© Statistical Research and Training Center

مجله‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **12** (2015): 163–178 دوره‌ی ۱۲، شماره‌ی ۲، پاییز و زمستان ۱۳۹۴، صص ۱۶۳–۱۷۸

The Location-Scale Mixture of Generalized Gamma Distribution: Estimation and Case Influence Diagnostics

Z. Rahnamaei

Firoozkooh Branch, Islamic Azad University

Received: 2/17/2015 Approved: 1/23/2016

Abstract. One of the most interesting problems in distribution theory is constructing the distributions, which are appropriate for fitting skewed and heavy-tailed data sets. In this paper, we introduce a skew-slash distribution by using the scale mixture of the generalized gamma distribution. Some properties of this distribution are obtained. An EM-type algorithm is presented to estimate the parameters. Finally, we provide a simulation study and an application to real data to illustrate the modeling strength of the proposed distribution.

Keywords. EM algorithm; generalized gamma distribution; location-scale mixture of distribution; skew-slash distribution.

MSC 2010: 60E05, 62F10.

1 Introduction

In real application, the distribution of data is often skew, while virtually all robust methods assume symmetry of the error distribution. Moreover, the distribution of real data is seldom as heavily tailed, where the ones employed in theoretical robustness studies. To handle both skewness and heavy tails simulate, Stacy (1962) proposed the Generalized Gamma (GG) distribution, which is a flexible family of distributions. This distribution has varieties of shapes and hazard functions for modeling duration. It is appropriate for modeling data with skewness and heavy tail structure. Khodabin and Ahmadabadi (2010) obtained some properties of this distribution.

The probability density function (p. d. f.) of GG distribution is given by:

$$g_X(x) = \frac{k(x-\gamma)^{k\tau-1}}{\delta^{K\tau}\Gamma(\tau)} \exp\left(-\left(\frac{x-\gamma}{\delta}\right)^k\right), \tag{1}$$

where $-\infty < \gamma < \infty$, $\delta > 0$, $k > 0$ and $-\infty < x < \infty$ (denoted by $X \sim GG(\gamma, \delta, k, \tau)$).

Another distribution with skewness and heavy tail proposed by Arsalan (2008, 2009). It is constructed by location-scale mixture of a non-negative random variable with normal random variable as

$$Y = \mu + V^{-1}\beta + V^{-\frac{1}{2}}\sigma X, \tag{2}$$

where $-\infty < \mu, \beta < \infty$, $\sigma > 0$, $\alpha > 0$, $X \sim N(0,1)$, $V \sim \text{beta}(\alpha, 1)$ and X and V are independent. Arsalan called (2) as the Generalized Hyperbolic Skew-Slash ($GHSSE$) distribution, which is denoted by $Y \sim GHSSE(\mu, \sigma, \beta, \alpha)$. He investigated some properties of this distribution and estimated its parameters by the EM algorithm.

In this paper, by replacing X in (2) with GG distribution, we construct a new flexible distribution for fitting skewed and heavy-tailed data sets, which fits better than GG and $GHSSE$ distribution to such data sets.

This paper is organized as follows:

In Section 2, we introduce the Location-Scale Mixture of Generalized Gamma ($LSMGG$) distribution then investigate some distributional properties. Moreover a maximum likelihood method is presented to estimate its parameters. In Section 3, in order to investigate the performance of the proposed model, we present some simulation studies and a real data application. Conclusion is given in Section 4.

2 Location-Scale Mixture of Generalized Gamma Distribution

In this section, we define $LSMGG$ distribution and derive some distributional properties. We implement an EM-type algorithm for parameter estimation by Maximum Likelihood (ML) method and we utilize local influence approach to detect observations that affect on the ML estimators.

Definition 1. A random variable Y have $LSMGG$ distribution with location parameter μ, scale parameter σ, skew parameters β and τ, tail parameter α and shape parameter k, denoted by $Y \sim LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$, if X in (2) replaces by $GG(0, 1, k, \tau)$.

From (1) and independence of X and V, we have

$$f_Y(y) = \begin{cases} \frac{\alpha k}{\sigma^{k\tau}\Gamma(\tau)} \int_0^1 v^{(\alpha-1)+\frac{k\tau}{2}} (y-\mu-v^{-1}\beta)^{k\tau-1} \\ \times \exp\left(-\frac{v^{\frac{k}{2}}}{\sigma^k}(y-\mu-v^{-1}\beta)^k\right) dv; & y > \mu + v^{-1}\beta, \quad \beta \neq 0 \\ 0; & y \leqslant \mu + v^{-1}\beta, \end{cases} \tag{3}$$

where $-\infty < y < \infty$.

2.1 Some Properties of *LSMGG* Distribution

In this section, we consider some properties of $LSMGG$ distribution.

Theorem 1. *(An Invariance Result) If $Y \sim LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$ and $T = aY + b$, $(a > 0, b \in R)$ then $T \sim LSMGG(a\mu + b, a\sigma, a\beta, \alpha, k, \tau)$.*

Theorem 2. *If $Y \sim LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$ and $V \sim \text{beta}(\alpha, 1)$, then $Y|V = v \sim GG\left(\mu + v^{-1}\beta, \frac{\sigma}{\sqrt{v}}, k, \tau\right)$.*

Theorem 3. *If $Y|V = v \sim GG\left(v^{-1}\beta, \frac{1}{\sqrt{v}}, k, \tau\right)$ and $V \sim \text{beta}(\alpha, 1)$, then $Y \sim LSMGG(0, 1, \beta, \alpha, k, \tau)$*

Theorem 4. *If* $Y \sim LSMGG(0,1,\beta,\alpha,k,\tau)$*, then*

$$\mu_n = E(Y^n) = \frac{\beta^n}{\Gamma(\tau)}\sum_{t=0}^{n}\binom{n}{t}\beta^{-t}\left(1+\frac{2n-t}{2\alpha-(2n-t)}\right)\Gamma\left(\frac{k\tau+t}{k}\right), \qquad \alpha > \frac{2n-t}{2}.$$

The proof of above theorems is easily derived from (1) and (3) and is omitted.

From Theorem 4 we get

$$\begin{aligned}
E(Y) &= \frac{\alpha\beta}{\alpha-1} + \frac{2\alpha}{2\alpha-1}A_1, \quad \alpha > 1,\\
V(Y) &= \alpha\beta^2\left(\frac{1}{\alpha-2} - \frac{\alpha}{(\alpha-1)^2}\right)\\
&\quad + 4\alpha\beta A_1\left(\frac{1}{2\alpha-3} - \frac{\alpha}{(\alpha-1)(2\alpha-1)}\right)\\
&\quad + \frac{\alpha}{\alpha-1}A_2 - \left(\frac{2\alpha}{2\alpha-1}A_1\right)^2, \quad \alpha > 2.
\end{aligned}$$

To find the measure of skewness of $LSMGG$, without loss of generality, let $\mu = 0$ and $\sigma = 1$, then the skewness of $LSMGG(\mu,\sigma,\beta,\alpha,k,\tau)$ can be computed by $r_1 = \dfrac{E(Y-E(Y))^3}{\mathrm{Var}^{\frac{3}{2}}(Y)}$, where

$$\begin{aligned}
E(Y-E(Y))^3 &= \frac{2\alpha(\alpha+1)}{(\alpha-1)^3(\alpha-2)(\alpha-3)}\beta^3\\
&\quad + \frac{24\alpha^6 - 144\alpha^5 + 306\alpha^4 - 252\alpha^3 + 66\alpha^2 - 36\alpha}{(\alpha-1)^2(\alpha-2)(2\alpha-1)(2\alpha-3)(2\alpha-5)}\beta^2 A_1\\
&\quad + \frac{3\alpha}{(\alpha-1)^2(\alpha-2)}A_2\beta\\
&\quad - \frac{24\alpha^2}{(\alpha-1)(2\alpha-1)^2(2\alpha-3)}A_1^2\beta\\
&\quad + \left(\frac{2\alpha}{2\alpha-3}A_3 - \frac{6\alpha^2}{(\alpha-1)(2\alpha-1)}A_1A_2 + \frac{16\alpha^3}{(2\alpha-1)^3}A_1^3\right), \quad \alpha > 3
\end{aligned}$$

and $A_i = \dfrac{\Gamma\left(\frac{k\tau+i}{k}\right)}{\Gamma(\tau)}$, $i = 1, 2, 3$.

Note that r_1 is complicated function of β, α, k and τ. For fixed $\alpha > 3$, $k > 0$ and $\tau > 0$, if $\beta > 0$ ($\beta < 0$) we have $0 < r_1 < \infty$ ($-\infty < r_1 < 0$) with $\lim_{\alpha\to 3^+} r_1 = \infty$ ($\lim_{\alpha\to 3^+} r_1 = -\infty$).

2.2 Maximum Likelihood Estimation via the *EM*-type Algorithm

In this section, we implement the maximum likelihood method to estimate the parameters of *LSMGG* distribution. We can see that the log-likelihood function of $LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$ is a very complicated function of the parameters and direct maximization of this function is very difficult. So, we use the *EM*-type algorithm (Dempster et al., 1977) to estimate these parameters. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$ be a random sample of size n from $LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$ distribution. Following the *EM* algorithm we consider Y_i as the observed data and V_i as the missing data. Therefore, (Y_i, V_i), $i = 1, 2, \cdots, n$ are the complete date. From (3), for $i = 1, 2, \ldots, n$ we can find the log-likelihood function for the complete data as

$$\begin{aligned} L_c(\boldsymbol{\theta}) &= n\ln(\alpha) + n\ln(k) - nk\tau\ln(\sigma) - n\ln(\Gamma(\tau)) \\ &\quad + \left(\alpha - 1 + \left(\frac{k\tau}{2}\right)\right) \times \sum_{i=1}^{n} \ln(v_i) + (k\tau - 1) \\ &\quad \times \sum_{i=1}^{n} \ln(y_i - \mu - v_i^{-1}\beta) - \frac{1}{\sigma^k} \\ &\quad \times \sum_{i=1}^{n} v_i^{\frac{k}{2}} (y_i - \mu - v_i^{-1}\beta)^k. \end{aligned}$$

For the simplicity in computation, we let $k = 2$ and μ and τ are known. Now the conditional expectation of $L_c(\boldsymbol{\theta})$ given the observed data y_i and the current estimates of the parameters, say $\widehat{\boldsymbol{\theta}}$, is given by

$$\begin{aligned} Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= E(L_c(\boldsymbol{\theta})|\boldsymbol{y}, \widehat{\boldsymbol{\theta}}) \\ &= c + n\ln(\alpha) - 2n\tau\ln(\sigma) - n\ln(\Gamma(\tau)) + (\alpha + \tau - 1) \end{aligned}$$

$$
\begin{aligned}
&\times \sum_{i=1}^{n} E(\ln(v_i)|y_i,\widehat{\boldsymbol{\theta}}) + (2\tau-1) \\
&\times \sum_{i=1}^{n} E(\ln(y_i-\mu-v_i^{-1}\beta)|y_i,\widehat{\boldsymbol{\theta}}) - \frac{1}{\sigma^2} \\
&\times \sum_{i=1}^{n} ((y_i-\mu)^2 E(v_i|y_i,\widehat{\boldsymbol{\theta}}) + \beta^2 E(v_i^{-1}|y_i,\widehat{\boldsymbol{\theta}}) - 2\beta(y_i-\mu)). \qquad (4)
\end{aligned}
$$

Letting

$$
\begin{aligned}
\widehat{k}_i &= E(v_i|y_i,\widehat{\boldsymbol{\theta}}), \\
\widehat{s}_{2i} &= E(\ln(v_i)|y_i,\widehat{\boldsymbol{\theta}}), \\
\widehat{s}_{3i} &= E(v_i^{-1}|y_i,\widehat{\boldsymbol{\theta}}), \\
\widehat{s}_{4i} &= E(\ln(y_i-\mu-v_i^{-1}\beta)|y_i,\widehat{\boldsymbol{\theta}}). \qquad (5)
\end{aligned}
$$

It is easy to show that

$$
\widehat{k}_i = \frac{\int_0^1 v_i^{\widehat{\alpha}+\tau}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i}{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i},
$$

$$
\widehat{s}_{2i} = \frac{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}\ln(v_i)(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i}{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i},
$$

$$
\widehat{s}_{3i} = \frac{\int_0^1 v_i^{\widehat{\alpha}+\tau-2}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i}{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i},
$$

$$
\widehat{s}_{4i} = \frac{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}\ln(y_i-\mu-v_i^{-1}\widehat{\beta})(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i}{\int_0^1 v_i^{\widehat{\alpha}+\tau-1}(y_i-\mu-v_i^{-1}\widehat{\beta})^{2\tau-1}\exp\left(-\frac{v_i}{\widehat{\sigma}^2}(y_i-\mu-v_i^{-1}\widehat{\beta})^2\right)dv_i}, \qquad (6)
$$

For E-step of the algorithm, the conditional expectations $\widehat{k}_i$, $\widehat{s}_{2i}$ and $\widehat{s}_{3i}$ must be evaluated. We note that for computing $\widehat{k}_i$, $\widehat{s}_{2i}$ and $\widehat{s}_{3i}$, Monte Carlo integration can be employed, which yields the so called $MC-EM$ algorithm. For the M-step of the algorithm, which consist of maximizing the expected complete data function over $\boldsymbol{\theta}$ or the Q-function, we have from (4) and (5)

$$\begin{aligned} Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) &= E(L_c(\boldsymbol{\theta})|\boldsymbol{y},\widehat{\boldsymbol{\theta}}^{(k)}) \\ &= c + n\ln(\alpha) - 2n\tau\ln(\sigma) - n\ln(\Gamma(\tau)) \\ &\quad +(\alpha+\tau-1)\sum_{i=1}^{n}\widehat{s}_{2i}^{(k)} + (2\tau-1)\sum_{i=1}^{n}\widehat{s}_{4i}^{(k)} \\ &\quad -\frac{1}{\sigma^2}\sum_{i=1}^{n}((y_i-\mu)^2\widehat{k}_i^{(k)} + \beta^2\widehat{s}_{3i}^{(k)} \\ &\quad -2\beta(y_i-\mu)), \end{aligned}$$

where $\widehat{\boldsymbol{\theta}}^{(k)}$ is an updated value of $\widehat{\boldsymbol{\theta}}$. When M-step of EM algorithm turns out to be analytically intractable, it can be replaced with a sequence of conditional maximization (CM) steps. Such a modification is referred to the ECM algorithm (Meng and Rubin, 1993). The $ECME$ algorithm, which is maximize the constrained Q-function with some CM-steps that maximize the corresponding constrained actual marginal likelihood function, called CML-steps (Liu and Rubin, 1994). Similar to Basso et al. (2010) and Farnoosh et al. (2013), we describe ECME algorithm as follows.

E-step. Given a current estimate $\widehat{\boldsymbol{\theta}}^{(k)} = (\widehat{\beta}^{(k)}, \widehat{\sigma}^{(k)}, \widehat{\alpha}^{(k)})$ and observation $\boldsymbol{y} = (y_1, \ldots, y_n)$, evaluate $\widehat{k}_i^{(k)}$, $\widehat{s}_{2i}^{(k)}$ and $\widehat{s}_{3i}^{(k)}$ from (6) by Monte Carlo integration.

CM-steps. Update $\widehat{\boldsymbol{\theta}}^{(k+1)}$ by maximizing $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$ over $\boldsymbol{\theta}$, which are given by the following closed from expressions

$$\widehat{\sigma}^{2(k+1)} = \frac{\sum_{i=1}^{n}\Big((y_i-\mu)^2\widehat{k}_i^{(k)} + (\widehat{\beta}^{(k)})^2\widehat{s}_{3i}^{(k)} - 2\widehat{\beta}^{(k)}(y_i-\mu)\Big)}{n\tau}$$

$$\widehat{\alpha}^{(k+1)} = -\frac{n}{\sum_{i=1}^{n}\widehat{s}_{2i}^{(k)}}.$$

CML-step. Update $\widehat{\beta}^{(k+1)}$ by maximizing the actual marginal log-likelihood function as follows

$$\widehat{\beta}^{(k+1)} = \text{argmax} \sum_{i=1}^{n} \log(f_{\boldsymbol{Y}}(y_i; \mu, \widehat{\sigma}^{(k)}, \beta, \hat{\alpha}^{(k)}, 2, \tau)),$$

where $f_Y(y_i; \boldsymbol{\theta})$ is the $LSMGG$ p. d. f.

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

2.3 The Observed Information Matrix

In order to find the standard error of the ML estimators of the parameters of $LSMGG$ distribution, we obtain the observed information matrix of this model, which will be defined by

$$\boldsymbol{J}_0(\boldsymbol{\theta}|\boldsymbol{y}) = -\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{y})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}.$$

Under some regularity conditions, the covariance matrix of the maximum likelihood estimates $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\sigma}, \widehat{\beta}, \widehat{\alpha}, \widehat{k}, \widehat{\tau})$ can be approximated by the inverse of $\boldsymbol{J}_0(\boldsymbol{\theta}|\boldsymbol{y})$. We evaluated the observed information matrix as follows,

$$\boldsymbol{J}_0(\widehat{\boldsymbol{\theta}}|\boldsymbol{y}) = \sum_{i=1}^{n} \widehat{t}_i \widehat{t}_i^T,$$

where

$$\widehat{t}_i = \frac{\partial(\log f(y_i; \theta_j))}{\partial\theta_j}, \qquad j = 1, 2, 3, 4, 5, 6.$$

We refer the reader to reference Basford et al. (1997) and Lin et al. (2007).

We note that vector $\widehat{\boldsymbol{t}}_i$ can be partitioned into components corresponding to all the parameters in $\boldsymbol{\theta}$ as

$$\widehat{\boldsymbol{t}}_i = (\widehat{t}_{i,\mu}, \widehat{t}_{i,\sigma}, \widehat{t}_{i,\beta}, \widehat{t}_{i,\alpha}, \widehat{t}_{i,k}, \widehat{t}_{i,\tau})^T,$$

where

$$\widehat{t}_{i,\theta_j} = \frac{\partial \log f(y_i; \boldsymbol{\theta})}{\partial\theta_j}, \qquad j = 1, 2, 3, 4, 5, 6.$$

We define

$$
\begin{aligned}
I_i(u_1,u_2,u_3) &= \int_0^1 v^{(\alpha-u_1)+\frac{k(\tau+u_2)}{2}}(y_i-\mu-v^{-1}\beta)^{k(\tau+u_2)-u_3} \\
&\quad\times \exp\left(-\frac{v^{\frac{k}{2}}}{\sigma^k}(y_i-\mu-v^{-1}\beta)^k\right)dv, \\
I_i^v(u_1,u_2,u_3) &= \int_0^1 v^{(\alpha-u_1)+\frac{k(\tau+u_2)}{2}}\log v(y_i-\mu-v^{-1}\beta)^{k(\tau+u_2)-u_3} \\
&\quad\times \exp\left(-\frac{v^{\frac{k}{2}}}{\sigma^k}(y_i-\mu-v^{-1}\beta)^k\right)dv,
\end{aligned}
$$

and

$$
\begin{aligned}
I_i^{v,y}(u_1,u_2,u_3) &= \int_0^1 v^{(\alpha-u_1)+\frac{k(\tau+u_2)}{2}}\log(y_i-\mu-v^{-1}\beta) \\
&\quad\times(y_i-\mu-v^{-1}\beta)^{k(\tau+u_2)-u_3} \\
&\quad\times \exp\left(-\frac{v^{\frac{k}{2}}}{\sigma^k}(y_i-\mu-v^{-1}\beta)^k\right)dv.
\end{aligned}
$$

After some algebraic calculation, we obtain

$$
\begin{aligned}
\frac{\partial}{\partial\mu}(f(y_i;\boldsymbol{\theta})) &= \frac{\alpha k}{\sigma^{k\tau}\Gamma(\tau)}\left(\frac{k}{\sigma^k}I_i(1,1,2)-(k\tau-1)I_i(1,0,2)\right), \\
\frac{\partial}{\partial\sigma}(f(y_i;\boldsymbol{\theta})) &= \frac{\alpha k^2}{\sigma^{k\tau+1}\Gamma(\tau)}\left(\frac{1}{\sigma^k}I_i(1,1,1)-\tau I_i(1,0,1)\right), \\
\frac{\partial}{\partial\beta}(f(y_i;\boldsymbol{\theta})) &= \frac{\alpha k}{\sigma^{k\tau}\Gamma(\tau)}\left(\frac{k}{\sigma^k}I_i(2,1,2)-(k\tau-1)I_i(2,0,2)\right), \\
\frac{\partial}{\partial\alpha}(f(y_i;\boldsymbol{\theta})) &= \frac{k}{\sigma^{k\tau}\Gamma(\tau)}(I_i(1,0,1)-\alpha I_i^v(1,0,1)), \\
\frac{\partial}{\partial k}(f(y_i;\boldsymbol{\theta})) &= \frac{\alpha}{\sigma^{k\tau}\Gamma(\tau)}\Bigg(I_i(1,0,1)-k\tau\log\sigma I_i(1,0,1) \\
&\quad+\frac{k\tau}{2}I_i^v(1,0,1)+k\tau I_i^{v,y}(1,0,1)-\frac{k}{2\sigma^k}I_i^v(1,1,1) \\
&\quad+\frac{k\log\sigma}{\sigma^k}I_i(1,1,1)-\frac{k}{\sigma^k}I_i^{v,y}(1,1,1)\Bigg),
\end{aligned}
$$

$$\begin{aligned}\frac{\partial}{\partial\tau}(f(y_i;\boldsymbol{\theta})) &= \frac{\alpha k}{\sigma^{k\tau}\Gamma(\tau)}\Bigg(\frac{k}{2}I_i^v(1,0,1)+kI_i^{v,y}(1,0,1)-k\log\sigma I_i(1,0,1)\\ &\quad -\frac{\partial}{\partial\tau}\log\Gamma(\tau)I_i(1,0,1)\Bigg).\end{aligned}$$

In the next section, we use the above method to estimate the parameters.

2.4 Diagnostic Analysis

We use the best known perturbation schemes for detecting the influence of observations that under small perturbation of the model exert great influence on the ML estimators, which are based on case deletion approach.

Cook (1977), Bolfarine et al. (2007) and Lin et al. (2009) have used the above method. we use the case-deletion approach to detect the influence of removing the ith case from the analysis by evaluating the matrix such as Cooks distance (see Cook, 1977).

Let $\widehat{\boldsymbol{\theta}}_{(i)}$ be ML estimate of $\boldsymbol{\theta}$ without the ith observation in the sample. To determine the influence of the ith case on the ML estimate $\widehat{\boldsymbol{\theta}}$, the basic idea is to compare the difference between $\widehat{\boldsymbol{\theta}}_{(i)}$ and $\widehat{\boldsymbol{\theta}}$. If $\widehat{\boldsymbol{\theta}}_{(i)}$ is far from $\widehat{\boldsymbol{\theta}}$, then the ith case is regarded as an influential observation. Based on this idea the global influence is defined as the standardized norm of $\widehat{\boldsymbol{\theta}}_{(i)}-\widehat{\boldsymbol{\theta}}$, namely the generalized Cook distance

$$GD_i(\boldsymbol{\theta})=(\widehat{\boldsymbol{\theta}}_{(i)}-\boldsymbol{\theta})^T[-\ddot{L}(\boldsymbol{\theta})](\widehat{\boldsymbol{\theta}}_{(i)}-\boldsymbol{\theta}),$$

where $-\ddot{L}(\boldsymbol{\theta})=\dfrac{\partial^2 L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^2}$ is the observed information matrix evaluated for $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}$ in the previous section.

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

3 Applications

In order to examine the performance of the LSMGG distribution, results from simulation studies and a real data set are presented.

Table 1. Estimates of the parameter by choosing 100 samples of sizes $n = 50, 100, 500$ from the $LSMGG$ distribution with $\beta = 1$, $\sigma = 1$, $\alpha = 5$ and $(\mu, \tau) = (0, 2), (-1, 3), (1, 6.5)$.

n	$\widehat{\beta}$	bias($\widehat{\beta}$)	$MSE(\widehat{\beta})$	$\widehat{\sigma}$	bias($\widehat{\sigma}$)	$MSE(\widehat{\sigma})$	$\widehat{\alpha}$	bias($\widehat{\alpha}$)	$MSE(\widehat{\alpha})$
				$(\mu, \tau) = (0, 2)$					
50	0.8938	-0.1061	0.0113	1.1021	0.1021	0.0115	5.4909	0.4909	1.4347
100	0.9378	-0.0622	0.0096	1.0700	0.0700	0.0050	5.4833	0.4833	0.3712
500	0.9617	-0.0382	0.0015	0.9983	-0.0016	0.0036	5.2612	0.2612	0.0692
				$(\mu, \tau) = (-1, 3)$					
50	-0.7862	-0.2138	0.0537	1.1424	0.1424	0.0726	4.0074	-0.9926	1.0254
100	0.8891	-0.1108	0.0489	1.1305	0.1305	0.0381	5.7416	0.7416	0.8143
500	0.9721	-0.0278	0.0008	1.0683	0.0683	0.0051	5.6664	0.6664	0.6567
				$(\mu, \tau) = (1, 6.5)$					
50	1.2567	0.2574	0.0675	0.8678	-0.1322	0.0186	4.7555	-0.2445	0.1647
100	1.0076	0.0076	0.0107	0.9634	-0.0365	0.0019	4.7434	-0.2566	0.1192
500	1.0035	0.0035	0.0050	1.0038	0.0038	0.0013	5.1653	0.1653	0.0275

Simulation Study

We perform a simulation study to show the precision of the proposed method of estimation. In this case, we first generate 100 samples of different sizes from $LSMGG$ distribution for $k = 2$ and fixed μ and τ parameters. We compute the ML estimates of the other parameters by the iteration method that illustrate in Subsection 2.2. Then we report the bias and Mean Square Error (MSE) of these estimates, which for β are defined by

$$\text{bias}(\widehat{\beta}) = \frac{1}{100}\sum_{i=1}^{100}\widehat{\beta}_{(i)} - \beta \text{ and } MSE(\widehat{\beta}) = \frac{1}{100}\sum_{i=1}^{100}(\widehat{\beta}_{(i)} - \beta)^2, \text{ respectively,}$$

where $\widehat{\beta} = \frac{1}{100}\sum_{i=1}^{100}\widehat{\beta}_{(i)}$ and $\widehat{\beta}_{(i)}$ is the $MC - EM$ estimate of β when the data is sample i. Definition for bias and MSE of the σ and α parameters are obtained similarly.

Table 1 presents the estimate of parameters and their bias and MSE for different values of n. From this table we see that the bias and MSE of $\widehat{\beta}$, $\widehat{\sigma}$ and $\widehat{\alpha}$ converges to zero when the sample size is increase.

Table 2. Maximum likelihood estimates and their standard errors for fitting $LSMGG(\mu, \sigma, \beta, \alpha, k, \tau)$, $GHSSE(\mu, \sigma, \beta, \alpha)$ and $GG(\gamma, \delta, k, \tau)$ distributions to the fiber glass data.

parameter	MLE of LSMGG	Se	MLE of GHSSE	Se	MLE of GG	Se
γ	-	-	-	-	-49.2660	0.0012
δ	-	-	-	-	50.6242	0.0012
μ	1.4310	0.1673	1.8825	0.0485	-	-
σ	0.4400	0.0938	0.2362	0.0205	-	-
β	-0.9200	0.1183	-0.2383	0.0361	-	-
α	3.2801	0.9237	2.6992	0.0051	-	-
k	2.1201	0.2837	-	-	130.8905	8.2354
τ	8.1200	2.4638	-	-	1.9369	0.0231
log f	-12.0338	-	-15.0802	-	-14.2398	-
AIC	36.0676	-	38.1604	-	36.4796	-
EDC	33.5923	-	36.5103	-	34.8294	-

The Fiber Glass Data Set

In this section we use the fiber glass data set, which is analyzed by Jones and Faddy (2003) and Azzalini and Capitanio (2003) and Wang and Genton (2006) in order to compare our proposed model with the GG model of Stacy (1962) and $GHSSE$ model of Arsalan (2008, 2009). We use the optim routin in R software to find the ML estimates of the parameters for GG and LSMGG distributions and we apply $MC - EM$ algorithm to estimate parameters of $GHSSE$ distribution. Table 2 contains estimated parameters of the above distributions with their corresponding standard errors (Se), that are computed via the information based method presented in subsection 2.3.

For each fitted model, we computed the log-likelihood values, the Akaike Information Criterion (AIC) (Akaike, 1974) and the Efficient Determination Criterion (EDC) (Bai et al., 1989). In addition, we display the fitted distributions in Figure 1. From these criteria and Figure 1, it appears that $LSMGG$ distribution fits better than the others to this data set.

Sensitivity Analysis

In this section we use the diagnostic analysis to detect the influence of observation on the ML estimates.

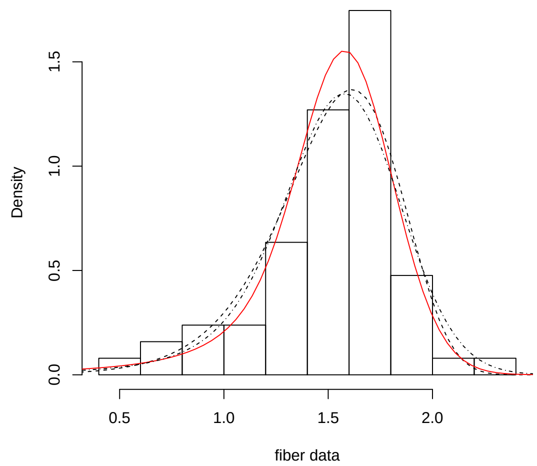

Figure 1. Histogram of fiber glass data set with fitted $LSMGG(\mu,\sigma,\beta,\alpha,k,\tau)$ distribution (solid line), $GHSSE(\mu,\sigma,\beta,\alpha)$ distribution (dashed-dotted line) and $GG(\gamma,\delta,k,\tau)$ distribution (dashed line).

To obtain the diagnostic measure, let $\widehat{\boldsymbol{\theta}}$ be the ML estimate of $\boldsymbol{\theta}$ in fiber glass data and $\widehat{\boldsymbol{\theta}}_i$ be the ML estimate of $\boldsymbol{\theta}$ without the i-th observation, then we compute the GD_i measures presented in subsection 2.4. Figure 2 depicts the index plot of GD_i for case weights perturbation. From Figure 2 we observe that the cases 1, 17, 18, 45, 63 are influential in parameters estimation indicating that the methodology works very well when suspicious points are present in the data set.

4 Conclusion

In this paper, we proposed a new asymmetric distribution, called LSMGG distribution. We investigate some properties of the presented model. Then, in order to find ML estimates of the parameters of distribution, we utilize EM-type algorithm. Furthermore, we illustrate our method with a real data set and show that the $LSMGG$ model has better performance than the other competitors' ones.

Finally, we use the local influence approach to detect influence of observations on ML estimators. Results are obtained from the real data set show

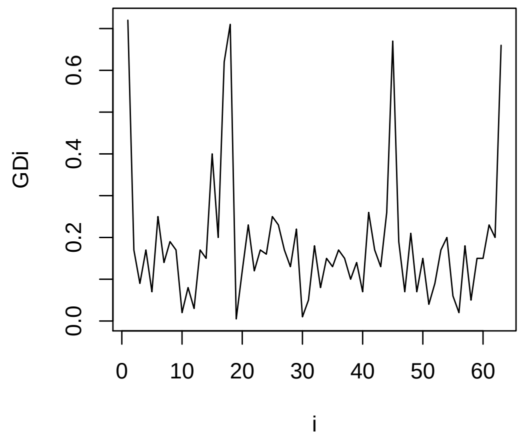

Figure 2. Index plot of GD_i for case weights perturbation for fiber glass data.

the methodology works very well when suspicious points are present in the data set.

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

Azzalini, A. (1985). A Class of Distributions which Includes the Normal Ones, *Scand. J. Statist.*, **12**, 171-178.

Azzalini, A. and Capitanio, A. (2003). Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t Distribution, *J. Roy. Statist. Soc.*, B **65**, 367-389.

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.

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 EMAalgorithm , *J. Roy. Statist. Soc.*, B **39**, 1-38.

Farnoosh, R., Nematollahi, N., Rahnamaei, Z. and Hajrajabi, A. (2013). A Family of Skew-slash Distributions and Estimation of its Parameters via an EM Algorithm, *JIRSS*, **12(2)**, 271-291.

Jones, M.C. and Faddy, M.J. (2003). A Skew Extension of the t-Distribution with Applications, *J. Roy. Statist. Sec. Ser.*, B **65**, 159-174.

Khodabin, M. and Ahmadabadi, A. (2010). Some Properties of Generalized Gamma Distribution, *Mathematical Sciences*, **4(1)**, 9-28.

Lin, T.L., Lee, J.C. and Yen, S.Y. (2007). Finite Mixture Modeling using the Skew Normal Distribution, *Statistica. Sinica*, B **17**, 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.

Stacy, E.W. (1962). A Generalization of the Gamma Distribution, *The Annals of Mathematical Statistics*, **33**, 1187-1192.

Wang, J. and Genton, M.G. (2006). The Multivariate Skew-slash Distribution, *J. Statist. Plann. Infer.*, **136**, 209-220.

Z. Rahnamaei
Department of Mathematics,
Firoozkooh Branch, Islamic Azad University,
Firoozkooh, Iran.
email: *rahnamaei@iaufb.ac.ir*