Statistical Methodology 32 (2016) 131–146

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Location-scale mixture of skew-elliptical distributions: Looking at the robust modeling

N. Nematollahi [a,*], R. Farnoosh [b], Z. Rahnamaei [c]

[a] Department of Statistics, Allameh Tabataba'i University, Tehran, Iran

[b] School of Mathematics, Iran university of Science and Technology, Narmak, Tehran, Iran

[c] Department of Mathematics, Firoozkooh Branch, Islamic Azad University, Firoozkooh, Iran

ARTICLE INFO

Article history:
Received 3 October 2014
Received in revised form
2 May 2016
Accepted 3 May 2016
Available online 2 June 2016

MSC:
60E05
62F10

Keywords:
Heavy tailed distributions
Location-scale mixture distributions
Skew elliptical distributions
Slash distribution

ABSTRACT

A flexible class of skew-slash distributions which is a location-scale mixture of skew-elliptically distributed random variable with power of a beta random variable is presented. This family of distributions, which is a generalization of location-scale mixture of normal and beta distributions, contain some existing and important distributions and is appropriate for modeling data with skewness and heavy tail structure. Some distributional properties and the moments of this new family of distributions are obtained. In the special case of location-scale mixture of skew-normal distribution, we estimate the parameters via an EM-type algorithm and a simulation study and an application to real data are provided for illustration. Finally we extend some results to multivariate case.

1. Introduction

Normal distribution is one of the classical and most important distributions for analyzing a real data set. Most of the statistical analysis are developed based on normal distribution, because of its analytical tractability and the central limit theorem effect. But, when the data set is not symmetric and/or has heavy tail, the normal distribution does not fit well to it. In the past three decades, researchers construct some family of distributions which are skew and/or have heavy tail for analyzing such data

* Corresponding author.
E-mail address: nematollahi@atu.ac.ir (N. Nematollahi).

http://dx.doi.org/10.1016/j.stamet.2016.05.001
1572-3127/

sets. Most of these family of distributions have some properties of normal distribution and contain normal distribution as a proper member. See for example [15] and references cited therein.

Azzalini [4,5] defined the skew-normal distribution($SN_1(\mu, \sigma, \lambda)$) which have the probability density function (p.d.f.)

$$f_X(x) = \frac{2}{\sigma} g\left(\frac{x-\mu}{\sigma}\right) G\left(\lambda \frac{x-\mu}{\sigma}\right), \tag{1}$$

where $g(.)$ and $G(.)$ are the p.d.f. and cumulative distribution function (c.d.f.) of standard normal distribution, respectively. It is a skew distribution with skew parameter λ and contain normal distribution as a proper member. Some authors extended and applied p.d.f. (1) to construct other distributions. Among them Azzalini and Dalla Valle [7] extend it to the multivariate skew-normal distribution and Branco and Dey [11] and Sahu et al. [28] extend it to the skew-elliptical distribution in which $g(.)$ and $G(.)$ in (1) are the p.d.f. and c.d.f. of an elliptical distribution, respectively.

One of the heavy-tailed distribution which is alternative to the normal distributions for modeling data sets with heavier tails than normal is slash distribution. The random variable

$$X = \mu + \sigma \frac{Z}{U^{\frac{1}{q}}} \tag{2}$$

is said to have slash distribution with parameters 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. It is a symmetric distribution which have heavier tails than normal distribution when q is small, and close to normal distribution when $q \to \infty$. Rogers and Tukey [27] and Mosteller and Tukey [25] are studied general properties of this distribution. Note that (2) is a scale mixture of normal and uniform random variables. Some researchers by changing the distribution of Z in (2), extended this family of distributions. Among them, Wang and Genton [29] and Gomez et al. [17] extended it to the multivariate Slash (SL) and the Slash-Elliptical (SLEL) distributions by replacing Z in (2) with the multivariate normal and elliptical distributions, respectively. Wang and Genton [29] and Farnoosh et al. [14] extended it to the Slash Skew-Normal (SLSN) and Slash Skew-Elliptical (SLSEL) distributions, by replacing Z in (2) with the multivariate skew-normal and skew-elliptical distributions, respectively, which are not symmetric and have heavier tail than normal distribution.

Some authors by location-scale mixture of a no-negative random variable with normal random variable, construct flexible distributions that are heavier tail than normal and have skewness (see, e.g., [24]). Arslan [2,3] considered the following location-scale mixture random vector

$$\mathbf{Y} = \mu + V^{-1}\beta + V^{-\frac{1}{2}} \Sigma^{\frac{1}{2}} \mathbf{X}, \tag{3}$$

where $\mu, \beta \in \Re^p$, Σ is a positive definite matrix, V have a $beta(\alpha, 1)$ distribution and $\mathbf{X}$ have a standard multivariate elliptical distribution. When X have $N_p(\mathbf{0}, I_p)$ distribution, he called (3) as the Generalized Hyperbolic Skew-Slash (GHSS) distribution, denoted by $\mathbf{Y} \sim GHSS_p(\mu, \Sigma, \beta, \alpha)$, and derived some properties of this distribution and estimate its parameters by an EM algorithm.

In this paper, by replacing $\mathbf{X}$ in (3) with skew-elliptical distributions, we define a new family of location-scale mixture distributions in univariate and multivariate cases. These distributions are appropriate for fitting skewed and heavy-tailed data sets, and contain SL, GHSS, SLSN, SLEL and some other existing distributions. To this end, in Section 2 we state some preliminary definitions and results. In Section 3, in univariate case, the Location-Scale Mixture of Skew-Elliptical (LSMSEL) distribution is introduced and some properties are given. In the special case of location-scale mixture of skew-normal distribution, an EM-type algorithm is constructed to estimate its parameters and a simulation study and an application to real data set are given for illustration. In Section 4, we define multivariate LSMSEL distribution and give some distributional properties.

2. Preliminaries

In this section some definitions and results that are needed for subsequent sections are given.