

Research Article

A Mixture of Generalized Tukey's g Distributions

José Alfredo Jiménez and Viswanathan Arunachalam

Department of Statistics, Universidad Nacional de Colombia, Carrera 45 No. 26-85, Bogotá, Colombia

Correspondence should be addressed to Viswanathan Arunachalam; varunachalam@unal.edu.co

Received 8 April 2016; Revised 1 July 2016; Accepted 14 July 2016

Academic Editor: Chin-Shang Li

Copyright © 2016 J. A. Jiménez and V. Arunachalam. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mixtures of symmetric distributions, in particular normal mixtures as a tool in statistical modeling, have been widely studied. In recent years, mixtures of asymmetric distributions have emerged as a top contender for analyzing statistical data. Tukey's g family of generalized distributions depend on the parameters, namely, g , which controls the skewness. This paper presents the probability density function (pdf) associated with a mixture of Tukey's g family of generalized distributions. The mixture of this class of skewed distributions is a generalization of Tukey's g family of distributions. In this paper, we calculate a closed form expression for the density and distribution of the mixture of two Tukey's g families of generalized distributions, which allows us to easily compute probabilities, moments, and related measures. This class of distributions contains the mixture of Log-symmetric distributions as a special case.

1. Introduction

The main focus of interest in financial economics is the distribution of stock market returns. Mandelbrot [1] suggested the family of stable Paretoian distributions for stock market returns. Fama [2] established that the normality assumption of the empirical data does not hold as the distribution is fat tailed. Kon [3] and Tse [4] used a mixture of normal distributions for stock return. Fielitz and Rozelle [5] proposed a mixture of nonnormal stable distributions for stock price. Consequently, greater emphasis has been placed on using distributions which have asymmetry and leptokurtic properties. Recently Jiménez et al. [6] proposed option pricing based mixture of log-skew-normal distributions. If extreme events tend to occur more frequently than normal events, then skewness and kurtosis of nonnormal distributions play an essential role for the volatility smile.

The most important and useful characteristic of Tukey's g - h family of distributions introduced by Tukey [7] is that it covers most of the Pearsonian family of distributions. It can also generate several known distributions, for example, log-normal, Cauchy, exponential, and Chi-squared (see Martínez and Iglewicz [8], page 363). From Tukey's g - h family of distribution, we obtain g distribution, which is closely related

to lognormal distribution and possesses similar properties of moments. Tukey's g - h family of distributions have been used to study financial markets. Badrinath and Chatterjee [9, 10] and Mills [11] used g - h to model the return on a stock index, as well as the return on shares in several markets. Dutta and Babbal [12] found that the skewness and leptokurtic behavior of LIBOR were modeled effectively using g - h distribution. Dutta and Babbal [13] used g - h to model interest rates and options on interest rates, while Tang and Wu [14] proposed a new method for the Decomposition of Portfolio VaR. Dutta and Perry [15] and recently Jiménez and Arunachalam [16] used g - h distribution to study the operational risk for heavy tailed severity models. Jiménez and Arunachalam [17] provided explicit expressions for VaR and CVaR calculations using the family of Tukey's g - h distributions. Currently Jiménez et al. [18] studied generalization of Tukey's g - h family of distributions, when the standard normal random variable is replaced by a continuous random variable U with mean 0 and variance 1.

The subfamily of g distributions exhibits skewness and has great importance in the study of asymmetric distributions for analyzing data. This kind of distribution allows us to obtain scaled Log-symmetric distributions. Vitiello and Poon [19] considered a simple mixture of two g distributions

for option pricing data. The purpose of this paper is to present a mixture of Tukey's g distributions and derive some statistical properties including the pdf and moment generating function and its properties.

The paper is organized as follows: Section 2 presents Tukey's g - h family of generalized distributions and its pdf, as well as the cumulative distribution function (cdf). In Section 3, some theoretical results of the mixture of two Tukey's g families of generalized distributions are presented and Section 4 explains the methodology of calculating estimation of parameters by the method of moments. Section 5 discusses the adjustment methodology of our proposed model to real data of Heating-Degree-Days (HDD) indices and finally, in Section 6, we conclude.

2. Tukey's g Family of Generalized Distributions

Tukey [7] introduced the family g - h distributions by means of two nonlinear transformations given by

$$Y = T_{g,h}(Z) = \frac{1}{g} (\exp \{gZ\} - 1) \exp \left\{ \frac{hZ^2}{2} \right\} \quad (1)$$

with $g \neq 0$, $h \in \mathbb{R}$, where the distribution of Z is standard normal. When these transformations are applied to a continuous random variable U with mean 0 and variance 1 such that its pdf $f_U(\cdot)$ is symmetric about the origin and cdf $F_U(\cdot)$, the transformation $T_{g,h}(U)$ is obtained, which henceforth will be termed Tukey's g - h generalized distribution. If $h = 0$, Tukey's g - h generalized distribution reduces to

$$T_{g,0}(U) = \frac{1}{g} (\exp \{gU\} - 1) \quad (2)$$

which is known as Tukey's g generalized distribution.

In order to model an arbitrary random variable X using the transformation given in (2), Hoaglin and Peters [20] introduced two new parameters, A (location) and B (scale), and proposed the following linear transformation:

$$X = A + BY \quad \text{with } Y = T_{g,0}(U). \quad (3)$$

The following properties for pdf, cdf, and quantile functions of Tukey's g generalized distribution were established by Jiménez et al. [18] in terms of the pdf and cdf of X as follows:

$$\begin{aligned} \text{pdf: } f_X(x; \lambda, g, \theta) &= \frac{1}{g(x-\theta)} f_U \left(\frac{1}{g} \ln \left(\frac{x-\theta}{B/g} \right) \right) \\ &\quad \text{if } g(x-\theta) > 0, \\ \text{cdf: } F_X(x; \lambda, g, \theta) &= F_U \left(\frac{1}{g} \ln \left(\frac{x-\theta}{B/g} \right) \right) \\ &\quad \text{if } g(x-\theta) > 0, \\ \text{qf: } F_U^{-1}(q) &= u_q = \frac{1}{g} \ln \left(\frac{x_q - \theta}{B/g} \right) \\ &\quad \text{if } g(x_q - \theta) > 0, \end{aligned} \quad (4)$$

where $\lambda = \ln(B/|g|)$ and $\theta = A - B/g$. We say that the random variable X has a Log-symmetric distribution (such distributions are all asymmetric; see for reference Johnson et al. [21] and Stuart and Ord [22]) with three parameters: threshold (θ), scale (λ), and shape (g), denoted by $X \sim \text{LS}(\lambda, g, \theta)$.

The first expression of (4) allows us to obtain the following pdf associated with Tukey's g distribution. Table 1 shows the parameters of the pdf of X that we obtain using a selected set of well known symmetrical distributions (from Jiménez et al. [18]).

The n th moment of the random variable $Y = T_{g,0}(U)$ is given by

$$\mu'_n(Y) = \frac{1}{g^n} \sum_{k=0}^n (-1)^k \binom{n}{k} M_U(\tilde{g}), \quad \text{if } g \neq 0, \quad (5)$$

where $\tilde{g} = (n-k)g$ and $M_U(t)$ is the moment generating function of the random variable U , which are even function; that is, $M_U(t) = M_U(-t)$. Table 2 shows parameters of the pdf and the moment generating function for a random variable U , using a selected set of well known symmetrical distributions.

Expression (5) allows us to obtain the moments of Tukey's g generalized distribution. The n th moment of the random variable X given by (3) can be obtained using the formula

$$\begin{aligned} \mathbb{E}[(X - \mathbb{E}[X])^n] &= \mu_n(X) \\ &= \left(\frac{B}{g} \right)^n \sum_{k=0}^n (-1)^k \binom{n}{k} M_U(\tilde{g}) M_U^k(g), \end{aligned} \quad (6)$$

where $B/g = \text{sgn}(g)e^\lambda$. Note that the above expression of the n th moment does not depend on the parameter θ . Formulas for calculating the standardized skewness, $\beta_1(X)$, and standardized excess kurtosis, $\beta_2(X)$, are given by

$$\begin{aligned} \beta_1(X) &= \text{sgn}(g) \\ &\quad \cdot \frac{(M_U(3g) - M_U^3(g)) - 3(M_U(2g) - M_U^2(g))M_U(g)}{[M_U(2g) - M_U^2(g)]^{3/2}}, \quad (7) \\ \beta_2(X) &= \frac{M_U(4g) - 4M_U(3g)M_U(g) + 3M_U^2(2g)}{[M_U(2g) - M_U^2(g)]^2} - 3, \end{aligned}$$

where $\text{sgn}(\cdot)$ denote the signum function. Note that these expressions only depend on the parameter g and its sign, respectively. Any LS distribution should satisfy the following test given in Stuart and Ord [22]:

$$\beta_2(X) - \beta_1^2(X) - 1 \geq 0. \quad (8)$$

3. The Mixture of Two g Distributions

We assume that Y follows a Log-Symmetric Mixture (LSMIX) distribution. Let us assume that $f_Y(y)$ is the weighted sum of m -component LS densities; that is,

$$f_Y(y; \Lambda) = \sum_{j=1}^m \omega_j f_U(y; \lambda_j, g_j, \theta_j). \quad (9)$$

TABLE 1: Parameters of the pdf of the random variable $Z = \ln(X)$.

Distribution of the r.v. U	Parameters	Distribution of the r.v. Z	Parameters			
	μ, a	σ, b	$g \neq 0$		μ, a	σ, b
Laplace	0	$\frac{\sqrt{2}}{2}$	$ g < \frac{\sqrt{2}}{n}$	Log-Laplace	$\ln\left(\frac{B}{ g }\right)$	$\frac{\sqrt{2}}{2} g $
Logistic	0	$\frac{\sqrt{3}}{\pi}$	$ g < \frac{\pi}{\sqrt{3}n}$	Log-Logistic	$\ln\left(\frac{B}{ g }\right)$	$\frac{\sqrt{3}}{\pi} g $
Normal	0	1	$g \in \mathbb{R}$	Lognormal	$\ln\left(\frac{B}{ g }\right)$	$ g $
HyperSec	0	$\frac{2}{\pi}$	$ g < \frac{\pi}{2n}$	LoghyperSec	$\ln\left(\frac{B}{ g }\right)$	$\frac{2}{\pi} g $
HyperCsc	0	$\frac{\sqrt{2}}{\pi}$	$ g < \frac{\pi}{\sqrt{2}n}$	LoghyperCsc	$\ln\left(\frac{B}{ g }\right)$	$\frac{\sqrt{2}}{\pi} g $

TABLE 2: Parameters of the pdf and moment generating functions of the random variable U .

Distribution of the r.v. U	Parameters	$M_U(g)$		
	μ, a	σ, b	$g \neq 0$	
Laplace	0	$\frac{\sqrt{2}}{2}$	$ g < \frac{\sqrt{2}}{n}$	$\frac{2}{2 - g^2}$
Logistic	0	$\frac{\sqrt{3}}{\pi}$	$ g < \frac{\pi}{\sqrt{3}n}$	$\sqrt{3}g \csc(\sqrt{3}g)$
Normal	0	1	$g \in \mathbb{R}$	$\exp\left\{\frac{1}{2}g^2\right\}$
HyperSec	0	$\frac{2}{\pi}$	$ g < \frac{\pi}{2n}$	$\sec(g)$
HyperCsc	0	$\frac{\sqrt{2}}{\pi}$	$ g < \frac{\pi}{\sqrt{2}n}$	$\sec^2\left(\frac{g}{\sqrt{2}}\right)$

We use the notation $Y \sim \text{LSMIX}(\Lambda)$, where $\Lambda = (\xi_1, \dots, \xi_m)$, and each element $\xi_j = (\omega_j, \lambda_j, g_j, \theta_j)$ is the parameter vector that defines the j th component and probability weights, ω_j , satisfying the conditions

$$\sum_{j=1}^m \omega_j = 1, \quad 0 < \omega_j < 1, \quad \text{for each } j. \quad (10)$$

$$F_X(x) = \begin{cases} 0, & \text{if } x \leq \theta_1, \\ \omega_1 F_U\left(\frac{\ln(x - \theta_1) - \lambda_1}{g_1}\right), & \text{if } \theta_1 < x \leq \theta_2, \\ \omega_1 F_U\left(\frac{\ln(x - \theta_1) - \lambda_1}{g_1}\right) + \omega_2 F_U\left(\frac{\ln(x - \theta_2) - \lambda_2}{g_2}\right), & \text{if } x > \theta_2, \end{cases} \quad (14)$$

where $\omega_2 = 1 - \omega_1$. Begin with the fact that the quartile function is the inverse of the cdf. Thus, replacing $x_q > \theta_2$ in (14), we obtain

$$F_X(x_q) = \omega_1 F_U\left(\frac{\ln(x_q - \theta_1) - \lambda_1}{g_1}\right) + \omega_2 F_U\left(\frac{\ln(x_q - \theta_2) - \lambda_2}{g_2}\right). \quad (15)$$

According to Titterton et al. [23] the two-component mixture of known distributions is set by two weights. Let

$$X = A + BY \quad \text{with } Y \sim \text{LSMIX}(\Lambda). \quad (11)$$

Then we can assume that $f_X(x)$ is the weighted sum of two Tukey's g mixture densities such that $g_1 g_2 > 0$. Thus

$$f_X(x) = \begin{cases} 0, & \text{if } x \leq \theta_1, \\ \frac{\omega_1}{g_1(x - \theta_1)} f_U(z_1), & \text{if } \theta_1 < x \leq \theta_2, \\ \frac{\omega_1}{g_1(x - \theta_1)} f_U(z_1) + \frac{1 - \omega_1}{g_2(x - \theta_2)} f_U(z_2), & \text{if } x > \theta_2, \end{cases} \quad (12)$$

where, without loss of generality, we let $\theta_1 < \theta_2$, $0 \leq \omega_1 \leq 1$ and for $j = 1, 2$

$$z_j = \frac{1}{g_j} \ln\left(\frac{x - \theta_j}{B/g_j}\right) \quad (13)$$

with $\theta_j = A - (B/g_j)$, $\lambda_j = (B/|g_j|)$. We use the notation $X \sim \text{GTMIX}(A, B, g_1, g_2, \omega_1)$. Vitiello and Poon [19] did not provide the piecewise nature of the mixture density function above in (12). In this case the cdf of X is given by

If we assume that $U \sim N(0, 1)$, (12) can be written as

$$f_X(x) = \begin{cases} 0, & \text{if } x \leq \theta_1, \\ \frac{\omega_1}{g_1(x - \theta_1)} \varphi(z_1), & \text{if } \theta_1 < x \leq \theta_2, \\ \frac{\omega_1}{g_1(x - \theta_1)} \varphi(z_1) + \frac{1 - \omega_1}{g_2(x - \theta_2)} \varphi(z_2), & \text{if } x > \theta_2, \end{cases} \quad (16)$$

where $\varphi(z)$ is the standard normal pdf. Note that the expression above matches the pdf of a mixture of three-parameter lognormal distributions. Letting $\theta_1 = \theta_2 = 0$, the above pdf reduces to that of a mixture of two-parameter lognormal distributions.

Given that every normal pdf is a version of the standard normal pdf then if $U \sim N(\mu, \sigma^2)$ we have

$$f_U(u, \mu, \sigma) = \frac{1}{\sigma} \varphi\left(\frac{u - \mu}{\sigma}\right), \quad \text{with } \mu \in \mathbb{R}, \sigma > 0, \quad (17)$$

and (12) can be written as

$$f_X(x) = \begin{cases} 0, & \text{if } x \leq \theta_1^*, \\ \frac{\omega_1}{g_1^* (x - \theta_1^*)} \varphi(z_1), & \text{if } \theta_1^* < x \leq \theta_2^*, \\ \frac{\omega_1}{g_1^* (x - \theta_1^*)} \varphi(z_1) + \frac{1 - \omega_1}{g_2^* (x - \theta_2^*)} \varphi(z_2), & \text{if } x > \theta_2^*. \end{cases} \quad (18)$$

If the parameters g_j are scaled by σ , that is, $g_j^* = \sigma g_j$, then

$$z_j = \frac{\ln(x - \theta_j^*) - \lambda_j^*}{g_j^*} \quad (19)$$

with $\theta_j^* = A - (B\sigma/g_j^*)$, $\lambda_j^* = \ln(B\sigma/|g_j^*|) + (\mu/\sigma)g_j^*$. Note that the expression above matches the pdf of a mixture of three-parameter lognormal distributions, which is a generalization of the pdf given in (16), and we use the notation $X \sim \text{LSMIX}(\lambda_1^*, \lambda_2^*, g_1^*, g_2^*, \theta_1^*, \theta_2^*, \omega_1)$. Similarly, we can obtain pdf of a mixture of distributions for the random variables listed in Table 1.

4. Estimation of the Mixtures of Two Tukey's g Distributions

In this section, we explain the estimation of the mixture of two Tukey's g distributions. The expected value of X is given by

$$\mathbb{E}[X] = \mu'_1 = \sum_{j=1}^2 \omega_j \left(\theta_j + \frac{B}{g_j} M_U(g_j) \right). \quad (20)$$

The n th raw moment of the random variable X is given by

$$\mathbb{E}[X^n] = \sum_{j=1}^2 \omega_j \sum_{k=0}^n \binom{n}{k} \left(\frac{B}{g_j} \right)^{n-k} \theta_j^k M_U(\bar{g}_j), \quad (21)$$

where $g_1 g_2 \neq 0$, $\bar{g}_j = (n - k)g_j$ and $M_U(t)$ is the moment generating function of the random variable U . The central moments μ_n of the random variable X are given by

$$\begin{aligned} \mathbb{E}[(X - \mu'_1)^n] &= \mu_n(X) \\ &= \sum_{j=1}^2 \omega_j \sum_{k=0}^n \binom{n}{k} \left(\frac{B}{g_j} \right)^{n-k} (\theta_j - \mu'_1)^k M_U(\bar{g}_j). \end{aligned} \quad (22)$$

The first five central moments are as follows:

$$\begin{aligned} \mu_1 &= \omega_1 \eta_1 + \omega_2 \eta_2 = 0, \\ \mu_2 &= \omega_1 (\sigma_1^2 + \eta_1^2) + \omega_2 (\sigma_2^2 + \eta_2^2), \\ \mu_3 &= \omega_1 (3\sigma_1^2 + \eta_1^2) \eta_1 + \omega_2 (3\sigma_2^2 + \eta_2^2) \eta_2, \\ \mu_4 &= \omega_1 (3\sigma_1^4 + 6\eta_1^2 \sigma_1^2 + \eta_1^4) \\ &\quad + \omega_2 (3\sigma_2^4 + 6\eta_2^2 \sigma_2^2 + \eta_2^4), \\ \mu_5 &= \omega_1 (15\sigma_1^4 + 10\eta_1^2 \sigma_1^2 + \eta_1^4) \eta_1 \\ &\quad + \omega_2 (15\sigma_2^4 + 10\eta_2^2 \sigma_2^2 + \eta_2^4) \eta_2, \end{aligned} \quad (23)$$

where for $j = 1, 2$

$$\begin{aligned} \eta_j &= \frac{B}{g_j} M_U(g_j) + \theta_j - \mu'_1, \\ \sigma_j^2 &= \left(\frac{B}{g_j} \right)^2 [M_U(2g_j) - M_U^2(g_j)]. \end{aligned} \quad (24)$$

Because $\theta_1 < \theta_2$, upon equating population moments to the corresponding sample moments, it follows from (23) that

$$\begin{aligned} \omega_1 \left(\frac{B}{g_1} M_U(g_1) + \theta_1 - m_1 \right) \\ + \omega_2 \left(\frac{B}{g_2} M_U(g_2) + \theta_2 - m_1 \right) = 0. \end{aligned} \quad (25)$$

Left-hand side of system (23) is multiplied by $\omega_1 + \omega_2 = 1$; the equations take the following form:

$$\begin{aligned} \omega_1 \eta_1 + \omega_2 \eta_2 &= 0, \\ \omega_1 (\sigma_1^2 + \eta_1^2 - m_2) + \omega_2 (\sigma_2^2 + \eta_2^2 - m_2) &= 0, \\ \omega_1 (3\eta_1 \sigma_1^2 + \eta_1^3 - m_3) + \omega_2 (3\eta_2 \sigma_2^2 + \eta_2^3 - m_3) &= 0, \\ \omega_1 (3\sigma_1^4 + 6\eta_1^2 \sigma_1^2 + \eta_1^4 - m_4) \\ + \omega_2 (3\sigma_2^4 + 6\eta_2^2 \sigma_2^2 + \eta_2^4 - m_4) &= 0, \\ \omega_1 (15\eta_1 \sigma_1^4 + 10\eta_1^3 \sigma_1^2 + \eta_1^5 - m_5) \\ + \omega_2 (15\eta_2 \sigma_2^4 + 10\eta_2^3 \sigma_2^2 + \eta_2^5 - m_5) &= 0, \end{aligned} \quad (26)$$

where m_i ($i = 1, 2, \dots$) denote the i th central moment of the sample. Equations (26) accordingly constitute a system of five equations to be solved simultaneously for the estimates of the five parameters A , B , g_1 , g_2 , and ω_1 .

Note that, from the first equation of system of (26), it follows that

$$\omega_1 = \frac{\eta_2}{\eta_2 - \eta_1}. \quad (27)$$

We eliminate ω_1 between the first and the subsequent equations of (26) in turn and thereby reduce the system to the following four equations in four unknowns A, B, g_1 , and g_2 :

$$\begin{aligned} \eta_1^{-1} (\sigma_1^2 + \eta_1^2 - m_2) &= \eta_2^{-1} (\sigma_2^2 + \eta_2^2 - m_2), \\ \eta_1^{-1} (3\eta_1\sigma_1^2 + \eta_1^3 - m_3) &= \eta_2^{-1} (3\eta_2\sigma_2^2 + \eta_2^3 - m_3), \\ \eta_1^{-1} (3\sigma_1^4 + 6\eta_1^2\sigma_1^2 + \eta_1^4 - m_4) \\ &= \eta_2^{-1} (3\sigma_2^4 + 6\eta_2^2\sigma_2^2 + \eta_2^4 - m_4), \\ \eta_1^{-1} (15\eta_1\sigma_1^4 + 10\eta_1^3\sigma_1^2 + \eta_1^5 - m_5) \\ &= \eta_2^{-1} (15\eta_2\sigma_2^4 + 10\eta_2^3\sigma_2^2 + \eta_2^5 - m_5). \end{aligned} \quad (28)$$

These systems of equations are solved computationally by using scientific software package and we do not need to verify the unique solution of the system as the parameter estimates. We skip further details and numerical illustration owing to space constraint.

5. Illustration

In this section we discuss some examples and applications of the results derived in Section 3 with two examples. In the first example, we discuss the pricing of a call option using a mixture of two Tukey's g -generalized distributions as an example to illustrate the results of Section 3. In the second example, we examine the empirical real data of Heating-Degree-Day to demonstrate usefulness of our approach of mixture of LS distributions.

Jiménez et al. [24] derived the option price of an European option assuming that the terminal price distribution follows a g -generalized distribution. Instead if we use a mixture of two Tukey's classes of g -generalized distributions, then the price of the call option denoted by $C(t, \tau; K)$ with a strike price K and maturity date $T = t + \tau$ can be expressed as follows:

$$\begin{aligned} C(t, \tau; K) &= \sum_{j=1}^2 \omega_j e^{-r\tau} \\ &\cdot \left[\int_{-\delta_j}^{\infty} e^{\lambda_j + g_j u} f_U(u) du - (K - \theta_j) F_U(\delta_j) \right], \end{aligned} \quad (29)$$

where $K > \theta_2$ and

$$\delta_j = \frac{\lambda_j - \ln(K - \theta_j)}{g_j} \quad \text{for } j = 1, 2. \quad (30)$$

When $U \sim N(0, 1)$, (29) reduces to

$$\begin{aligned} C(t, \tau; K) &= \sum_{j=1}^2 \omega_j e^{-r\tau} \\ &\cdot \left[e^{\lambda_j + (1/2)g_j^2} \Phi(\delta_j - g_j) - (K - \theta_j) \Phi(\delta_j) \right], \end{aligned} \quad (31)$$

where $\Phi(\cdot)$ denotes the cdf of a standard univariate normal variable. If we assume that $\theta_1 = \theta_2 = 0$, then (31) reduces to

$$\begin{aligned} C(t, \tau; K) \\ &= \sum_{j=1}^2 \omega_j e^{-r\tau} \left[e^{\lambda_j + (1/2)g_j^2} \Phi(\delta_j - g_j) - K \Phi(\delta_j) \right]. \end{aligned} \quad (32)$$

Note that when $g_j = \sigma_j \sqrt{\tau}$, these expressions coincide with the option pricing formula given in Bahra [25]. The authors also established closed form formula for the calculation of the sensitives measures of option pricing (Greek parameters of the option). Here we wish to observe that our mixture model uses less unknown parameters for calculating the option pricing, whereas Vitiello and Poon [19] used nine unknown parameters to obtain the same for the mixture of two g -distributions. It has been known that when we increase the number of parameters, we lose degrees of freedom and it is no longer acceptable for the best fit of data. This gives an advantage of our approach for the mixture of two g -generalized distributions.

We now present, as an example, the use of Heating-Degree-Days (HDD) in relation to winter temperature risk as a substitute for gas demand. HDD based contracts are listed on the Chicago Mercantile Exchange (CME). We consider an example that consists of monthly aggregate Heating-Degree-Day (HDD) data values at the Chicago O'Hare International Airport from December 1979 to December 2000 given in Wang [26] and explored also by Vitiello and Poon [19]. We describe first a LS distribution with three parameters based method to infer the implied risk-neutral probability density (RND). In Table 3, we present the estimated values of the three parameters of lognormal and Log-Logistic distributions; our interest is to compare with Vitiello and Poon [19] risk-neutral densities with our proposed mixture model.

The smaller value of the Kolmogorov-Smirnov (KS) test confirms that the data obeys the LS distributions with three parameters. We wish to observe that Anderson-Darling (AD) test is more sensitive to the tails of the LS distributions in comparison with KS test. In this case, we choose the Log-Logistic distribution as the best fit for the HDD data.

The implicit risk-neutral densities (RND) of LS distributions are shown in Figure 1 and compared with Figure 6 of Vitiello and Poon [19]. We have obtained a similar plot by our method with less unknown parameters than method given by Vitiello and Poon [19]. Furthermore, their KS test value of 13.6326% which is higher than the KS test values of Table 3 favors the best fit for the frequency of the LS distributions. Therefore, finite mixtures are attractive from the application viewpoint because of its flexibility and permit us to model various kinds of shaped distributions. In Table 4,

TABLE 3: Estimates for adjusting the LS(Λ).

Distribution of three parameters	Parameters			Test of adjusted	
	λ	g	θ	AD (%)	KS (%)
Lognormal	5.7407	0.5265	798.2540	37.53	12.3041
Log-Logistic	3.0225	283.8185	824.1814	32.68	11.1367

TABLE 4: Estimates for adjusting the mixture of LS(Λ).

Mixture of distributions	Parameters				KS (%) test
	$(\lambda_1; \lambda_2)$	$(g_1; g_2)$	$(\theta_1; \theta_2)$	ω	
Lognormal	6.5749	-0.1268	1797.9995	0.8188	8.0866
	11.0933	5.7574×10^{-4}	-64213.6284		
Log-Logistic	10.1338	-538.2433	1620.7711	0.8182	8.5928
	303.7429	6818.2894	-5309.4256		

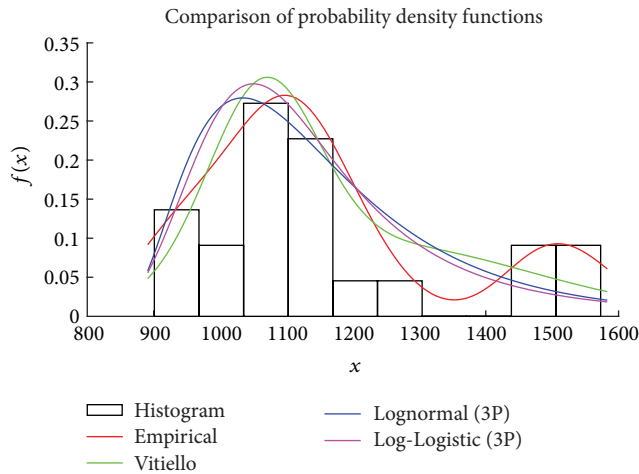


FIGURE 1: Empirical and LS(Λ) densities estimated from HDD.

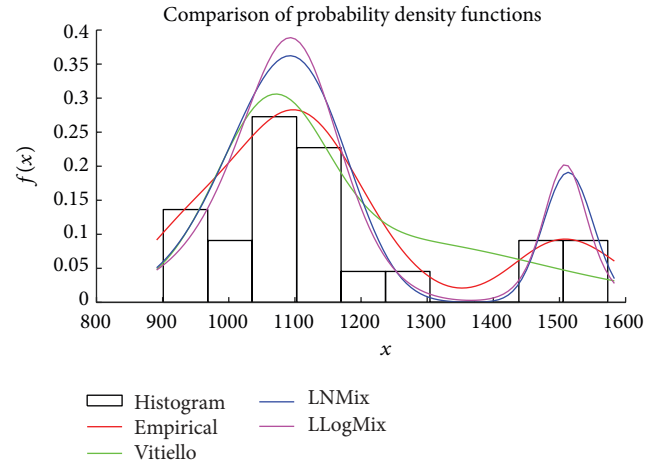


FIGURE 2: Empirical and two- g densities estimated from HDD.

we give the estimate values of the parameters of the mixture LS distributions. These parameters are estimated using (28). The estimated two g -densities and the implied risk-neutral densities (RND) are shown in Figure 2.

We observe that the bimodal LS mixture distribution has same fitting performance of the empirical distribution function (EDF) and lognormal mixture distribution gives best goodness of fit using the KS test.

6. Conclusions

This paper presents a mixture of Tukey’s g -generalized distributions and its properties. The methodology of estimating the unknown parameters by the method of moments is also presented. The proposed model has the advantage that it provides flexibility, when skewness, kurtosis, or other moments of the underlying distribution do not follow a normal distribution.

Some special cases of well known distributions are obtained from the proposed model.

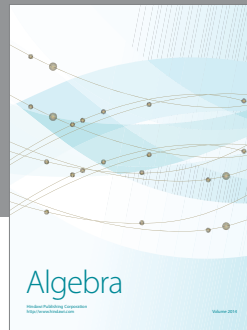
Competing Interests

The authors declare that they have no competing interests.

References

- [1] B. Mandelbrot, “The variation of certain speculative prices,” *The Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.
- [2] E. F. Fama, “The behavior of stock-market prices,” *The Journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.
- [3] S. J. Kon, “Models of stock returns—a comparison,” *The Journal of Finance*, vol. 39, no. 1, pp. 147–165, 1984.
- [4] Y. K. Tse, “Price and volume in the tokyo stock exchange: an exploratory study,” Working Paper 1573, University of Illinois at Urbana-Champaign, 1989.

- [5] B. D. Fielitz and J. P. Rozelle, "Stable distributions and the mixtures of distributions hypotheses for common stock returns," *Journal of the American Statistical Association*, vol. 78, no. 381, pp. 28–36, 1983.
- [6] J. A. Jiménez, V. Arunachalam, and G. M. Serna, "Option pricing based on a log-skew-normal mixture," *International Journal of Theoretical and Applied Finance*, vol. 18, no. 8, Article ID 1550051, 22 pages, 2015.
- [7] J. W. Tukey, "Modern techniques in data analysis," in *Proceedings of the NSF Sponsored Regional Research Conference*, Southeastern Massachusetts University, North Dartmouth, Mass, USA, 1977.
- [8] J. Martínez and B. Iglewicz, "Some properties of the tukey g and h family of distributions," *Communications in Statistics—Theory and Methods*, vol. 13, no. 3, pp. 353–369, 1984.
- [9] S. G. Badrinath and S. Chatterjee, "On measuring skewness and elongation in common stock return distributions: the case of the market index," *The Journal of Business*, vol. 61, no. 4, pp. 451–472, 1988.
- [10] S. G. Badrinath and S. Chatterjee, "A data-analytic look at skewness and elongation in common-stock-return distributions," *Journal of Business & Economic Statistics*, vol. 9, no. 2, pp. 223–233, 1991.
- [11] T. C. Mills, "Modelling skewness and kurtosis in the london stock exchange *ft - se* index return distributions," *The Statistician*, vol. 44, no. 3, pp. 323–332, 1995.
- [12] K. K. Dutta and D. F. Babbel, *On Measuring Skewness and Kurtosis in Short Rate Distributions: The Case of the US Dollar London Inter Bank Offer Rates*, The Wharton School, University of Pennsylvania, Philadelphia, Pa, USA, 2004.
- [13] K. K. Dutta and D. F. Babbel, "Extracting probabilistic information from the prices of interest rate options: tests of distributional assumptions," *Journal of Business*, vol. 78, no. 3, pp. 841–870, 2005.
- [14] X. Tang and X. Wu, "A new method for the decomposition of portfolio VaR," *Journal of Systems Science and Information*, vol. 4, no. 4, pp. 721–727, 2006.
- [15] K. K. Dutta and J. Perry, "A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital," Working Paper 06-13, Federal Reserve Bank of Boston, 2007.
- [16] J. A. Jiménez and V. Arunachalam, "Evaluating operational risk by an inhomogeneous counting process based on Panjer recursion," *The Journal of Operational Risk*, vol. 11, no. 1, pp. 1–21, 2016.
- [17] J. A. Jiménez and V. Arunachalam, "Using Tukey's g and h family of distributions to calculate value at risk and conditional value at risk," *Journal of Risk*, vol. 13, no. 4, pp. 95–116, 2011.
- [18] J. A. Jiménez, V. Arunachalam, and G. M. Serna, "A generalization of Tukey's g-h family of distributions," *Journal of Statistical Theory and Applications*, vol. 14, no. 1, pp. 28–44, 2015.
- [19] L. Vitiello and S.-H. Poon, "General equilibrium and risk neutral framework for option pricing with a mixture of distributions," *The Journal of Derivatives*, vol. 15, no. 4, pp. 48–60, 2008.
- [20] D. C. Hoaglin and S. C. Peters, *Software for Exploring Distribution Shape*, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 1979.
- [21] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons, New York, NY, USA, 1994.
- [22] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics: Distribution Theory*, vol. 1, Edward Arnold, London, UK; Halsted Press, New York, NY, USA, 6th edition, 1994.
- [23] D. M. Titterton, A. F. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*, vol. 7, John Wiley & Sons, New York, NY, USA, 1985.
- [24] J. A. Jiménez, V. Arunachalam, and G. M. Serna, "Option pricing based on the generalised Tukey distribution," *International Journal of Financial Markets and Derivatives*, vol. 3, no. 3, pp. 191–221, 2014.
- [25] B. Bahra, "Implied risk-neutral probability density functions from option prices: a central bank perspective," in *Forecasting Volatility in the Financial Markets*, J. Knight and S. Satchell, Eds., pp. 201–226, Elsevier, 3rd edition, 1997.
- [26] S. S. Wang, "A universal framework for pricing financial and insurance risks," *ASTIN Bulletin*, vol. 32, no. 2, pp. 213–234, 2002.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

