Journal of the Korean Statistical Society 39 (2010) 545–554

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

On a class of symmetric distributions associated with the product of two correlated t variables☆

Hea-Jung Kim

Department of Statistics, Dongguk University, 3 Pil-Dong, Chung-Gu, Seoul 100-715, Republic of Korea

ARTICLE INFO

Article history:
Received 27 May 2009
Accepted 7 January 2010
Available online 25 January 2010

AMS 2000 subject classifications:
primary 62E10
secondary 62F30

Keywords:
Symmetric distribution
Nonlinearly truncated distribution
Product of t variables
Scale mixture
Selection mechanism
Bayesian estimation

ABSTRACT

This article introduces a class of symmetric non-Gaussian distributions associated with the product of two correlated t variables. The distributions are obtained from the conditioning method. They are symmetric, possibly bimodal, and the marginal distributions of a nonlinearly truncated bivariate Student t_ν-distribution. The class is studied from several aspects, such as association with other distributions, nonlinear selection mechanism, and scale mixture scheme. The relationships among these aspects are given, and various properties of the class are also considered. Bayesian estimation of the class is discussed and three applications utilizing the class of distributions are also provided.

1. Introduction

For analyzing non-normal data sets, construction of flexible non-Gaussian distributions has seen a growing interest in recent years. Several classes of distributions have been put forward in the literature, an overview of which can be found in the book edited by Genton (2004), in Kim (2008), and from a unified point of view in Arellano-Valle, Branco, and Genton (2006).

As a class of such distributions, we may define a new class consisting of symmetric distributions called two-piece-t distributions. Specifically, we say that a random variable T has a two-piece-t distribution with ν degrees of freedom, and location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$, and shape parameter $\lambda \in \mathbb{R}$, if its probability density function (pdf) is given by

$$h_T(t;\nu,\mu,\sigma,\lambda) = \frac{c_\lambda}{\sigma} f_\nu(z) F_{\nu+1}\left(\lambda|z|\sqrt{(\nu+1)/(\nu+z^2)}\right), \quad t \in \mathbb{R}, \tag{1.1}$$

where $z = (t-\mu)/\sigma$, $c_\lambda = (1/2 + \pi^{-1}\tan^{-1}\lambda)^{-1}$, and $f_\nu(z)$ and $F_\nu(z)$ are, respectively, the pdf and cdf of the Student $t_\nu \equiv t_\nu(0,1)$ distribution. In such case, we write $T \sim TPt_\nu(\mu,\sigma,\lambda)$, and we note that $T \overset{d}{=} \mu + \sigma Z$, where $Z \sim TPt_\nu(0,1,\lambda) \equiv TPt_\nu(\lambda)$. The distribution arises from a non-linear selection scheme associated with the product of two correlated t variables. Although cases of non-linear selection scheme to obtain new symmetric/asymmetric distributions are generally addressed in Arellano-Valle et al. (2006), so far as we know, there are few results that fully explore the distributions obtained from the selection scheme. A particular case of (1.1) with $\nu \to \infty$ has been studied by Kim (2005).

☆ This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2009-0071000).

E-mail address: kim3hj@dongguk.edu.

1226-3192/$ – see front matter
doi:10.1016/j.jkss.2010.01.001

The interest in studying the class of $TPt_\nu(\mu, \sigma, \lambda)$ distributions comes from both theoretical and applied directions. On the theoretical side, it involves a random constraint that allows for non-linear hidden truncations, as discussed in Arellano-Valle et al. (2006). Furthermore, it defines yet another class of uni/bi-modal symmetric distributions whose properties are mathematically tractable and flexible enough to include the Student t_ν-distributions. From the applied point of view, (1.1) is a symmetric distribution with the presence of a heavy tail and possibly bimodal. Thus the class can be used to obtain many more families of new skewed distributions by using the skewing function approach given in Wang, Boyer, and Genton (2004). In particular, the class is useful for modelling various types of skew-bimodal distributions. Moreover, the class arises from a selection mechanism where underlying distribution is a non-linearly truncated bivariate Student-t distribution. Many screening and quality control problems arise in practice which can be modelled using such selection distribution. Such a model can also be used for an exact Bayesian inference for linear models with a nonlinear constraint.

This paper is organized as follows. Section 2 proposes important probabilistic characteristics, including a stochastic representation, of the $TPt_\nu(\mu, \sigma, \lambda)$ distribution. Then some distributional properties and moment calculations for the distribution are presented. Section 3 considers a Bayesian estimation of the $TPt_\nu(\mu, \sigma, \lambda)$ model. We provide an expression for the full conditional posterior distributions of the parameters of interest using a data augmentation technique. Then Markov chain Monte Carlo strategies for the estimation are given and examined by using a simulation study. Section 4 provides three applications that demonstrate the utility of the class. Finally, Section 5 concludes the paper with brief concluding remarks.

2. The class of symmetric distributions

2.1. Derivation of the class and its characteristics

Suppose Z_1 and Z_2 are two independent $N(0, 1/\eta)$ variables. Conditionally on $\eta \sim \text{Gamma}(\nu/2, \nu/2)$, the pdf of Z_1 given that Z_1 satisfies a random constraint of the form $Z_2 < \lambda|Z_1|$ is

$$f_{[Z_1|Z_2<\lambda|Z_1|,\eta]}(z) = \frac{f_{Z_1|\eta}(z)P(Z_2 \le \lambda|z| \mid Z_1 = z, \eta)}{P(Z_2 < \lambda|Z_1| \mid \eta)} = \frac{\eta^{1/2}\phi(\eta^{1/2}z)\Phi(\eta^{1/2}\lambda|z|)}{P(Z_2 < \lambda|Z_1|)}, \tag{2.1}$$

where $P(Z_2 < \lambda|Z_1| \mid \eta) = P(Z_2 < \lambda|Z_1|)$ for all η. The scale mixture of (2.1) is given by

$$h_Z(z; \nu, \lambda) = \int_0^\infty f_{[Z_1|Z_2<\lambda|Z_1|,\eta]}(z)\mathrm{d}G(\eta) = c_\lambda f_\nu(z)F_{\nu+1}\left(\lambda|z|\sqrt{(\nu+1)/(\nu+z^2)}\right), \tag{2.2}$$

where $G(\eta)$ is the cdf of $\eta \sim \text{Gamma}(\nu/2, \nu/2)$, $c_\lambda^{-1} = \int_0^\infty E[\Phi(\eta^{1/2}\lambda|Z_2|)]\mathrm{d}G(\eta) = 1/2 + \pi^{-1}\tan^{-1}\lambda$ (see, Kim (2005), for the evaluation of c_λ). This gives the following theorem.

Theorem 2.1. *Let $Z \stackrel{d}{=} [T_1 \mid T_2 < \lambda|T_1|]$ and $U \stackrel{d}{=} [T_1 \mid T_2 > \lambda|T_1|]$, where $(T_1, T_2)^\top \sim t_\nu(\mathbf{0}, I_2)$. Then $Z \sim TPt_\nu(\lambda)$ and $U \sim TPt_\nu(-\lambda)$, where $t_\nu(\mathbf{0}, I_2)$ denotes a bivariate t_ν-distribution with location parameter $\mathbf{0}$, scale matrix I_2, and degrees of freedom ν.*

From now on, we shall use $t_\nu(\mathbf{0}, \Omega)$ to denote a bivariate t_ν-distribution with location parameter $\mathbf{0}$, scale matrix Ω, and degrees of freedom ν, where

$$\Omega = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{2.3}$$

The following theorem shows that the $PTt_\nu(\lambda)$ distribution defined in (1.1) is associated with the product of two correlated t_ν variables, and it involves a random constraint that allows for non-linear hidden truncations, as discussed in Arellano-Valle et al. (2006).

Theorem 2.2. *Let $(X, Y)^\top \sim t_\nu(\mathbf{0}, \Omega)$, and consider the selection random variables $Z \stackrel{d}{=} [X|XY > 0]$ and $U \stackrel{d}{=} [X|XY < 0]$. Then $Z \sim TPt_\nu(\lambda)$ and $U \sim TPt_\nu(-\lambda)$, where $\lambda = \rho/\sqrt{1-\rho^2}$.*

Proof. We see that

$$f_{[X|XY>0]}(z) = f_\nu(z)P(XY > 0 \mid X = z)/P(XY > 0). \tag{2.4}$$

Here $[XY|X = z] \sim t_{\nu+1}(\rho z^2, (1-\rho)^2(\nu+z^2)z^2/(\nu+1))$, a $t_{\nu+1}$-distribution with location $\mu(z) = \rho z^2$ and scale parameter $\sigma^2(z) = (1-\rho^2)(\nu+z^2)z^2/(\nu+1)$, so that

$$P(XY > 0|X = z) = F_{\nu+1}(\lambda|z|\sqrt{(\nu+1)/(\nu+z^2)}).$$

By symmetry, $P(XY > 0) = 2P(X < 0, Y < 0) = 2\Phi(0, 0; \rho)$ and it is known that $2\Phi(0, 0; \rho) = 1/2 + \pi^{-1}\tan^{-1}\lambda$, where $\Phi(y_1, y_2; \rho)$ denotes the joint cumulative distribution function of $(Y_1, Y_2)^\top \sim N_2(\mathbf{0}, \Omega)$, a standard bivariate normal distribution (see, for example Johnson & Kotz, 1972). Thus (2.4) is equivalent to the pdf of $TPt_\nu(\lambda)$ in (2.2). Similar proof holds for the distribution of U. □