i,\hat{\theta})\\ \hat{S}_{2i} &= E(V_iU_iT_i|y_i,\hat{\theta}), & \hat{S}_{5i} &= E(V_i|y_i,\hat{\theta})\\ \hat{S}_{3i} &= E(V_iU_iT_i^2|y_i,\hat{\theta}), & \hat{S}_i &= E(V_i^{\frac{1}{2}}U_i^{\frac{1}{2}}W_\Phi(\frac{\mu_{T_i}V_i^{\frac{1}{2}}U_i^{\frac{1}{2}}}{M_T})|y_i,\hat{\theta}) \qquad (4.3)\end{aligned}$$

where $W_{\Phi(x)}=\frac{\phi(x)}{\Phi(x)}$ and ϕ, Φ are standard normal density and cumulative distribution function, respectively.

Since $Y_i|u_i,v_i\sim SN_1(\mu,u_i^{-\frac{1}{2}}v_i^{-\frac{1}{2}}\sigma,\lambda)$, so it is easy to show that

$$\begin{aligned} \hat{k}_i &= \frac{2}{\hat{\sigma}\hat{r}\sqrt{\hat{r}}f(y_i)}\int_0^1\int_0^\infty u_i^{\frac{3}{2}}w_i^{\frac{3}{2}}p(w_i)\phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))\\ &\quad \times\Phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}\hat{\lambda}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))dw_idu_i,\\ \hat{S}_i &= \frac{2}{\hat{\sigma}\hat{r}f(y_i)}\int_0^1\int_0^\infty u_iw_ip(w_i)\phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))\\ &\quad \times\phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}\hat{\lambda}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))dw_idu_i,\\ \hat{S}_{2i} &= \hat{\mu}_{T_i}\hat{k}_i+\hat{M}_T\hat{S}_i,\\ \hat{S}_{3i} &= \hat{\mu}_{T_i}^2\hat{k}_i+\hat{\mu}_{T_i}\hat{M}_T\hat{S}_i+\hat{M}_T^2,\end{aligned}$$

$$\begin{aligned}
\hat{S}_{4i} &= \frac{2}{\hat{\sigma}\hat{r}\sqrt{\hat{r}}f(y_i)}\int_0^1\int_0^\infty u_i^{\frac{1}{2}}w_i^{\frac{3}{2}}p(w_i)\phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))\\
&\quad\times\Phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}\hat{\lambda}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))dw_idu_i,,\\
\hat{S}_{5i} &= \frac{2}{\hat{\sigma}\hat{r}\sqrt{\hat{r}}f(y_i)}\int_0^1\int_0^\infty u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}\ln(\frac{w_i}{r})p(w_i)\phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))\\
&\quad\times\Phi(u_i^{\frac{1}{2}}w_i^{\frac{1}{2}}\hat{\lambda}(\frac{y_i-\hat{\mu}}{\hat{\sigma}\sqrt{\hat{r}}}))dw_idu_i, \qquad (4.4)
\end{aligned}$$

where $f(y_i)$ is the p.d.f of $SLST_1(\hat{\mu},\hat{\sigma},\hat{\lambda},2,\hat{r})$ (given by (3.2) with generator (3.6)) and $p(w_i)$ is the p.d.f of χ^2_r.

Now consider the E-step of the algorithm. In this step, given the observation y_i and current estimates $\hat{\theta}$, the conditional expectations $\hat{k_i}$, $\hat{S_i}$, $\hat{S_{2i}}$, $\hat{S_{3i}}$, $\hat{S_{4i}}$ and $\hat{S_{5i}}$ must be computed. Note that for computing $\hat{k_i}$, $\hat{S_i}$, $\hat{S_{4i}}$ and $\hat{S_{5i}}$, Monte Carlo integration can be employed, which yields the so called MC-EM algorithm. For the M-step of the algorithm, we maximize the expected complete-data function over θ, or the Q-function which from (4.2) and (4.3) is given by

$$\begin{aligned}
Q(\theta|\hat{\theta}^{(k)}) &= E(L_c(\theta)|\mathbf{y},\hat{\theta}^{(k)})\\
&= c-\frac{n}{2}\ln(\beta)-\frac{nr}{2}\ln(\frac{r}{2})-n\ln(\Gamma(\frac{r}{2}))\\
&\quad-\frac{1}{2\beta}\sum_{i=1}^n(\hat{k}_i^{(k)}(y_i-\mu)^2+\alpha^2\hat{S}_{3i}^{(k)}-2\alpha(y_i-\mu)\hat{S}_{2i}^{(k)})\\
&\quad+(\frac{r}{2}-1)\sum_{i=1}^n\hat{S}_{4i}^{(k)}-\frac{r}{2}\sum_{i=1}^n\hat{S}_{5i}^{(k)}, \qquad (4.5)
\end{aligned}$$

where $\hat{\theta}^{(k)}$ is an updated value of $\hat{\theta}$. From (4.5), we can derive explicit relations between μ, α and β which maximizes (4.5), but this is not the case for r. In this case the M-step can be replaced with a sequence of conditional maximization (CM) steps. The resulting method known as ECM algorithm (Meng and Rubin, 1993). Liu and Rubin (1994) introduced an ECME algorithm which is maximizing the constrained Q-function with some CM-steps that maximizes the corresponding constrained actual marginal likelihood function, called CML-steps. Similar to Basso et al. (2010), we use ECME algorithm as follows.

E-step: Given a current estimate $\hat{\theta}^{(k)}=(\hat{\mu}^{(k)},\hat{\sigma}^{(k)},\hat{\lambda}^{(k)},\hat{r}^{(k)})$ and observation $\mathbf{y}=(y_1,...,y_n)$, compute $\hat{k}_i^{(k)}$ and $\hat{S}_i^{(k)}$, $i=1,2,...,n$ from (4.4) by Monte Carlo integration and then compute $\hat{S}_{2i}^{(k)}$ and $\hat{S}_{3i}^{(k)}$.

CM-steps: Derive $\hat{\theta}^{(k+1)}$ by maximizing $Q(\theta|\hat{\theta}^{(k)})$ over θ, which are given by the following closed form expressions

$$\begin{aligned}
\hat{\mu}^{(k+1)} &= (\sum_{i=1}^{n}(y_i\hat{k}_i^{(k)} - \hat{\alpha}^{(k)}\hat{S}_{2i}^{(k)}))(\sum_{i=1}^{n}\hat{k}_i^{(k)})^{-1} \\
\hat{\alpha}^{(k+1)} &= (\sum_{i=1}^{n}(y_i - \hat{\mu}^{(k+1)})\hat{S}_{2i}^{(k)})(\sum_{i=1}^{n}\hat{S}_{3i}^{(k)})^{-1} \\
\hat{\beta}^{(k+1)} &= \frac{1}{n}\sum_{i=1}^{n}(\hat{k}_i^{(k)}(y_i - \hat{\mu}^{(k+1)})^2 + (\hat{\alpha}^{(k+1)})^2\hat{S}_{3i}^{(k)} \\
&\quad -2\hat{\alpha}^{(k+1)}(y_i - \hat{\mu}^{(k+1)})\hat{S}_{2i}^{(k)}) \\
\hat{\sigma^2}^{(k+1)} &= (\hat{\alpha}^{(k+1)})^2 + \hat{\beta}^{(k+1)} \qquad \hat{\lambda}^{(k+1)} = \frac{\hat{\alpha}^{(k+1)}}{\sqrt{\hat{\beta}^{(k+1)}}}.
\end{aligned}$$

CML-step: Derive $\hat{r}^{(k+1)}$ by maximizing the actual marginal log-likelihood function as follows

$$\hat{r}^{(k+1)} = argmax_r \sum_{i=1}^{n} \log(f(y_i; \hat{\mu}^{(k+1)}, \hat{\sigma}^{(k+1)}, \hat{\lambda}^{(k+1)}, 2, r)),$$

where $f(y_i;\theta)$ is the $SLST_1$ density.

The algorithm iterates between the E and M steps until reach convergence.

4.1 The observed information matrix

In this subsection we evaluate the observed information matrix of the SLST distribution, defined by

$$\mathbf{J}_0(\Theta|\mathbf{y}) = -\frac{\partial^2 l(\Theta|\mathbf{y})}{\partial\Theta\partial\Theta^T},$$

where $l(\Theta|\mathbf{y})$ is the incomplete likelihood function based on observation $\mathbf{y}$. Under some regularity conditions, the covariance matrix of the maximum likelihood estimates $\hat{\Theta}$ can be approximated by the inverse of $\mathbf{J}_0(\Theta|\mathbf{y})$. The observed information matrix can be obtained as follows,

$$\mathbf{J}_0(\hat{\Theta}|\mathbf{y}) = \sum_{i=1}^{n} \hat{\mathbf{t}}_i\hat{\mathbf{t}}_i^T,$$

where

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

see Basford et al. (1997) and Lin et al. (2007).
Now, partitioned $\hat{\mathbf{t}}_i$ into components corresponding to all the parameters in Θ, i.e.,

$$\hat{\mathbf{t}}_i = (\hat{t}_{i,\mu}, \hat{t}_{i,\sigma}, \hat{t}_{i,\lambda}, \hat{t}_{i,r})^T,$$

where

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

Define

$$\begin{aligned}
I_{i1}^{F_1}(v_1,v_2) &= \int_0^1 u^{\frac{1}{2}+v_1}(1+\frac{u(y_i-\mu)^2}{r\sigma^2})^{-\frac{r+v_2}{2}} \\
&\quad \times F_1(\lambda u^{\frac{1}{2}}(y_i-\mu)(\frac{(r+1)}{r(r\sigma^2+u(y_i-\mu)^2)})^{\frac{1}{2}};r+1)du, \\
I_{i1}^{f_1}(v_1,v_2,v_3) &= \int_0^1 u^{\frac{1}{2}+v_1}(r\sigma^2+u(y_i-\mu)^2)^{-\frac{v_2}{2}} \\
&\quad \times f_1(\lambda u^{\frac{1}{2}}(y_i-\mu)(\frac{(r+1)}{r(r\sigma^2+u(y_i-\mu)^2)})^{\frac{1}{2}};r+1)du, \\
I_{i2}^{F_1}(v_1,v_2) &= \int_0^1 u^{\frac{1}{2}+v_1}(\sigma^{\frac{2}{r+1}}+u\frac{\sigma^{-\frac{2r}{r+1}}}{r}(y_i-\mu)^2)^{-\frac{r+v_2}{2}} \\
&\quad \times F_1(\lambda u^{\frac{1}{2}}(y_i-\mu)(\frac{(r+1)}{r(r\sigma^2+u(y_i-\mu)^2)})^{\frac{1}{2}};r+1)du, \\
I_{i2}^{f_1}(v_1,v_2,v_3) &= \int_0^1 u^{\frac{1}{2}+v_1}(\sigma^{\frac{2}{r+1}}+u\frac{\sigma^{-\frac{2r}{r+1}}}{r}(y_i-\mu)^2)^{-\frac{r+v_2}{2}} \\
&\quad \times (r\sigma^2+u(y_i-\mu)^2)^{-\frac{v_2}{2}} \\
&\quad \times f_1(\lambda u^{\frac{1}{2}}(y_i-\mu)(\frac{(r+1)}{r(r\sigma^2+u(y_i-\mu)^2)})^{\frac{1}{2}};r+1)du,
\end{aligned}$$

where $f_1(x;r+1)$ and $F_1(x;r+1)$ are p.d.f. and c.d.f. of Student-t distribution with $r+1$ degree of freedom.

After some algebraic manipulation, we obtain

$$\begin{aligned}
\frac{\partial}{\partial\mu}(f(y_i;\Theta)) &= \frac{2}{\sigma}\frac{\Gamma(\frac{r+1}{2})}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\Big(\frac{(r+1)}{r}\frac{(y_i-\mu)}{\sigma^2}I_{i1}^{F_1}(1,3)\\
&\quad -\lambda\Big(\frac{r+1}{r}\Big)^{\frac{1}{2}}I_{i1}^{f_1}\Big(\frac{1}{2},1,3\Big)\\
&\quad +\lambda\Big(\frac{r+1}{r}\Big)^{\frac{1}{2}}(y_i-\mu)^2I_{i1}^{f_1}\Big(\frac{3}{2},1,3\Big)\Big),\\
\frac{\partial}{\partial\sigma}(f(y_i;\Theta)) &= \frac{2\Gamma(\frac{r+1}{2})}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\Big((y_i-\mu)^2\sigma^{-\frac{-3r+1}{r+1}}I_{i2}^{F_1}(1,3)\\
&\quad -\sigma^{-\frac{1-r}{r+1}}I_{i2}^{F_1}(0,3)-\lambda\frac{(r(r+1))^{\frac{1}{2}}}{2}(y_i-\mu)I_{i2}^{f_1}\Big(\frac{1}{2},1,3\Big)\Big),\\
\frac{\partial}{\partial\lambda}(f(y_i;\Theta)) &= \frac{2}{\sigma}\frac{(r+1)^{\frac{1}{2}}\Gamma(\frac{r+1}{2})}{r\pi^{\frac{1}{2}}\Gamma(\frac{r}{2})}(y_i-\mu)I_{i1}^{f_1}\Big(\frac{1}{2},1,1\Big),\\
\frac{\partial}{\partial r}(f(y_i;\Theta)) &= \frac{2}{\sigma}\Big(\frac{(\frac{\partial}{\partial r}\Gamma(\frac{r+1}{2}))\Gamma(\frac{r}{2})-\frac{1}{2r}\Gamma(\frac{r+1}{2})\Gamma(\frac{r}{2})-\Gamma(\frac{r+1}{2})(\frac{\partial}{\partial r}\Gamma(\frac{r}{2}))}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\\
&\quad \times I_{i1}^{F_1}(0,1)+\frac{\Gamma(\frac{r+1}{2})}{2\pi^{\frac{1}{2}}\Gamma(\frac{r}{2})}\frac{(r+1)}{\sigma^2r^{\frac{5}{2}}}(y_i-\mu)^2I_{i1}^{F_1}(1,3)\\
&\quad +\frac{\lambda}{r^2(r+1)^{\frac{1}{2}}}\frac{\Gamma(\frac{r+1}{2})}{2\pi^{\frac{1}{2}}\Gamma(\frac{r}{2})}(y_i-\mu)I_{i1}^{f_1}\Big(\frac{1}{2},1,1\Big)\\
&\quad +\lambda\sigma^2\frac{\Gamma(\frac{r+1}{2})}{2\pi^{\frac{1}{2}}\Gamma(\frac{r}{2})}\frac{(r+1)^{\frac{1}{2}}}{r}(y_i-\mu)I_{i1}^{f_1}\Big(\frac{1}{2},1,3\Big)\Big).
\end{aligned}$$

Note that the elements of the information matrix are very complicated. So, we cannot establish the non singularity of the observed information matrix. But for the SLST model, from the simulation study we observe that this matrix is nonsingular for finite degrees of freedom.
In the next section we use the above techniques to estimate the parameters.

4.2 Sensitivity Analysis

For checking the influence of observations on the ML estimators, we use a sensitivity analysis with scale-deletion method in this section and

Section 5.3, to detect observations that under small perturbation of the model exert great influence on the ML estimates. This method has been used in some papers, such as Cook (1977), Bolfarine et al. (2007) and Lin et al. (2009). We use the case-deletion approach to detect the influence of removing the i-th case from the analysis by evaluating the metrics such as the likelihood distance and Cook's distance (see Cook, 1977).

Let $\hat{\Theta}_{(i)}$ be the ML estimate of Θ 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 in 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

$$GD_i(\Theta) = (\hat{\Theta}_{(i)} - \hat{\Theta})^T[-\ddot{L}(\Theta)](\hat{\Theta}_{(i)} - \hat{\Theta}),$$

where $-\ddot{L}(\Theta) = \frac{\partial^2 l(\Theta|\mathbf{y})}{\partial\Theta\partial\Theta^T}$ is the observed information matrix (in $\Theta = \hat{\Theta}$) evaluated in this section.
Another measure of case-deletion approach is the likelihood distance which is defined by

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

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

5 Applications

In this section, we present two applications of SLSEL distribution. The first one is a small simulation study and the other is a statistical analysis of real data sets.

5.1 Small Simulation Studies

We perform a small simulation study to investigate the ML estimators that proposed in Section 4. In the simulation study, we investigate asymptotic properties of the estimate obtained by ECME algorithm. We generate random sample of sizes $n = 100, 500, 1000, 5000$ and 10000 from

$SLST_1(0, 1, 1, 2, 5)$. For each sample of size n, we repeat the sampling 100 times and compute the simulation bias and MSE which are defined as

$$\begin{aligned} Bias(\hat{\gamma}^{(j)}) &= \frac{1}{100}\sum_{i=1}^{100}\hat{\gamma}_i^{(j)} - \gamma^{(j)} \\ MSE(\hat{\gamma}^{(j)}) &= \frac{1}{100}\sum_{i=1}^{100}(\hat{\gamma}_i^{(j)} - \gamma^{(j)})^2, \quad j = 1, 2, 3, 4, \end{aligned}$$

where $\gamma^{(j)}$ and $\hat{\gamma_i}^{(j)}$, $j = 1, 2, 3, 4$ stands for parameters μ, σ, λ, r and their ECME estimates when the data is sample i, respectively. Figure 3 shows the graph of $Bias(\hat{\gamma}^{(j)})$ versus the sample size n. From this figure we see that the bias of $\hat{\mu}, \hat{\sigma}, \hat{\lambda}, \hat{r}$ converges to zero when the sample size is increased. A similar result was happening for the MSE of these estimators.

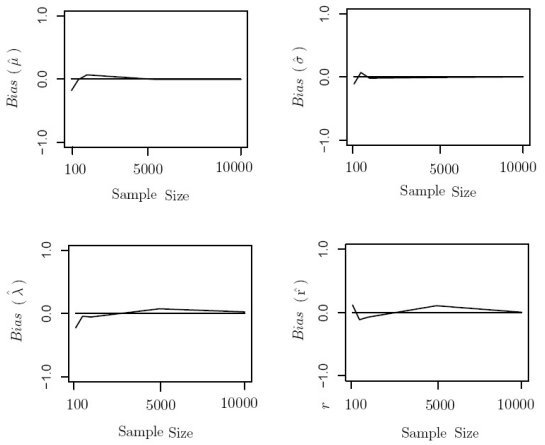


FIGURE. 3 Graph of bias versus sample size n for ECME estimates of **a** $\hat{\mu}$, **b** $\hat{\sigma}$, **c** $\hat{\lambda}$ and **d** $\hat{r}$

5.2 Real Data Application

We consider here the fiber-glass data set analyzed by Jones and Faddy (2003) and Azzalini and Capitanio (2003) in two forms of skew-t distribution and Wang and Genton (2006) in the slash skew-normal (SLSN) distribution. They note skewness on the left as well as heavy tail behavior. We fit a $SLSN_1(\mu, \sigma, \lambda, 2)$ and a $SLST_1(\mu, \sigma, \lambda, 2, r)$ distribution to this data set. The maximum likelihood estimates of the pa-

rameters via ECME algorithm and their standard errors (in parenthesis) are $\hat{\mu} = 1.77(0.6703)$, $\hat{\sigma} = 0.21(0.0481)$, $\hat{\lambda} = -1.73(0.6829)$ and $\hat{\mu} = 1.69(0.0622)$, $\hat{\sigma} = 0.15(0.0628)$, $\hat{\lambda} = -1.1(1.3475)$, $\hat{r} = 4.1(2.3732)$ for SLSN and SLST distributions, respectively. The histogram of this data set and the fitted density curves are plotted in Figure 4. For model comparison, we also computed the Akaike Information Criterion (AIC) (Akaike, 1974) and the Efficient Determination Criterion (EDC) (Bai et al., 1989). These values are $AIC = 35.41$ and $EDC = 34.17$ for SLSN distribution and $AIC = 34.59$ and $EDC = 32.94$ for SLST distribution. From these criteria and Figure 4, we see that SLST distribution has a better fit than the SLSN distribution to this data set. Note that we use a similar ECM algorithm as in Section 4 to estimate the parameters of SLSN distribution.

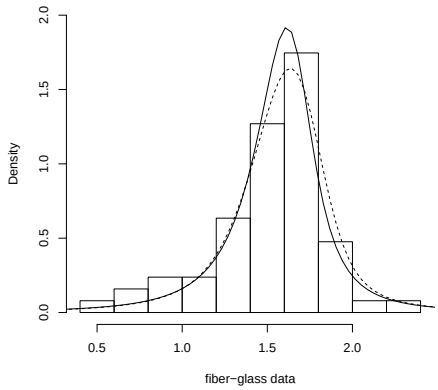
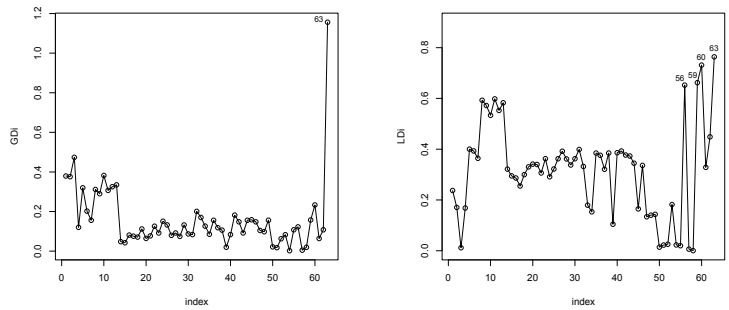


FIGURE. 4 Histogram of fiber-glass data set with fitted $SLSN_1(\mu, \sigma, \lambda, 2)$ distribution (dashed line) and $SLST_1(\mu, \sigma, \lambda, 2, r)$ distribution (solid line)

5.3 Sensitivity Analysis

In this section, we use the real data set to find the points which are influential in parameter estimation. Let $\hat{\Theta}$ be the ML estimate of Θ in fiber-glass data and $\hat{\Theta}_i$ be the ML estimate of Θ without the i-th observation, then we compute the GD_i and LD_i as diagnostics for global influence. The measures GD_i and LD_i computed and presented in Figures 5.a and 5.b, respectively. From these figures we observe that the case 63 is identified as the most influential and cases 56, 59 and 60 are influential.



(a) **(b)**

FIGURE.5 a Index plot of GD_i for case weights perturbation for fiber-glass data

b Index plot of LD_i for case weights perturbation for fiber-glass data

6 Conclusion and Future Work

In this article, we present a new class of asymmetric distributions, called SLSEL distributions, to find more flexible models which present features of the data. We investigate some properties of the proposed model. Also, we use the EM-type algorithm to obtain the maximum likelihood estimates. Furthermore, we illustrate our method with a real data set and show that the SLSEL model has better performance than the other competitors' ones. Finally, we use the case-deletion approach to the proposed model to check the influence of observations on ML estimators. Results obtained from the real data sets show the usefulness of the approach.

Another special case of this work is to define the Skew-Slash Contaminated Normal (SLSCN) distribution which can be defined as follows:
Let Y be given by (3.1), where X has the p.d.f. given by (2.2) with $\mu = 0$, $\sigma = 1$ and the generator function

$$g(t) = \upsilon\phi(t; 0, \frac{1}{\gamma}) + (1 - \upsilon)\phi(t; 0, 1), \qquad t\epsilon R.$$

We denoted it by $Y \sim SLSCN_1(\mu, \sigma, \lambda, q, \upsilon, \gamma)$. The SLSCN is a flexible distribution which can be used to model the data with skewness and/or heavy tailed. Because of the complexity of this model, further works

are needed for finding the ML estimates via the EM algorithm of the SLSCN distribution.

Acknowledgments

The authors are grateful to the Associate Editor, and anonymous referees for making helpful comments and suggestions on an earlier version of this article. Research of the second author was supported by the research council of Allameh Tabataba'i University.

References

Akaike, H. (1974), A new look at the statistical model identification. IEEE Transactions on Automatic Control, **19**, 716-723.

Arslan, O. (2008), An alternative multivariate skew-slash distribution. Statist. Probab. Lett., **78**, 2756-2761.

Arslan, O. (2009), Maximum likelihood parameter estimation for the multivariate skew-slash distribution. Statist. Probab. Lett., **79**, 2158-2165.

Arslan, O. and Genc, A. I. (2009), A generalization of the multivariate slash distribution. J. Statist. plann. Inference, **139**, 1164-1170.

Azzalini, A. (1985), A class of distributions which includes the normal ones. Scand. J. Statist., **12**, 171-178.

Azzalini, A. (1986), Further results on a class of distributions which includes the normal ones. Statistica, **46**, 199-208.

Azzalini, A. and Capitanio, A. (1999), Statistical applications of the multivariate skew-normal distribution. J. Roy. Statist. Soc. Ser. B, **61**, 579-602.

Azzalini, A. and Capitanio, A. (2003), Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution. J. Roy. Statist. Soc., Ser. B, **65**, 367-389.

Azzalini, A. and Dalla Valle, A. (1996), The multivariate skew-normal distribution. Biometrika, **83**, 715-726.

Bai, Z. D., Krishnaiah, P. R., and Zhao, L. C. (1989), On rates of convergence of efficient detection criteria in signal processing with white noise. IEEE Transactions on Information Theory, **35**, 380-388.

Basford K. E., Greenway, D. R., Mclachlan, G. J., and Peel, D. (1997), Standard errors of fitted component means of normal mixtures. Comput. Statist., **12**, 1-17.

Basso, R. M., Lachos, V. H., Cabral, C. R. B., and Ghosh, P., (2010), Robust mixture modeling based on the scale mixtures of skew-normal distributions. Comput. Statist. Data Analysis **54**, 2926-2941.

Bolfarine, H., Montenegro, L. C., and Lachos, V.H. (2007), Inference diagnostics for skew-normal linear mixed models. Sankhya, **69**, 648-670.

Branco, M. D. and Dey, D. K., (2001), A general class of multivariate skew-elliptical distributions. J. Multivariate Anal., **79**, 99-113.

Cook, R. D. (1977), Detection of influential observation in linear regression. Technometrics, **19**, 5-18.

Dempster, A. P., Liard, N. M., and Rubin, D. B. (1977), Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc., Ser. B, **39**, 1-38.

Fang, K. T., Kotz, S., and Ng, K. W. (1990), Symmetric multivariate and related distributions. New York: Chapman and Hall.

Genc, A. I. (2007), A generalization of the univariate slash by a scale-mixtured exponential power distribution. Commun. Statist. Simulation Comput., **36**, 937-947.

Genton, M. G. (2004), Skew-elliptical distributions and their applications: A Journey Beyond Normality. chapman and Hall/ CRC, Boca Raton, FL, Edited, Volume.

Gomez, H. W., Quintana, F. A., and Torres, J. (2007), A new family of slash-distributions with elliptical contours. Statist. Probab. Lett., **77**, 717-725.

Gomez, H. W. and Venegas, O. (2008), Erratum to: A new family of slash-distributions with elliptical contours. Statist. Probab. Lett., **78**, 2273-2274.

Henz., N. (1986), A probabilistic representation of the skew-normal distribution. Scan. J. Statist., **13**, 271-275.

Jones, M. C. and Faddy, M. J. (2003), A skew extension of the t distribution with applications. J. Roy. Statist. Soc. Ser. B, **65**, 159-174.

Kafadar, K. (1982), A biweight approach to the one-sample problem. J. Amer. Statist. Assoc., **77**, 416-424.

Kafadar, K. (2001), Slash distribution. Encyclopedia.

Kelker, D. (1970), Distribution theory of spherical distributions and a location-scale parameter generalization. Sankhya, **32**, 419-430.

Lin, T. L., Lee, J. C., and Yen, S. Y. (2007), Finite mixture modeling using the skew-normal distribution. Statistica Sinica, **17**(b), 909-927.

Lin, J. G., Xie, F. C., and Wei, B. C. (2009), Statistical Diagnostics for skew-t-normal nonlinear models. Communications in Statistics - Simulation and Computation, **38**(10), 2096-2110.

Liu, C. and Rubin, D. B. (1994), The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. Biometrika, **81**, 633-648.

Meng, X. L. and Rubin, D. B. (1993), Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika, **80**, 267-278.

Mosteller, F. and Tukey, J. W. (1977), Data analysis and regression. Addison-Wesley, Reading, MA.

Rogers, W. H. and Tukey, J. W. (1972), Understanding some long-tailed symmetrical distributions. Statist. Neerlandica, **26**, 211-226.

Sahu, S. K., Dey, D. K. and Branco, M. D. (2003), A new class of multivariate skew distributions with applications to Bayesian regression models. Canad. J. Statist., **41**, 129-150.

Wang, J., Genton, M. G. (2006), The multivariate skew-slash distribution. J. Statist. plann. Inference, **136**, 209-220.