A Special Weight for Inverse Gaussian Mixing Distribution in Normal Variance Mean Mixture with Application

Calvin B. Maina[1]*, Patrick G. O. Weke[2], Carolyne A. Ogutu[2], Joseph A. M. Ottieno[2]

[1]Department of Mathematics and Actuarial Science, Kisii University, Kisii, Kenya
[2]School of Mathematics, University of Nairobi, Nairobi, Kenya
Email: *calvinmaina@kisiiuniversity.ac.ke

Abstract

Normal Variance-Mean Mixture (NVMM) provides a general framework for deriving models with desirable properties for modelling financial market variables such as exchange rates, equity prices, and interest rates measured over short time intervals, *i.e.* daily or weekly. Such data sets are characterized by non-normality and are usually skewed, fat-tailed and exhibit excess kurtosis. The Generalised Hyperbolic distribution (GHD) introduced by Barndorff-Nielsen (1977) which act as Normal variance-mean mixtures with Generalised Inverse Gaussian (GIG) mixing distribution nest a number of special and limiting case distributions. The Normal Inverse Gaussian (NIG) distribution is obtained when the Inverse Gaussian is the mixing distribution, *i.e.*, the index parameter of the GIG is $-\frac{1}{2}$. The NIG is very popular because of its analytical tractability. In the mixing mechanism, the mixing distribution characterizes the prior information of the random variable of the conditional distribution. Therefore, considering finite mixture models is one way of extending the work. The GIG is a three parameter distribution denoted by $GIG(\lambda, \delta, \gamma)$ and nest several special and limiting cases. When $\lambda = -\frac{1}{2}$, we have $GIG\left(-\frac{1}{2}, \delta, \gamma\right)$ which is called an Inverse Gaussian (IG) distribution. When $\lambda = \frac{1}{2}$, $\lambda = -\frac{3}{2}$, $\lambda = \frac{3}{2}$, we have $GIG\left(\frac{1}{2}, \delta, \gamma\right)$, $GIG\left(-\frac{3}{2}, \delta, \gamma\right)$ and $GIG\left(\frac{3}{2}, \delta, \gamma\right)$ distributions respectively. These distributions are related to $GIG\left(-\frac{1}{2}, \delta, \gamma\right)$ and are called weighted inverse Gaussian distributions. In

this work, we consider a finite mixture of $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ and $GIG\left(\frac{3}{2}, \delta, \gamma\right)$ and show that the mixture is also a weighted Inverse Gaussian distribution and use it to construct a NVMM. Due to the complexity of the likelihood, direct maximization is difficult. An EM type algorithm is provided for the Maximum Likelihood estimation of the parameters of the proposed model. We adopt an iterative scheme which is not based on explicit solution to the normal equations. This subtle approach reduces the computational difficulty of solving the complicated quantities involved directly to designing an iterative scheme based on a representation of the normal equation. The algorithm is easily programmable and we obtained a monotonic convergence for the data sets used.

Keywords

1. Introduction

Normal Inverse Gaussian (NIG) distribution introduced by Barndorff-Nielsen [1] is the popular normal variance-mean mixture, with an inverse Gaussian mixing distribution that has been used to handle non-normal distributed data.

Efforts have been done and are still going on to identify alternatives to the NIG distributions: Aas and Haff [2] [3] considered Skew Student's t distribution as an alternative. Corlu and Corlu [4] used Generalized Lambda distribution, a generalization of Tukey's Lambda as an alternative.

Finite mixtures are more flexible (robust) than single mixing distributions; a result that has been stated a lot in statistical literature.

Very few studies have used finite mixtures as mixing distributions. Hence an attempt has been done here in this paper of ours. This idea is motivated by the fact that finite mixtures are more flexible than single distributions. Nadarajah, Zhang and Chan [5] have stated that finite mixtures of normal distributions are flexible than single normal distribution; finite mixtures of stable distributions are flexible than a single stable distribution; finite mixtures of student's t distributions are flexible than single student's distribution.

The second motivation is that very few studies on continuous mixtures have used finite mixtures as mixing distributions; with the exception of Lindley [6] distribution and its generalization in Poisson mixtures (e.g., Sankaran [7]; Mahmoudi and Zakerzadeh [8]). Lindley distribution and its generalizations are basically finite gamma mixtures.

GIG is a three parameter distribution denoted by $GIG(\lambda, \delta, \gamma)$. When $\lambda = -\frac{1}{2}$, the mixing distribution is $GIG\left(-\frac{1}{2}, \delta, \gamma\right)$ which is Inverse Gaussian and the mixture is Normal Inverse Gaussian (NIG). In our present work, we con-

sider two other special cases: $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ and $GIG\left(\frac{3}{2}, \delta, \gamma\right)$. A finite mixture of these cases is shown to be a weighted Inverse Gaussian distribution. The concept of weighted distribution was introduced by Fisher [9] and elaborated by Patil and Rao [10]. The reciprocal Inverse Gaussian, *i.e.*, $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ and the finite mixture of $GIG\left(-\frac{1}{2}, \delta, \gamma\right)$ and $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ are shown to be weighted inverse Gaussian distribution by Gupta and Kundu [11].

The finite mixture model is used as a mixing distribution in Normal Variance Mean mixture to construct the proposed model. The maximum likelihood parameter estimates of the mixture are obtained via the Expectation-Maximization (EM) algorithm. For data analysis we have used three datasets: s&p500 index, Range Resource Corporation (RRC) and Shares of Chevron Corporation (CVX). We consider log returns for the period 3/01/2000 to 1/07/2013 with 702 observations for each dataset.

2. Weighted Inverse Gaussian Distribution

Consider a random variable Z with a probability distribution $f(z)$. Let $w(Z)$ be a random function of Z. We can construct a new probability distribution

$$g(z) = \frac{w(z)}{E[w(Z)]} f(z), -\infty < x < \infty \tag{1}$$

which is a weighted distribution of $f(z)$. The concept of weighted distribution was introduced by Fisher [9] and elaborated by Patil and Rao [10].

In this work we consider weighted distribution for the Inverse Gaussian (IG) distribution.

Now, suppose $Z \sim IG(\gamma, \delta)$ the Inverse Gaussian distribution with parameters γ and δ and probability density function given by

$$f(z) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta\gamma) z^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right) \tag{2}$$

Let

$$w(Z) = Z + \frac{Z^2}{1 + \delta\gamma}$$

$$E[w(Z)] = \frac{\delta(\gamma^2 + 1)}{\gamma^3}$$

$$\therefore g(z) = \frac{\gamma^3}{\delta(\gamma^2 + 1)}\left[z + \frac{z^2}{1 + \delta\gamma}\right] f(z) \tag{3}$$

which is also a finite mixture of $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ and $GIG\left(\frac{3}{2}, \delta, \gamma\right)$. That is

$$g(z) = pGIG\left(\frac{1}{2}, \delta, \gamma\right) + (1-p)GIG\left(\frac{3}{2}, \delta, \gamma\right)$$

with

$$p = \frac{\gamma^2}{\gamma^2 + 1}$$

The mean and variance for the weighted distribution are

$$E[Z] = \frac{\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)} \tag{4}$$

$$var(Z)$$
$$= \frac{\left(\left(\delta^2\gamma^2 + 3\delta\gamma + 3\right)(1+\delta\gamma)\gamma^2 + \left(\delta^3\gamma^3 + 6\delta^2\gamma^2 + 15\delta\gamma + 15\right)\right)(1+\delta\gamma)(\gamma^2+1)}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2} \tag{5}$$
$$- \frac{\left(\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3\right)^2}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2}$$

3. Normal Variance-Mean Mixture

A stochastic representation of a Normal Variance-Mean mixture is given by

$$X = \mu + \beta Z + \sqrt{Z}Y$$

where

$$Y \sim N(0,1)$$

which gives the hierarchical representation

$$X/Z = z \sim N(\mu + \beta z, z) \tag{6}$$

being the conditional pdf and *g(z)* the mixing distribution.

Construction of the Mixed Model

Suppose the mixing distribution follows formula (3) and noting that

$$\frac{w(z)}{E[w(Z)]} = \frac{\gamma^3}{\delta(\gamma^2+1)}\left(Z + \frac{Z^2}{1+\delta\gamma}\right) \tag{7}$$

the mixed model becomes

$$f(x) = \frac{\delta e^{\delta\gamma}}{2\pi} e^{\beta(x-\mu)} \int_0^\infty \frac{w(z)}{E[w(Z)]} z^{-2} e^{-\frac{1}{2}\left(\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right)} dz$$

$$= \frac{\delta e^{\delta\gamma}}{2\pi} e^{\beta(x-\mu)} \frac{\gamma^3}{\delta(\gamma^2+1)} \int_0^\infty \left(z + \frac{z^2}{1+\delta\gamma}\right) z^{-2} e^{-\frac{1}{2}\left(\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right)} dz$$

$$= \frac{\gamma^3 e^{\delta\gamma}}{2\pi(\gamma^2+1)} e^{\beta(x-\mu)} \int_0^\infty \left(z^{0-1} + \frac{z^{1-1}}{1+\delta\gamma}\right) e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z}\right)} dz$$

980

$$f(x) = \frac{\gamma^3 e^{\delta\gamma} e^{\beta(x-\mu)}}{\pi(\gamma^2+1)} \left\{ K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \right\}$$

$$= \frac{\gamma^3 e^{\delta\gamma} e^{\beta(x-\mu)}}{\alpha\pi(1+\delta\gamma)(\gamma^2+1)} \left\{ \alpha(1+\delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \right\}$$

(8)

With the following properties

$$E(X) = \mu + \beta \frac{\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)}$$

(9)

$$var(X)$$
$$= \frac{\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)}$$
$$- \beta^2 \frac{\left(\left(\delta^2\gamma^2 + 3\delta\gamma + 3\right)(1+\delta\gamma)\gamma^2 + \left(\delta^3\gamma^3 + 6\delta^2\gamma^2 + 15\delta\gamma + 15\right)\right)(1+\delta\gamma)(\gamma^2+1)}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2}$$
$$- \frac{\left(\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3\right)^2}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2}$$

(10)

Note

$K_\lambda(\omega)$ denotes modified Bessel function of the third kind with index λ evaluated at ω.

$$K_\lambda(\omega) = \frac{1}{2}\int_0^\infty x^{\lambda-1} e^{-\frac{1}{2}\left(x+\frac{1}{x}\right)} dx$$

(11)

with the following properties

a) $K_{\frac{1}{2}}(\omega) = K_{-\frac{1}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}$ (12)

b) $K_{\frac{3}{2}}(\omega) = K_{-\frac{3}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}\left(1 + \frac{1}{\omega}\right)$ (13)

c) $K_{\frac{5}{2}}(\omega) = K_{-\frac{5}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}\left(1 + \frac{3}{\omega} + \frac{3}{\omega^2}\right)$ (14)

d) $K_{\frac{7}{2}}(\omega) = K_{-\frac{7}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}\left(1 + \frac{6}{\omega} + \frac{15}{\omega^2} + \frac{15}{\omega^3}\right)$ (15)

e) $K_{\frac{9}{2}}(\omega) = K_{-\frac{9}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}\left(1 + \frac{10}{\omega} + \frac{45}{\omega^2} + \frac{105}{\omega^3} + \frac{105}{\omega^4}\right)$ (16)

f) $K_{\frac{11}{2}}(\omega) = K_{-\frac{11}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega}\left(1 + \frac{15}{\omega} + \frac{105}{\omega^2} + \frac{420}{\omega^3} + \frac{945}{\omega^4} + \frac{945}{\omega^5}\right)$ (17)

which are necessary in deriving the properties and estimates of the proposed models. For more definition and properties (see Abramowitz and Stegun [12]).

4. Parameter Estimation

Given a random sample of size n from our proposed model the log-likelihood function is given by

$$l = \log L = \sum_{i=1}^{n} \log f(x_i)$$

$$= \sum_{i=1}^{n} \left\{ 3\log \gamma + \delta\gamma + \beta(x_i - \mu) - \log\left(\alpha\pi(\gamma^2 + 1)(1 + \delta\gamma)\right) \right.$$

$$\left. + \log\left[\alpha(1 + \delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \right] \right\} \qquad (18)$$

$$= 3n\log \gamma + n\delta\gamma + \beta\sum_{i=1}^{n} x_i - n\beta\mu - n\log\left(\alpha\pi(\gamma^2 + 1)(1 + \delta\gamma)\right)$$

$$+ \sum_{i=1}^{n} \log\left[\alpha(1 + \delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \right]$$

The derivatives of the log-likelihood function involve the Bessel function and the parameters are difficult to separate. Direct maximization is not an easy. In this work we adopt the Expectation-Maximization (EM) algorithm. The EM-algorithm, introduced by Dempster *et al.* [13], is a powerful technique for maximum likelihood estimation for data containing missing values or data that can be considered as containing missing values.

Karlis [14] considered the mixing operation responsible for producing missing values. He applied the algorithm in estimating the parameters for the Normal Inverse Gaussian. The iterative scheme was based on explicit solution to the normal equations. In our present work the iterative scheme is based on a representation of the normal equation. This subtle approach overcomes the computational difficulty of solving the complicated quantities numerically.

Assume that the true data are made of an observed part X and unobserved part Z. This then ensures the log likelihood of the complete data (x_i, z_i) for $i = 1, 2, 3, \cdots, n$ factorizes into two parts, *i.e.*,

$$\log L = \log \prod_{i=1}^{n} f(x_i / z_i) + \log \prod_{i=1}^{n} g(z_i)$$

$$= \sum_{i=1}^{n} \log f(x_i / z_i) + \sum_{i=1}^{n} \log g(z_i)$$

where

$$l_1 = \sum_{i=1}^{n} \log f(x_i / z_i)$$

and

$$l_2 = \sum_{i=1}^{n} \log g(z_i)$$

4.1. M-Step for Conditional Distribution

Since the conditional distribution for the six models is normal distribution as presented in formula (6), we have

$$l_1 = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log z_i - \sum_{i=1}^{n}\frac{(x_i - \mu - \beta z_i)^2}{2z_i}$$

Therefore

$$\frac{\partial}{\partial\beta}l_1 = 0 \Rightarrow \sum_{i=1}^{n}(x_i - \hat{\mu} - \hat{\beta}z_i) = 0$$

$$i.e., \quad \sum_{i=1}^{n}x_i - n\hat{\mu} - \hat{\beta}\sum_{i=1}^{n}z_i = 0$$

$$\therefore \hat{\mu} = \bar{x} - \hat{\beta}\bar{z}$$

where $\bar{x} = \sum_{i=1}^{n}\frac{x_i}{n}$ and $\bar{z} = \sum_{i=1}^{n}\frac{z_i}{n}$.

Similarly,

$$\frac{\partial}{\partial\mu}l_1 = 0 \Rightarrow \sum_{i=1}^{n}\frac{x_i}{z_i} - \hat{\mu}\sum_{i=1}^{n}\frac{1}{z_i} - n\hat{\beta} = 0$$

$$\therefore \sum_{i=1}^{n}\frac{x_i}{z_i} - \bar{x}\sum_{i=1}^{n}\frac{1}{z_i} + \hat{\beta}\bar{z}\sum_{i=1}^{n}\frac{1}{z_i} - n\hat{\beta} = 0$$

$$\therefore \hat{\beta} = \frac{\sum_{i=1}^{n}\frac{x_i}{z_i} - \bar{x}\sum_{i=1}^{n}\frac{1}{z_i}}{n - \bar{z}\sum_{i=1}^{n}\frac{1}{z_i}}$$

4.2. M-Step for the Mixing Distribution

$$l_2 = 3n\log\gamma + n\delta\gamma - \frac{n}{2}\log(2\pi) - n\log(\gamma^2 + 1) - n\log(1 + \delta\gamma)$$

$$- \frac{1}{2}\sum_{i=1}^{n}\log z_i + \sum_{i=1}^{n}\left[\log(1 + \delta\gamma + z_i)\right] - \frac{\gamma^2}{2}\sum_{i=1}^{n}z_i - \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} \tag{19}$$

Maximizing with respect to δ and γ we have the following representation

$$\frac{3n + 2n\delta\gamma}{\gamma(1+\delta)} - \gamma\left(\frac{2n}{\gamma^2 + 1} + \sum_{i=1}^{n}z_i\right) + \delta\left(n + \sum_{i=1}^{n}\frac{1}{1 + \delta\gamma + z_i}\right) = 0 \tag{20}$$

$$\frac{n\delta\gamma^2}{1 + \delta\gamma} + \gamma\sum_{i=1}^{n}\frac{1}{1 + \delta\gamma + z_i} - \delta\sum_{i=1}^{n}\frac{1}{z_i} = 0 \tag{21}$$

There is need to estimate the values for Z_i, $\frac{1}{Z_i}$ and Z_i^2 using the posterior expectations as follows:

Posterior Expectation

$$E(Z/X) = \frac{\int_0^{\infty} z\left(z + \frac{z^2}{1 + \delta\gamma}\right)z^{-2}e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\mathrm{d}z}{\int_0^{\infty}\left(z + \frac{z^2}{1 + \delta\gamma}\right)z^{-2}e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\mathrm{d}z}$$

$$= \frac{\int_0^\infty \left(z^{1-1} + \frac{z^{2-1}}{1+\delta\gamma} \right) e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}{\int_0^\infty \left(z^{0-1} + \frac{z^{1-1}}{1+\delta\gamma} \right) e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}$$

$$= \frac{\frac{\delta\sqrt{\phi(x)}}{\alpha} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \left[\frac{\delta\sqrt{\phi(x)}}{\alpha} \right]^2 \frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \tag{22}$$

$$= \frac{\delta\sqrt{\phi(x)}(1+\delta\gamma) K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \frac{\delta^2\phi(x)(1+\delta\gamma)}{\alpha} K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha(1+\delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

$$= \frac{\alpha\delta\sqrt{\phi(x)}(1+\delta\gamma) K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta\phi(x)(1+\delta\gamma) K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha^2(1+\delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \alpha\delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

Similarly,

$$E\left(\frac{1}{Z} \middle/ X \right) = \frac{\int_0^\infty z^{-1}\left(z + \frac{z^2}{1+\delta\gamma} \right) z^{-2} e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}{\int_0^\infty \left(z + \frac{z^2}{1+\delta\gamma} \right) z^{-2} e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}$$

$$= \frac{\frac{1}{2}\int_0^\infty \left(z^{-1-1} + \frac{z^{0-1}}{1+\delta\gamma} \right) e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}{\frac{1}{2}\int_0^\infty \left(z^{0-1} + \frac{z^{1-1}}{1+\delta\gamma} \right) e^{-\frac{\alpha^2}{2}\left(z + \frac{\delta^2\phi(x)}{\alpha^2 z} \right)} dz}$$

$$= \frac{\left[\frac{\delta\sqrt{\phi(x)}}{\alpha} \right]^{-1} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \frac{K_0\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

$$E\left(\frac{1}{Z} \middle/ X \right) = \frac{\frac{\alpha^2(1+\delta\gamma)}{\delta\sqrt{\phi(x)}} K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \alpha K_0\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha(1+\delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \alpha\delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \tag{23}$$

$$= \frac{\alpha^2(1+\delta\gamma) K_1\left(\alpha\delta\sqrt{\phi(x)}\right) + \alpha\delta\sqrt{\phi(x)} K_0\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha\delta\sqrt{\phi(x)} K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \delta^2\phi(x) K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

Similarly

$$E\left(Z^2 \middle/ X \right) = \frac{\alpha(1+\delta\gamma)\delta^2\phi(x) K_2\left(\alpha\delta\sqrt{\phi(x)}\right) + \left(\delta\sqrt{\phi(x)}\right)^3 K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha^3(1+\delta\gamma) K_0\left(\alpha\delta\sqrt{\phi(x)}\right) + \alpha^2\delta\sqrt{\phi(x)} K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \tag{24}$$

Now let $s_i = E\left(Z_i/X_i\right)$, $w_i = E\left(\dfrac{1}{Z_i}\Big/X_i\right)$ and $v_i = E\left(Z_i^2/X_i\right)$ and

$$\overline{s} = \sum_0^n \frac{s_i}{n}$$

$$\overline{w} = \sum_0^n \frac{w_i}{n}$$

Therefore the k-th iterations are:

$s_i^{(k+1)}$
$$= \frac{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\left(1+\delta^{(k)}\gamma^{(k)}\right)K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\delta^{(k)}\phi^{(k)}(x)\left(1+\delta^{(k)}\gamma^{(k)}\right)K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}{\left(\alpha^{(k)}\right)^2\left(1+\delta^{(k)}\gamma^{(k)}\right)K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}$$

$$w_i^{(k+1)} = \frac{\left(\alpha^{(k)}\right)^2\left(1+\delta^{(k)}\gamma^{(k)}\right)K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\left(\delta^{(k)}\right)^2\phi^{(k)}(x)K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}$$

$$v_i^{(k+1)} = \frac{\alpha^{(k)}\left(1+\delta^{(k)}\gamma^{(k)}\right)\left(\delta^{(k)}\right)^2\phi^{(k)}(x)K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\left(\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)^3K_3\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}{\left(\alpha^{(k)}\right)^3\left(1+\delta^{(k)}\gamma^{(k)}\right)K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+\left(\alpha^{(k)}\right)^2\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)}$$

These can be used to obtain the $(k+1)$-th values as follows

$$\hat{\gamma}^{(k+1)} = \frac{\dfrac{3+2\delta^{(k)}\gamma^{(k)}}{\gamma^{(k)}\left(1+\delta^{(k)}\gamma^{(k)}\right)}+\delta^{(k)}\left[1+\dfrac{1}{n}\sum_{i=1}^n\dfrac{1}{1+\delta^{(k)}\gamma^{(k)}+s_i^{(k)}}\right]}{\dfrac{2}{\left(\gamma^{(k)}\right)^2+1}+\overline{s}^{(k)}} \tag{25}$$

$$\hat{\delta}^{(k+1)} = \frac{\dfrac{n\delta^{(k)}\left(\gamma^{(k+1)}\right)^2}{1+\delta^{(k)}\gamma^{(k+1)}}+\gamma^{(k+1)}\sum_{i=1}^n\dfrac{1}{1+\delta^{(k)}\gamma^{(k+1)}+s_i}}{n\overline{w}^{(k)}} \tag{26}$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^n x_i w_i^{(k+1)}-\overline{x}\sum_{i=1}^n w_i^{(k+1)}}{n-\overline{s}^{(k+1)}\sum_{i=1}^n w_i^{(k+1)}} \tag{27}$$

$$\hat{\mu}^{(k+1)} = \overline{x}-\hat{\beta}^{(k+1)}\overline{s}^{(k+1)} \tag{28}$$

$$\hat{\alpha}^{(k+1)} = \sqrt{\left(\hat{\beta}^{(k+1)}\right)^2+\left(\hat{\gamma}^{(k+1)}\right)^2} \tag{29}$$

The $(k+1)$-th iteration of the log-likelihood function becomes

$$
\begin{aligned}
l^{(k+1)} = {}& 3n\log\gamma^{(k+1)}+n\delta^{(k+1)}\gamma^{(k+1)}+\beta^{(k+1)}\sum_{i=1}^n x_i-n\beta^{(k+1)}\mu^{(k+1)} \\
& -n\log\left(\alpha^{(k+1)}\pi\left(\left(\gamma^{(k+1)}\right)^2+1\right)\left(1+\delta^{(k+1)}\gamma^{(k+1)}\right)\right) \\
& +\sum_{i=1}^n\log\left[\alpha^{(k+1)}\left(1+\delta^{(k+1)}\gamma^{(k+1)}\right)K_0\left(\alpha^{(k+1)}\delta^{(k+1)}\sqrt{\phi^{(k+1)}(x)}\right)\right. \\
& \left.+\delta^{(k+1)}\sqrt{\phi^{(k+1)}(x)}K_1\left(\alpha^{(k+1)}\delta^{(k+1)}\sqrt{\phi^{(k+1)}(x)}\right)\right]
\end{aligned} \tag{30}
$$

5. Application

Let (P_t) denote the price process of a security at time t, in particular of a stock. In order to allow comparison of investments in different securities we shall investigate the rates of return defined by

$$X_t = \log P_t - \log P_{t-1}$$

In this section we consider three datasets for data analysis: Range Resource Corporation (RRC), Shares of Chevron Corporation (CVX) and s&p500 index. The period 3/01/2000 to 1/07/2013 with 702 observations for each data set is considered. The histogram for the weekly log-returns in **Figures 1-3** for s&p 500, RRC and CVX show that the data is negatively skewed and exhibiting heavy tails. The Q-Q plot shows that the normal distribution is not a good fit for the data especially at the tails. This is typical the other data sets.

Table 1 provides descriptive statistics of the data sets for the return series in consideration. We observe that the excess kurtosis that indicates the leptokurtic behaviour of the returns. The log-returns has a distributions with relatively heavier
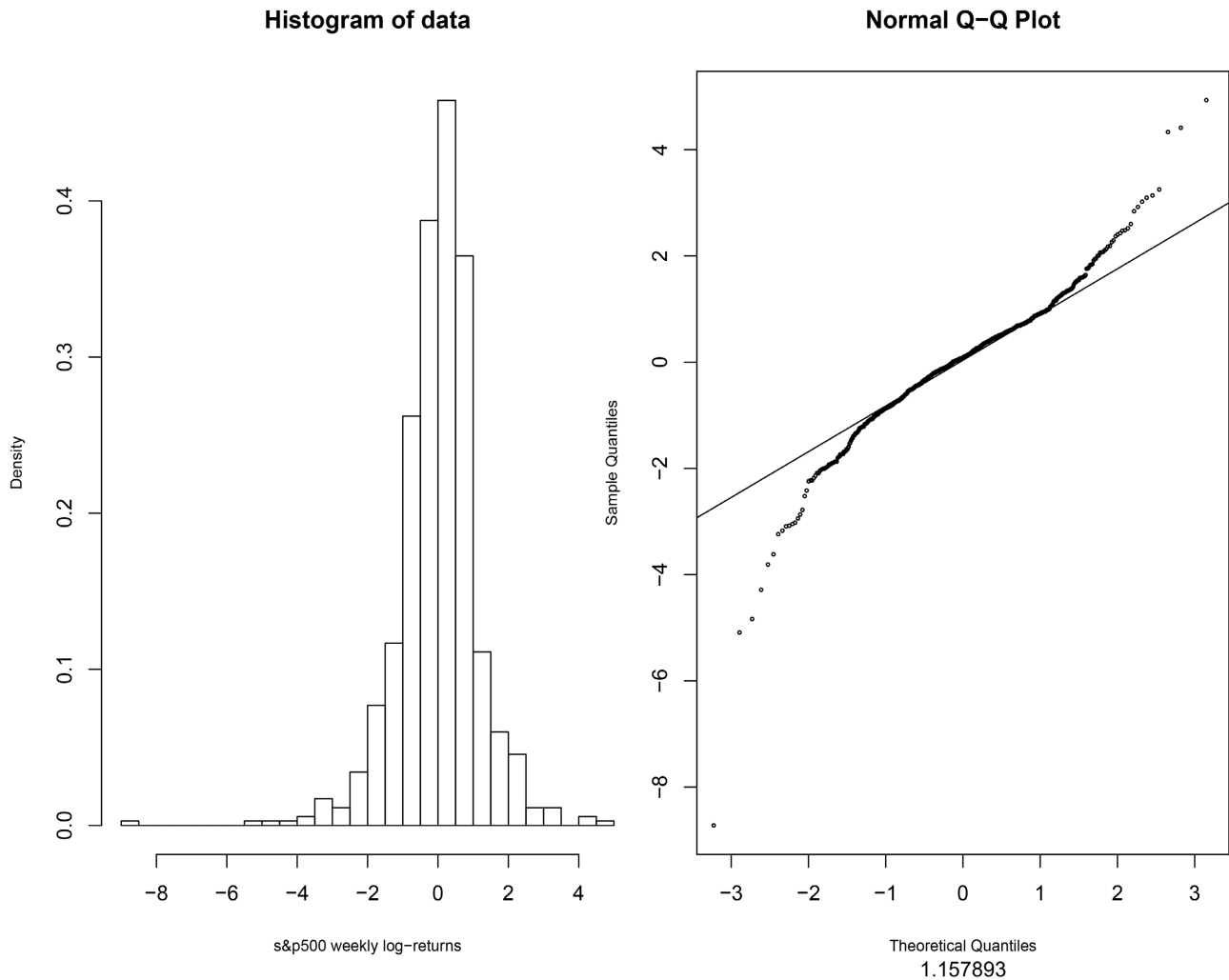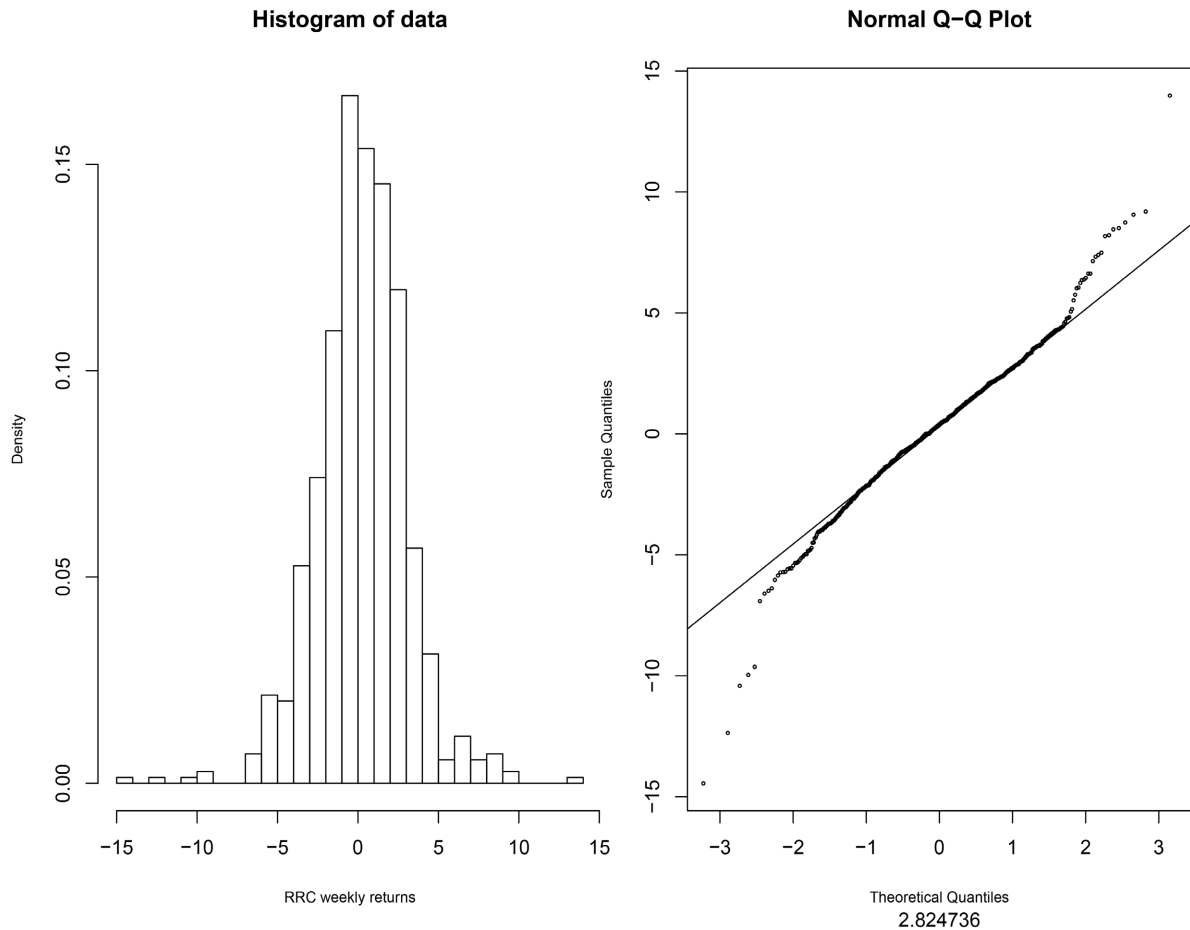

Figure 1. Histogram and Q-Q plot for s&p500 weekly log-returns.

Table 1. Summary statistics for RRC weekly log-returns.

dataset	Minimum	Standard.dev	skewness	exc.kurtosis	Maximum	Mean	N
RRC	−14.4465	2.824736	−0.1886714	2.768252	13.9830	0.2333	702
CVX	−13.76112	1.480436	−1.297339	11.10113	6.71410	0.08711	702
s&p500	−8.722261	1.157893	−0.7851156	6.408709	4.931805	0.006697	702

Figure 2. Histogram and Q-Q plot for RRC weekly log-returns.

tails than the normal distribution. We observe skewness for the data sets which indicates that the two tails of the returns behave slightly differently.

We use Karlis [14] formulation to obtain the initial values for the EM algorithm. For this dataset, the values obtained are in Table 2.

The stopping criterion is when

$$\frac{l^{(k)} - l^{(k-1)}}{l^{(k)}} < tol \tag{31}$$

where *tol* is the tolerance level. Table 3 below shows the NIG maximum likelihood parameter estimates using the EM-algorithm at $tol = 10^{-3}$.

We now wish to obtain the maximum likelihood parameter estimates for the proposed model via the EM algorithm. Tables 3-5 illustrate monotonic convergence
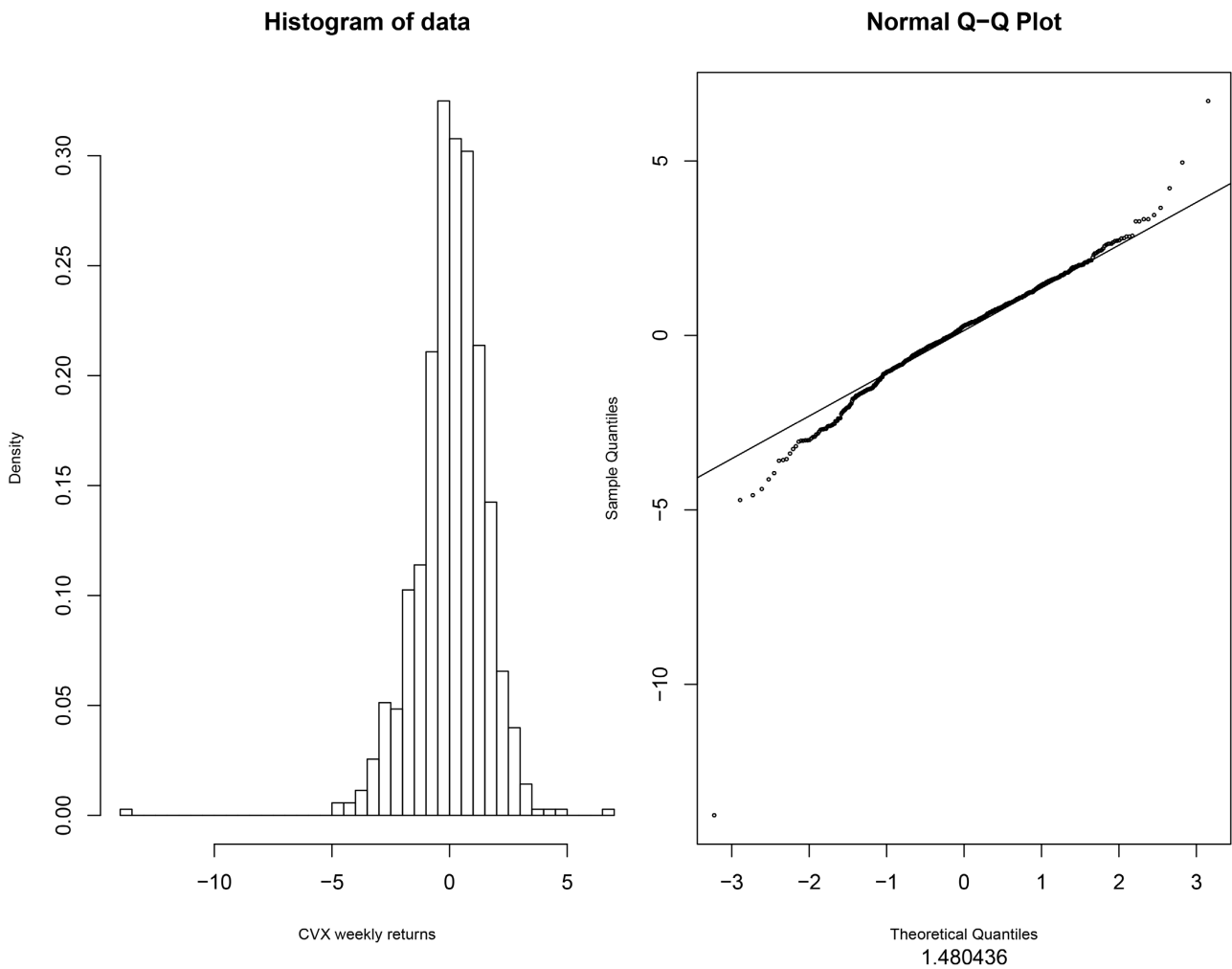
Histogram of data

Normal Q-Q Plot

1.480436

Figure 3. Histogram and Q-Q plot for CVX weekly log-returns.

Table 2. NIG method of moment estimates for the data sets.

dataset	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$
RRC	0.3722511	−0.02456226	2.950864	0.4284473
CVX	0.4190067	−0.1054991	0.8324058	0.3036691
s&p500	0.6556607	−0.1257455	0.8310044	0.1690855

Table 3. Maximum likelihood estimates of the proposed model for RRC data set.

Parameter	Starting Values	$EM(tol=10^{-4})$	$EM(tol=10^{-6})$	$EM(tol=10^{-8})$
$\hat{\alpha}$	0.3722511	0.6090961	0.60882855	0.6088255
$\hat{\beta}$	−0.02456226	−0.03574201	−0.03559926	−0.03559527
$\hat{\delta}$	2.950864	1.33822	1.33576	1.335739
$\hat{\mu}$	0.4284473	0.5176707	0.516677	0.5166471
Log-likelihood		−1697.415	−1697.81	−1697.814
No. iteration		32	50	67
AIC		3402.83	3403.62	3403.628

at different levels. The log-likelihood and AIC for each data sets are also provided.

Figures 4-6 show how the proposed models fit the data sets. It is clear that the proposed model is a good fit compared to the normal distribution and hence a good alternative to the Normal Inverse Gaussian distribution.

Remark:

Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\gamma^2}{\gamma^2 + 1} \times GHD\left(\frac{1}{2}, \alpha, \delta, \beta, \mu\right) + \frac{1}{\gamma^2 + 1} \times GHD\left(\frac{3}{2}, \alpha, \delta, \beta, \mu\right) \quad (32)$$

The finite mixture therefore is flexible in determining between $GHD\left(\frac{1}{2}, \alpha, \delta, \beta, \mu\right)$ and $GHD\left(\frac{3}{2}, \alpha, \delta, \beta, \mu\right)$ depending on the nature of the data. Table 6 below gives different values of weights for the three data sets considered based on the maximum likelihood estimates.

Table 4. Maximum likelihood estimates of the proposed model for CVX data set.

Parameter	Starting Values	$EM(tol = 10^{-4})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.4190067	1.368124	1.394767	1.395091
$\hat{\beta}$	−0.1054991	−0.2778858	−0.297205	−0.2974406
$\hat{\delta}$	0.8324058	1.452515	1.480648	1.480994
$\hat{\mu}$	0.3036691	0.6180421	0.6452299	0.6455598
Log-likelihood		−1224.464	−1223.515	−1223.504
No. iteration		30	77	126
AIC		2456.928	2455.030	2455.008

Table 5. Maximum likelihood estimates of the proposed model for s&p500 index.

Parameter	Starting Values	$EM(tol = 10^{-4})$	$EM(tol = 10^{-5})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.6556607	2.172327	2.32012	2.727979
$\hat{\beta}$	−0.1257455	−0.2766788	−0.3220001	−0.4674433
$\hat{\delta}$	0.8310044	2.026605	2.190487	2.627292
$\hat{\mu}$	0.1690855	0.3394356	0.3866598	0.5386185
Log-likelihood		−1061.985	−1061.673	−1063.643
No. iteration		31	43	510
AIC		2131.970	2131.346	2135.286

Table 6. Estimates for proposition p of the data sets.

dataset	$\hat{p}$
RRC	0.26975
CVX	0.65008
s&p500	0.878395

Proposed Model & Normal
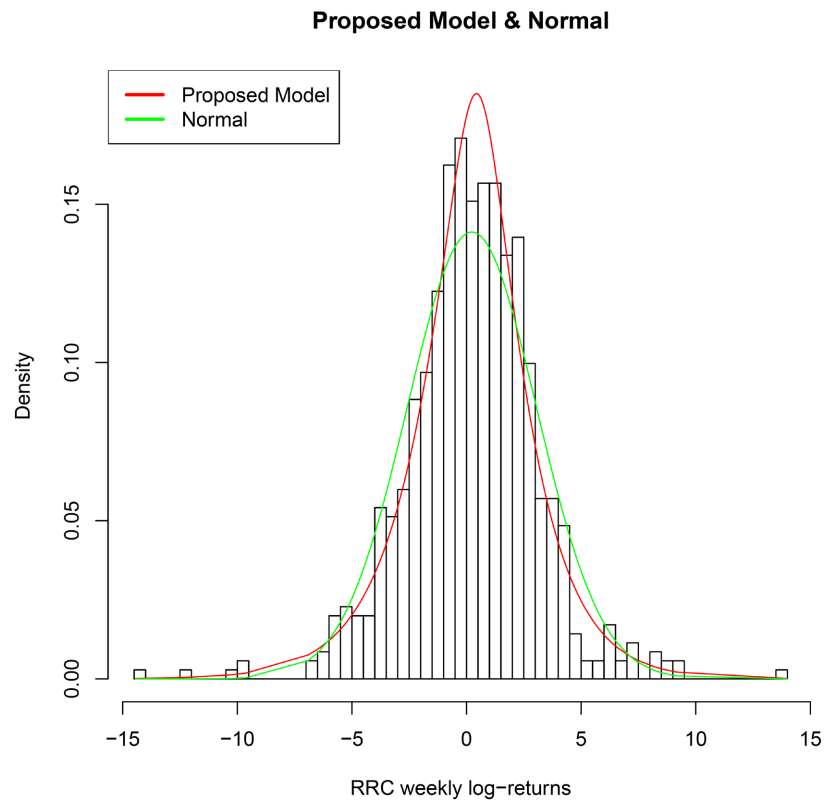
Figure 4. Fitting Proposed Model to RRC weekly returns.

Proposed Model & Normal
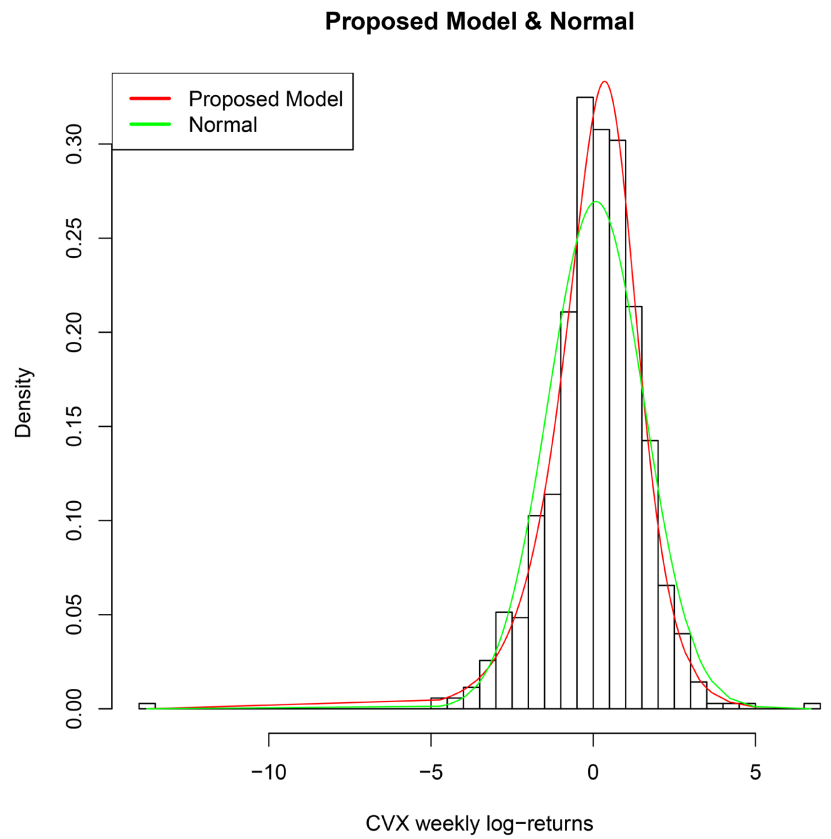
Figure 5. Fitting proposed model to CVX weekly returns.
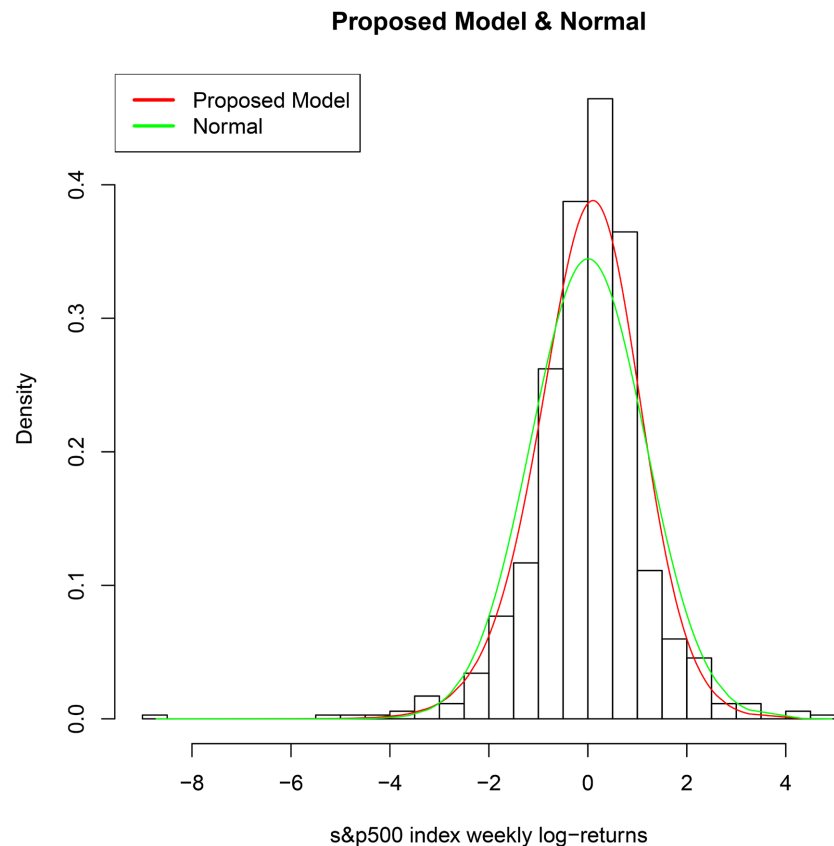
Proposed Model & Normal

Figure 6. Fitting proposed model to s&p500 index weekly returns.

6. Conclusions

Two special cases of the Generalized Inverse Gaussian with indexes $\frac{1}{2}$ and $\frac{3}{2}$ have been used to construct a finite mixture model. The model has been used as a mixing distribution for Normal Variance-Mean mixture to a Normal Weighted Inverse Gaussian Model. The mean and variance of the proposed model have been obtained.

Three data sets: Range Resource Corporation (RRC), Shares of Chevron Corporation (CVX) and s&p500 index for the period 3/01/2000 to 1/07/2013 with 702 observations have been used for data analysis. An iterative scheme has been presented for parameter estimation by the EM algorithm. The iterative scheme demonstrates a monotonic convergence. The method of moment estimates for NIG worked well for the three data sets. The model fits the data sets well. Hence the model is a good alternative to NIG.

Acknowledgements

The authors gratefully acknowledge the financial support from Kisii University.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Barndorff-Nielsen, O. E. (1997) Normal Inverse Gaussian Distribution and Stochastic Volatility Modelling. *Scandinavian Journal of Statistics*, **24**, 1-13. https://doi.org/10.1111/1467-9469.00045

[2] Aas, K. and Haff, H. (2005) NIG and Skew Student's t: Two Special Cases of the Generalised Hyperbolic. Norwegian Computing Center, Oslo, Norway.

[3] Aas, K. and Haff, H. (2006) The Generalised Hyperbolic Skew Student's t-Distribution. *Journal of Financial Econometrics*, **4**, 275-309. https://doi.org/10.1093/jjfinec/nbj006

[4] Corlu, C.G. and Corlu, A. (2014) Modeling Exchange Rate Returns: Which Flexible Distributions to Use? *Quantitative Finance*, **15**, 1851-1864. https://doi.org/10.1080/14697688.2014.942231

[5] Nadarajah, S., Zhang, B. and Chan, S. (2014) Estimation Methods for Expected Shortfall. *Quantitative Finance*, **14**, 271-291. https://doi.org/10.1080/14697688.2013.816767

[6] Lindley, D.V. (1958) Fiducial Distributions and Bayes Theorem. *Journal of the Royal Statistical Society*: *Series B*, **20**, 102-107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x

[7] Sankaran, M. (1970) The Discrete Poisson-Lindley Distribution. *Biometrics*, **26**, 145-149. https://doi.org/10.2307/2529053

[8] Mahmoudi, E. and Zakerzadeh, H. (2010). Generalized Poisson-Lindley Distribution. *Communications in Statistics*: *Theory and Methods*, **39**, 1785-1798. https://doi.org/10.1080/03610920902898514

[9] Fisher, R.A. (1934) The Effect of Methods of Ascertainment upon the Estimation of Frequencies. *The Annals of Human Genetics*, **6**, 13-25. https://doi.org/10.1111/j.1469-1809.1934.tb02105.x

[10] Patil, G.P. and Rao, C.R. (1978) Weighted Distributions and Size-Biased Sampling with Applications to Wildlife Populations and Human Families. *Biometrics*, **34**, 179-189. https://doi.org/10.2307/2530008

[11] Gupta, R.C. and Kundu, D. (2011) Weighted Inverse Gaussian—A Versatile Lifetime Model. *Journal of Applied Statistics*, **38**, 2695-2708. https://doi.org/10.1080/02664763.2011.567251

[12] Abramowitz, M. and Stegun, I.A. (1972) Handbook of Mathematical Function. Dover, New York.

[13] Dempster, A.P., Laird, N.M. and Rubin, D. (1977) Maximum Likelihood from Incomplete Data via the *EM* Algorithm. *Journal of the Royal Statistical Society*: *Series B*, **39**, 1-38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

[14] Karlis, D. (2002) An EM Type Algorithm for Maximum Likelihood Estimation of the Normal-Inverse Gaussian Distribution. *Statistics and Probability Letters*, **57**, 43-52. https://doi.org/10.1016/S0167-7152(02)00040-8