The Korean Communications
in Statistics Vol. 8, No. 3, 2001
pp. 867-873

On a Skew-t Distribution

Hea-Jung Kim[1)]

Abstract

In this paper we propose a family of skew- t distributions. The family is derived by a scale mixtures of skew-normal distributions introduced by Azzalini (1985) and Henze (1986). The salient features of the family are mathematical tractability and strict inclusion of the normal law. Further it includes a shape parameter, to some extent, controls the index of skewness. Necessary theory involved in deriving the family of distributions is provided and main properties of the family are also studied.

Keywords : Skew-t distribution, Moments, Index of Skewness

1. Introduction

Azzalini (1985) and Henze (1986) worked on the so called skew-normal distribution, a family of distributions including the standard normal, but with an extra parameter to regulate skewness. A random variable Z is said to be skew-normal with parameter θ, written $Z \sim SN(\theta)$, if its density function is

$$\phi(z;\theta) = 2\phi(z)\Phi(\theta z), \quad -\infty < z < \infty, \tag{1}$$

where $\phi(z)$ and $\Phi(z)$ denote the $N(0,1)$ density and distribution function, respectively; the parameter θ which regulates the skewness varies in $(-\infty, \infty)$, and $\theta=0$ corresponds to the $N(0,1)$ density. We refer Arnold et al. (1993), Azzalini and Valle (1996) and Chen, Dey and Shao (1999), Kim (2001) for the applications of the distribution.

The purpose of the present paper is to introduce yet another family of distributions that includes the skew-normal as a special case. We consider using a scale mixtures of skew-normal densities to coming at a family of skew- t distributions. So that this gives rich family of parametric density functions that allow a continuous variation from normality to non-normality. Such an extension is potentially relevant for practical applications, since in data

1) Professor, Department of Statistics, Dongguk University, Seoul, Korea
E-mail : kim3hj@mail.dgu.ac.kr

analysis there are a few parametric distributions available to dealing with both symmetric and skewed data, especially for the problem of fitting heavy-tailed and skewed data. Necessary theory involved in deriving the family of distribution is provided.

2. The Family of Distributions

This section proposes a family of skew- t distributions by use of a scale mixtures of the skew-normal distributions in (1).

Lemma 1. Let $f(\cdot\,|\lambda)$ be a density function symmetric about 0, $G(\cdot\,|\lambda)$ an absolutely continuous distribution function such that $G'(\cdot\,|\lambda)$ is symmetric about 0, and λ is a random variable with density function $\mu(\lambda)$. Then

$$2E_{\lambda}[G(\theta y\,|\lambda)f(y\,|\lambda)], \quad -\infty < y < \infty \tag{2}$$

is a density function for any real θ.

Proof. Let $Y|\lambda$ and $W|\lambda$ be independent random variables with density f and G', respectively. Then

$$\begin{aligned} 1/2 &= E_{\lambda}[P(W-\theta Y < 0\,|\lambda)] = E_{\lambda}E_{Y|\lambda}[P(W < \theta y\,|Y=y,\lambda)] \\ &= \int_{\lambda\in\Lambda}\int_{-\infty}^{\infty} G(\theta y\,|\lambda)f(y\,|\lambda)\mu(\lambda)dyd\lambda = \int_{-\infty}^{\infty} E_{\lambda}[G(\theta y|\lambda)f(y\,|\lambda)]dy, \end{aligned}$$

where Λ is the space of λ.

Using Lemma 1, we can define the family of skew- t densities.

Definition 1. A random variable Z is a skew- t random variable with parameter θ and ν, written $Z \sim St(\theta,\nu)$, if its probability density function is

$$h(z;\theta,\nu) = 2E_{\lambda}[\lambda^{1/2}\phi(\lambda^{1/2}z)\Phi(\lambda^{1/2}\theta z)], \quad -\infty < z < \infty, \tag{3}$$

where $\lambda \sim$ Gamma $(\nu/2, 2/\nu)$ with $E\,\lambda = 1$, and ϕ and Φ are the standard normal density and distribution function, respectively.

Notice that the family $\{h_{\theta,\nu}, \theta\in\Theta, \nu>0\}$ defined by (3) denotes a family of scale mixtures of skew-normal densities, and hence leads to skewed and heavy-tailed distributions. The family can be represented in terms of scale mixture of normal and truncated normal laws.

Theorem 1. Conditional on $\lambda \sim$ Gamma $(\nu/2, 2/\nu)$, let $U \sim N(0, \lambda^{-1})$ and $V \sim N(0, \lambda^{-1})$ be independent normal variables, and let

$$Z = \frac{\theta}{(1+\theta^2)^{1/2}} |U| + \frac{1}{(1+\theta^2)^{1/2}} V. \tag{4}$$

Then the unconditional distribution of Z is $St(\theta, \nu)$.

Proof. For a fixed λ, the conditional density of $Z|\lambda$ is straightforward from the convolution formula. So that it has the density

$$2\lambda^{1/2}\phi(\lambda^{1/2}z)\Phi(\lambda^{1/2}\theta z), \quad -\infty < z < \infty, \quad \lambda > 0.$$

Taking expectation with respect to λ, we obtain the result.

Corollary 1. If $Y|\lambda$ and $W|\lambda$ are independent $N(0, \lambda^{-1})$ variables for some $\lambda \sim$ Gamma $(\nu/2, 2/\nu)$, and Z is set to the scale mixed random variable Y conditionally on $\theta Y > W$. Then the scale mixed distribution of Z is $St(\theta, \nu)$.

Proof. Upon performing the transformations $R|\lambda = (1+\theta^2)^{-1/2}(W|\lambda - \theta Y|\lambda)$ and $S|\lambda = (1+\theta^2)^{-1/2}(Y|\lambda + \theta W|\lambda)$, $R|\lambda$ and $S|\lambda$ are independent $N(0, \lambda^{-1})$ random variables, and $\theta Y|\lambda > W|\lambda$ is equivalent to $R|\lambda < 0$. Since $Y|\lambda = (1+\theta^2)^{-1/2}(-\theta R|\lambda + S|\lambda)$ the conditional distribution of $-R|\lambda$ given that $R|\lambda < 0$ equals that of $|U|\,|\lambda$ in (4), and hence the distribution of $Y|\lambda$ given that $R|\lambda < 0$ is equivalent to that of $Z|\lambda$ in (4). Thus the mixed distribution of Z gives the result.

Corollary 1 implies that the random variable Z with density (2) can be generated by the following acceptance-rejection technique. Sample λ from Gamma$(\nu/2, 2/\nu)$, and then independently sample W and Y from $N(0, \lambda^{-1})$. If $W < \theta Y$, then put $Z = Y$, otherwise restart sampling a new λ and pair of variables W and Y until the inequality is satisfied.

The next result follows immediately from Corollary 1 on setting $X|\lambda = (\theta Y|\lambda - W|\lambda)/$

$(1+\theta^2)^{1/2}$.

Corollary 2. If $(X|\lambda, Y|\lambda)$ is a bivariate normal variable with $N(0, \lambda^{-1})$ marginals and correlation ρ for $\lambda \sim \text{Gamma}(\nu/2,\ 2/\nu)$, then the scale mixed distribution of Y given $X > 0$ is $St(\theta(\rho), \nu)$, where $\theta(\rho) = \rho/(1-\rho^2)^{1/2}$.

The following properties follow immediately for the Theorem 1.

Property 1. If $\Pr(\lambda = 1) = 1$, $St\ (\theta,\ \nu)$ and $St\ (0,\ \nu)$ densities are the skew-normal density by Azzalini(1985) and $N(0,1)$ density, respectively.

Property 2. As $\theta \to \infty$, distribution of $St\ (\theta,\ \nu)$ tends to the half- t_ν. On the other hand $St\ (0,\ \nu)$ is the same as t_ν distribution.

Property 3. If Z is a $St\ (\theta,\ \nu)$ random variable, then $-Z$ is a $St\ (-\theta,\ \nu)$ random variable.

3. Moments

Since, conditional on λ, $Z^2 \sim \lambda^{-1}\chi^2_{(1)}$, for $\nu > 2k+2$, the even moments of Z are equal to

$$EZ^{2k+2} = 1\cdot 3\cdots(1+2k)\left(\frac{\nu}{2}\right)^{k+1}\Gamma\left(\frac{\nu-2}{2}-k\right)/\Gamma\left(\frac{\nu}{2}\right), \quad \text{for } k = 0,1,2,\cdots. \tag{5}$$

For computing the odd moments, we make use of Corollary 4 of Henze (1986). This gives

$$EZ^{2k+1} = \left(\frac{2}{\pi}\right)^{1/2}\left(\frac{\nu}{2}\right)^{k+1/2}\Gamma\left(\frac{\nu-1}{2}-k\right)/\Gamma\left(\frac{\nu}{2}\right)\frac{\theta}{(1+\theta^2)^{k+1/2}}\ \frac{(2k+1)!}{2^k}\sum_{j=0}^{k}\frac{j!\,(2\theta)^{2j}}{(2j+1)!\,(k-j)!}$$

for $k = 0,1,2,\cdots$ and $\nu > 2k+1$. (6)

Lemma 2 The moment generating function of Z is

$$M_Z(t) = 2E_\lambda\left[\exp\{\lambda^{-1}t^2/2\}\,\Phi\left(\frac{\theta\lambda^{-1/2}t}{(1+\theta^2)^{1/2}}\right)\right]. \tag{7}$$

Proof. From Definition 1, we see that the moment generating function $M_Z(t)$ is

$$2E_\lambda\int_{-\infty}^{\infty}\exp\left\{\frac{t^2}{2\lambda}\right\}\left(\frac{\lambda}{2\pi}\right)^{1/2}\exp\left\{-\frac{\lambda(z-t/\lambda)^2}{2}\right\}\Phi\{\theta z\lambda^{1/2}\}\,dz$$

$$= 2E_\lambda \int_{-\infty}^{\infty} \exp\left\{\frac{t^2}{2\lambda}\right\}\left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{u^2}{2}\right\} \Phi\{\theta(u + \lambda^{-1/2}t)\}du\ .$$

Using the well known result that, for $U \sim N(0,1)$, $E\{\Phi(hU + k)\} = \Phi\{k/(1+h^2)^{1/2}\}$ for any real h and k (see Zacks 1981, 99. 53-54), we have the result.

Hence, after some algebra, we obtain the mean and variance of $Z \sim St(\theta, \nu)$:

$$E[Z] = \left(\frac{\nu}{\pi}\right)^{1/2} \frac{\theta}{(1+\theta^2)^{1/2}} \Gamma\left(\frac{\nu-1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right), \quad \nu > 1$$

and

$$\sigma_Z^2 = \frac{\nu}{\nu-2} - \frac{\theta^2\nu}{\pi(1+\theta^2)}\left\{\Gamma\left(\frac{\nu-1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)\right\}^2, \quad \nu > 2.$$

The skewness of the distribution Z can be obtained by use of (4).

Theorem 2. Let $\sigma^2_{|U|}$ and σ^2_V be the variances of $|U|$ and V, and let $\mu^3_{|U|}$ be the standardized third moment of $|U|$; that is, $\mu^3_{|U|} = E\{[|U| - E(|U|)]/\sigma_{|U|}\}^3$. Then, the standardized third moment μ^3_Z of Z with the pdf (3) is given by

$$\mu^3_Z = E\left(\frac{Z - E(Z)}{\sigma_Z}\right)^3 = \frac{\theta(\theta^2\sigma^3_{|U|}\mu^3_{|U|} + K(\nu))}{\sigma^3}, \quad \nu > 3, \tag{8}$$

where $\sigma^2 = Var((1+\theta^2)^{1/2}Z) = (1+\theta^2)\sigma_z^2$.

Proof. Let $Z^* = (1+\theta^2)^{1/2}Z$, then the standardized third moment μ^3_Z is equivalent to that Z^*, i.e. $\mu^3_Z = \mu^3_{Z^*}$. From Theorem 1, we see that $E[Z^* - E(Z^*)]^3 = E_\lambda\{E[(\theta|U| - \theta E(|U|) + V)^3|\lambda]\} = \theta^3 E[(|U| - E(|U|))^3] + 3\theta E_\lambda E\{[(|U| - E|U|)V^2|\lambda]\} = \theta^3\sigma^3_{|U|}\mu^3_{|U|} + \theta K(\nu)$. Thus $\mu^3_Z = \mu^3_{Z^*} = (\theta^3\sigma^3_{|U|}\mu^3_{|U|} + \theta K(\nu))/\sigma^3_{Z^*}$, where

$$K(\nu) = 3\left(\frac{2}{\pi}\right)^{1/2}(E\lambda^{-3/2} - E\lambda^{-1}E\lambda^{-1/2})$$
$$= 3\left(\frac{2}{\pi}\right)^{1/2}\left(\frac{\nu}{2}\right)^{3/2}\Gamma\left(\frac{\nu-3}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)\left\{1 - \frac{\nu-3}{\nu-2}\right\}.$$

and $\sigma^2_{Z^*} = (1 + \theta^2)\sigma^2_Z$.

Note that $\mu^3_{|U|} > 0$ (i. e. the distribution of $|U|$ is skewed to the right as defined in (4)), and $K(\nu) > 0$. Consequently the skewness of the distribution of Z is characterized by θ. (8) implies that the distribution of Z, $St(\theta, \nu)$, is skewed to the right (the left) when $\theta > 0$ ($\theta < 0$) and symmetric for $\theta = 0$. Figure 1 and Figure 2 show the shape of the distribution for various values of θ and ν.

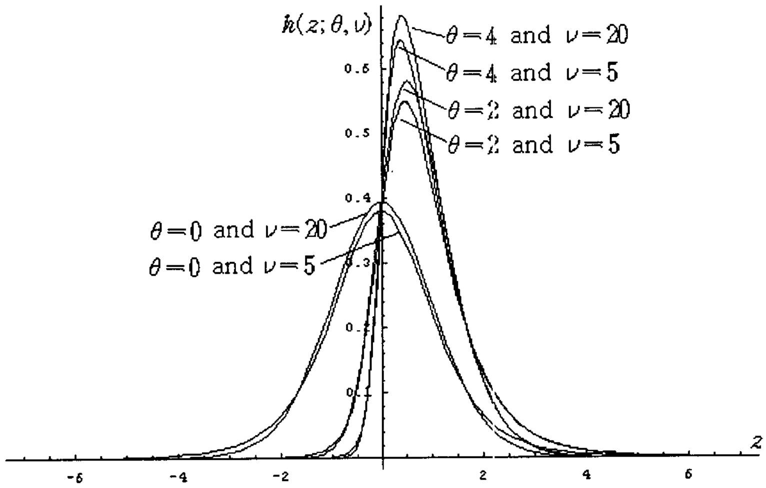

Figure 1. Shape of the Probability Density Function of $Z \sim St\,(\theta,\ \nu)$.

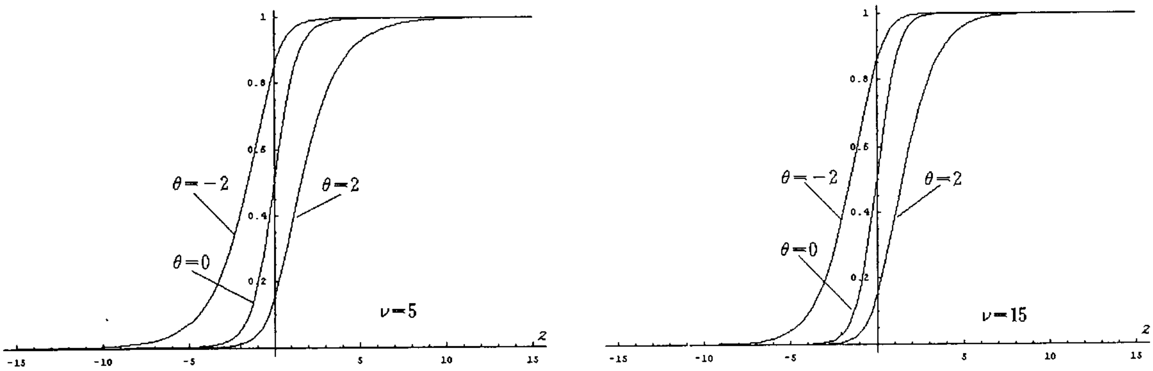

Figure 2. Shape of the Distribution Function of $Z \sim St\,(\theta,\ \nu)$.

4. Concluding Remark

This paper has proposed a family of skew $-t$ distributions, denoted by $St\,(\theta,\ \nu)$, the parameter θ which regulates the skewness. The special feature of the family is that it gives rich family of parametric density functions that allow a continuous variation from normality to non-normality. Therefore the family of skew $-t$ distributions is potentially relevant for practical applications, especially for the analysis of skewed data. Immediate applications of the distribution can be illustrated as follows: (i) Binary regression with an asymmetric link function; (ii) Regression analysis with asymmetric errors; (iii) Regression analysis with truncated errors in independent variables.

A study pertaining to the applications is an interesting research topic and it is left as a future study of interest.

References

[1] Arnold, B. C., Beaver, R. J., Groeneveld, R. A. and Meeker, W. Q. (1993). The nontruncated marginal of a truncated bivariate normal distribution. *Psychometrica*, 58, 471-478.

[2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.

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

[4] Chen, M. H., Dey, D. K., and Shao, Q. M. (1999). A new skewed link model for dichotomous quantal response model. *Journal of the American Statistical Association*, 94, 1172-1186.

[5] Henze, N. (1986). A probabilistic representation of the 'Skewed-normal' distribution. *Scandinavian Journal of Statistics*, 13, 271-275.

[6] Kim, H. J. (2001). Bayesian analysis of a new skewed multivariate probit model for correlated binary response data. *Journal of the Korean Statistical Society*, 30, to appear.

[7] Zacks, S. (1981). *Parametric Statistical Inference*, Pregamon Press, Oxford.