Journal of Mathematical Finance, 2022, 12, 20-45
https://www.scirp.org/journal/jmf
ISSN Online: 2162-2442
ISSN Print: 2162-2434

Value at Risk and Expected Shortfall for Normal Weighted Inverse Gaussian Distributions

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, Nairob, Kenya
Email: *calvinmaina@kisiiuniversity.ac.ke

How to cite this paper: Maina, C.B., Weke, P.G.O., Ogutu, C.A. and Ottieno, J.A.M. (2022) Value at Risk and Expected Shortfall for Normal Weighted Inverse Gaussian Distributions. *Journal of Mathematical Finance*, **12**, 20-45.
https://doi.org/10.4236/jmf.2022.121002

Received: June 9, 2021
Accepted: January 8, 2022
Published: January 11, 2022

Abstract

Value at Risk (VaR) and Expected Shortfall (ES) is commonly used measures of potential risk for losses in financial markets. In literature VaR and ES for the Normal Inverse Gaussian (NIG) distribution, a special case of Generalized Hyperbolic Distribution (GHD), is frequently used. There are however, Normal Inverse Gaussian related distributions, which are also special cases of GHD that can also be used. The objective of this paper is to calculate VaR for Normal Weighted Inverse Gaussian (NWIG) distributions. The Expectation-Maximization (EM) algorithm has been used to obtain the Maximum Likelihood (ML) estimates of the proposed models for the Range Resource Corporation (RRC) financial data. We used Kupiec likelihood ratio (LR) for backtesting of VaR. Kolmogorov-Smirnov test and Anderson-Darling test have been used for goodness of fit test. Akaike Information Creterion (AIC), Bayesian Information Creterion (BIC) and Log-likelihood have been used for model selection. The results clearly show that the NWIG distributions are good alternatives to NIG for determining VaR and ES.

Keywords

Risk Measures, Backtesting, Weighted Distribution, Normal Mixture, EM-Algorithm

1. Introduction

The most popular measures for financial risk are Value at Risk (VaR) and Expected Shortfall (ES). VaR was proposed by Till Guldimann in the late 1980s while working for J. P. Morgan as the head of global research. It is generally defined as possible maximum loss over a given holding period within a fixed confidence level. An attractive feature of VaR is the backtestability of the measure.

Backtesting is a method that uses historical data to gauge accuracy and effectiveness (Zhang and Nadarajah [1]). Backtesting VaR is used to compare the forecast/predicted losses from the actual calculated losses realised at the end of a fixed time horizon. However, the main shortcoming of VaR is that it ignores any loss beyond the value at risk level. That is, it fails to capture tail risk. It also lacks a mathematical property called subadditivity as stated by Wimmerstedt [2]. That is, VaR for two combined portfolios can be larger than VaR for the sum of the two portfolios independently. This implies that diversification could increase risk, a contradiction to standard beliefs in finance. Artzner, Delbaen, Eber and Heath ([3] [4]) have proposed the use of Expected Shortfall (ES) also called conditional Value at Risk (CVaR) to circumvent the problems inherent in VaR. Expected Shortfall is the conditional expectation of loss given that the loss is beyond the VaR level. However, the main problem with ES is that it lacks a mathematical property called elicitability, Gneting [5], necessary for the backtestability of the risk measure. The Basel Committee [6] proposed to replace Value at Risk with Expected Shortfall but concluded that the backtesting will still be done on VaR even though the capital would be based on Expected Shortfall. Therefore the two measures of risk still remain the most popular and useful in financial management.

Notebly, Nadarajah *et al.* [7] have given a detailed review of VaR and ES for various distributions. One of the distributions reviewed is the Generalized Hyperbolic Distribution (GHD) introduced by Barndorff-Nielsen [8] as a Normal Variance-Mean Mixture with the Generalized Inverse Gaussian (GIG) distribution as the mixing distribution. The GIG is a three-parameter distribution denoted as $GIG(\lambda,\delta,\gamma)$. It embraces a number of special and limiting cases. The GHD and its subclasses fit high frequency financial data well which are characterised by skewness, excess kurtosis and fat tail. The most common special case is Normal Inverse Gaussian (NIG) distribution introduced by Barndorff-Nielsen [9] with the Inverse Gaussian (IG) as the mixing distribution. However, there are other special cases of GHD which are related to the NIG distribution which have not been considered for VaR an ES. These special cases are Normal Weighted Inverse Gaussian Distributions.

The objective of this paper is to determine VaR and ES for Range Resource Corporation (RRC) financial data using Normal Weighted Inverse Gaussian (NWIG) distributions. In particular we consider Normal mixtures with $GIG\left(\frac{1}{2},\delta,\gamma\right)$, $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ and $GIG\left(\frac{3}{2},\delta,\gamma\right)$ as mixing distribution. We study their properties and estimate parameters using the Expectation Maximization algorithm introduced by Dempster *et al.* [10]. Kolmogorov-Smirnov test and Anderson-Darling test have been used for goodness of fit test. Akaike Information Creterion (AIC), Bayesian Information Creterion (BIC) and Log-likelihood have been used for model selection.

The concept of a weighted distribution was introduced by Fisher [11] and elaborated by Patil and Rao [12]. Reciprocal Inverse Gaussian and the finite mixture of Inverse Gaussian and Reciprocal Inverse Gaussian distribution are shown to be Weighted Inverse Gaussian (WIG) distributions by Akman and Gupta [13]; Gupta and Akman [14]; Gupta and Kundu [15]. The Special cases $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ and $GIG\left(\frac{3}{2},\delta,\gamma\right)$ are also WIG distributions.

For value at Risk of these models we use the Kupiec likelihood ratio (LR) introduced by Kupiec [16]. The remainder of this paper is structured as follows: Section 2 deals with the concept of VaR and ES. Section 3 deals with Generalised Inverse Gaussian and its special cases. Weighted distribution is covered in Section 4. The concept on Generalised hyperbolic distribution is illustrated in Section 5 while its special cases of interest are in Section 6. Parameter estimation of the proposed models is performed in Section 7. Application to Range Resource Corporation is done in Section 8 and Section 8 deals with conclusion.

2. Value at Risk and Expected Shortfall: Mathematical Background

The most important risk measures despite their drawbacks are Value at Risk (VaR) and Expected Shortfall (ES). VaR was proposed by Till Guldimann in the late 1980s, and at the time he was the head of global research at J. P. Morgan.

Value at Risk is generally defined as possible maximum loss over a given holding period within a fixed confidence level. Mathematically VaR at the $(100-\alpha)$ percent confidence level is defined as the lower 100α percentile of the profit-loss distribution.

In statistical terms, VaR is a quantile of distribution for financial asset returns. More formally, VaR is defined as

$$P\left\{X \le -VaR_{1-\alpha}^{X}\right\} = \alpha \tag{1}$$

where X represents the Asset's returns. In integral form it can be expressed as

$$\int_{-\infty}^{VaR_{\alpha}^{X}} f(x)\,\mathrm{d}x = \alpha \tag{2}$$

where $f(x)$ is the profit-loss distribution.

The concept of Expected Shortfall (ES) was first introduced in Rappoport [17]. Artzner *et al.* ([3] [4]) formally developed the concept. ES is the conditional expectation of loss given that the loss is beyond the VaR level and measures how much one can lose on average in the states beyond the VaR level.

From Equation (2.2)

$$\frac{1}{\alpha}\int_{-\infty}^{VaR_{\alpha}} f(x)\,\mathrm{d}x = 1 \tag{3}$$

Therefore $\frac{f(x)}{\alpha}$ is a pdf for $-\infty < x < VaR_{\alpha}$ and we refer to it as "Tail loss distribution".

Conditional Expectation

$$E\left[X \mid X < VaR_{\alpha}\right] = \int_{-\infty}^{VaR_{\alpha}} x\frac{f(x)}{\alpha}\mathrm{d}x \tag{4}$$

is the Expected Shortfall denoted as ES_α. This version was used by Yamai and Yoshiba [18] to obtain the ES for a normal distribution. Equation (2.4) can be expressed in a different version as follows: Defining $F(x)$ as the cdf of the random variable X, let

$$u = F(x) \Rightarrow x = F^{-1}(u)$$

$$\therefore \mathrm{d}u = f(x)\mathrm{d}x$$

when

$$x = -\infty \Rightarrow u = 0$$

$$x = VaR_\alpha \Rightarrow u = \alpha$$

$$\therefore ES_\alpha = \frac{1}{\alpha}\int_0^\alpha F^{-1}(u)\mathrm{d}u = \frac{1}{\alpha}\int_0^\alpha VaR_u \mathrm{d}u \tag{5}$$

as presented by Zhang *et al.* [19].

Remarks: Equation (2.3) is the mean of the loss distribution. Equation (2.4) represents the average of the VaR between 0 and α. The loss distribution, $\frac{f(x)}{\alpha}$, $-\infty < x < VaR_\alpha$ gives the tail distribution.

For the purpose of VaR and ES analysis, a model for the return distribution is important because it describes the potential behaviour of a financial security in the future Bams and Wilhouwer [20]. A Normal distribution supposedly underestimates the tail and hence VaR. Recently alternative distributions have been proposed that focus more on tail behaviour of the returns. One such candidate is the Normal Inverse Gaussian (NIG) distribution. We consider extensions of NIG distribution as Normal Weighted Inverse Gaussian (NWIG) distributions. In the next few sections we give a detailed illustration on their construction, properties and parameter estimation via EM-algorithm.

3. Generalised Inverse Gaussian Distribution

The Generalised Inverse Gaussian (GIG) Distribution is based on modified Bessel function of the third kind. Modified Bessel function of the third kind of order λ evaluated at ω denoted by $K_\lambda(\omega)$ is defined as

$$K_\lambda(\omega) = \frac{1}{2}\int_0^\infty x^{\lambda-1} \mathrm{e}^{-\frac{\omega}{2}\left(x+\frac{1}{x}\right)} \mathrm{d}x \tag{6}$$

with the following properties

$$\text{a)}\quad K_{\frac{1}{2}}(\omega) = K_{-\frac{1}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}}\mathrm{e}^{-\omega} \tag{7}$$

$$\text{b)}\quad K_{\frac{3}{2}}(\omega) = K_{-\frac{3}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}}\mathrm{e}^{-\omega}\left(1+\frac{1}{\omega}\right) \tag{8}$$

$$\text{c)}\quad K_{\frac{5}{2}}(\omega) = K_{-\frac{5}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}}\mathrm{e}^{-\omega}\left(1+\frac{3}{\omega}+\frac{3}{\omega^2}\right) \tag{9}$$

d) $$K_{\frac{7}{2}}(\omega) = K_{-\frac{7}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{6}{\omega} + \frac{15}{\omega^2} + \frac{15}{\omega^3}\right) \tag{10}$$

e) $$K_{\frac{9}{2}}(\omega) = K_{-\frac{9}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{10}{\omega} + \frac{45}{\omega^2} + \frac{105}{\omega^3} + \frac{105}{\omega^4}\right) \tag{11}$$

f) $$K_{\frac{11}{2}}(\omega) = K_{-\frac{11}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{15}{\omega} + \frac{105}{\omega^2} + \frac{420}{\omega^3} + \frac{945}{\omega^4} + \frac{945}{\omega^5}\right) \tag{12}$$

which are necessary in deriving the properties and estimates of the proposed models. For more definition and properties see Abramowitz and Stegun [21].

Using Parametrization $\omega = \delta\gamma$ and transformation $x = \frac{\gamma}{\delta} z$ then formula (2.1) becomes

$$K_\lambda(\delta\gamma) = \frac{1}{2}\int_0^\infty \left(\frac{\gamma}{\delta}\right)^\lambda z^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} dz$$

Hence

$$g(z) = \left(\frac{\gamma}{\delta}\right)^\lambda \frac{z^{\lambda-1}}{2K_\lambda(\delta\gamma)} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \tag{13}$$

$$z > 0; -\infty < \lambda < \infty, \delta > 0, \gamma > 0$$

which is a Generalized Inverse Gaussian (GIG) distribution with parameter λ, δ, γ.

Thus

$$Z \sim GIG(\lambda, \delta, \gamma)$$

with

$$E(Z^r) = \left(\frac{\delta}{\gamma}\right)^r \frac{K_{\lambda+r}(\delta\gamma)}{K_\lambda(\delta\gamma)}$$

where r can be positive or negative integers.

4. Weighted Inverse Gaussian Distribution

Let X be a random variable with pdf $f(x)$. A function of X, $w(X)$ is also a random variable with expectation

$$E[w(X)] = \int_{-\infty}^{\infty} w(x) f(x) dx$$

$$\therefore 1 = \int_{-\infty}^{\infty} \frac{w(x)}{E[w(X)]} f(x) dx$$

Thus

$$f_w(x) = \frac{w(x)}{E[w(X)]} f(x), -\infty < x < \infty \tag{14}$$

is a weighted distribution. It was introduced by Fisher [11] and elaborated by Patil and Rao [12].

From Equation (3.8) When $\lambda = -\frac{1}{2}$ we have $GIG\left(-\frac{1}{2}, \delta, \gamma\right)$ with pdf

$$g_1(z) = \frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \tag{15}$$

This is called Inverse Gaussian (IG) distribution.

The following special cases of GIG distribution can be expressed in terms of IG distribution. They are weighted Inverse Gaussian distributions.

Example 1: $GIG\left(\frac{1}{2}, \delta, \gamma\right)$

When $\lambda = \frac{1}{2}$, then the pdf of $GIG(\lambda, \delta, \gamma)$ becomes

$$g_2(z) = \frac{\gamma e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\}$$

This is a Reciprocal Inverse Gaussian (RIG) distribution. It can be written as

$$g_2(z) = \frac{\gamma}{\delta} z \left[\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\}\right]$$

$$\therefore g_2(z) = \left(\frac{\gamma}{\delta} z\right) g_1(z)$$

i.e.,

$$GIG\left(\frac{1}{2}, \delta, \gamma\right) = \left(\frac{\gamma}{\delta} z\right) GIG\left(-\frac{1}{2}, \delta, \gamma\right) \tag{16}$$

Thus a $GIG\left(\frac{1}{2}, \delta, \gamma\right)$ is a weighted Inverse Gaussian distribution with weights

$$w(Z) = Z \tag{17}$$

where

$$Z \sim GIG\left(-\frac{1}{2}, \delta, \gamma\right)$$

$$\therefore E[w(Z)] = E(Z) = \frac{\delta}{\gamma}$$

RIG distribution is called Length Biased Inverse Gaussian distribution.

Example 2: $GIG\left(-\frac{3}{2}, \delta, \gamma\right)$

The pdf is

$$g_3(z) = \frac{\delta^3 e^{\delta\gamma}}{\sqrt{2\pi}(1+\delta\gamma)} z^{-\frac{5}{2}} e^{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)} = \frac{\delta^2}{1+\delta\gamma} z^{-1} g_1(z)$$

Therefore $GIG\left(-\frac{3}{2}, \delta, \gamma\right)$ is a weighted Inverse Gaussian distribution with weights

$$w(Z)=Z^{-1} \tag{18}$$

and

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

Example 3: $GIG\left(\frac{3}{2},\delta,\gamma\right)$

The pdf of $GIG\left(\frac{3}{2},\delta,\gamma\right)$

$$\begin{aligned}g_4(z)&=\frac{\gamma^3}{1+\delta\gamma}\frac{\mathrm{e}^{\delta\gamma}}{\sqrt{2\pi}}z^{\frac{1}{2}}\mathrm{e}^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2z\right)}\\&=\frac{\gamma^3}{\delta(1+\delta\gamma)}z^2\left[\frac{\delta\mathrm{e}^{\delta\gamma}}{\sqrt{2\pi}}z^{-\frac{3}{2}}\exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2z\right)\right\}\right]\\&=\frac{\gamma^3}{\delta(1+\delta\gamma)}z^2g_1(z)\end{aligned}$$

Thus, $GIG\left(\frac{3}{2},\delta,\gamma\right)$ is a weighted Inverse Gaussian distribution with weights

$$w(Z)=Z^2 \tag{19}$$

where

$$Z\sim GIG\left(-\frac{1}{2},\delta,\gamma\right)$$

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

5. Generalized Hyperbolic Distribution

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

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

where

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

and Z, independent of Y, is a positive random variable.

If $F(x)$ is a cdf of X, then

$$\begin{aligned}F(x)&=prob\{X\le x\}\\&=\left\{Y\le\frac{x-\mu-\beta z}{\sqrt{z}},0<z<\infty\right\}\\&=\int_0^\infty\int_{-\infty}^{\frac{x-\mu-\beta z}{\sqrt{z}}}\phi(y)g(z)\mathrm{d}y\mathrm{d}z\\&=\int_0^\infty\Phi\left(\frac{x-\mu-\beta z}{\sqrt{z}}\right)g(z)\mathrm{d}z\end{aligned}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are pdf and cdf of a standard normal distribution, respectively.

$$\therefore f(x)=\int_0^{\infty}\frac{1}{\sqrt{z}}\phi\left(\frac{x-\mu-\beta z}{\sqrt{z}}\right)g(z)\mathrm{d}z=\int_0^{\infty}\frac{1}{\sqrt{2\pi z}}\mathrm{e}^{-\frac{[x-(\mu+\beta z)]^2}{2z}}g(z)\mathrm{d}z \quad (20)$$

Thus we have a hierarchical representation as

$$X/Z=z\sim N(\mu+\beta z,z) \quad (21)$$

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

If

$$Z\sim GIG(\lambda,\delta,\gamma) \quad (22)$$

then

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

where

$$\alpha^2=\gamma^2+\beta^2\Rightarrow\gamma=\sqrt{\alpha^2-\beta^2}$$

and

$$\phi(x)=1+\frac{(x-\mu)^2}{\delta^2}$$

This is an integral form of a GHD for $-\infty<x<\infty;-\infty<\mu<\infty,\delta>0,\alpha>0,\beta>0$.

Using the stochastic representation of NVM mixture, the properties are expressed in terms of the mixing distribution as shown in the following

Proposition 1

$$M_X(t)=\mathrm{e}^{\mu t}M_Z\left(\beta t+\frac{t^2}{2}\right)$$

$$E(X)=\mu+\beta E(Z)$$

$$Var(X)=E(Z)+\beta^2 Var(Z)$$

$$\mu_3(X)=3\beta var(Z)+\beta^3\mu_3(Z)$$

$$\mu_4(X)=\beta^4\mu_4(Z)+6\beta^2\mu_3(Z)+6\beta^2 E[Z]var(Z)+3E\left[Z^2\right]$$

6. Special Cases of Interest

When $\lambda=-\frac{1}{2}$, the GHD becomes NIG whose pdf can be expressed in the integral form as

$$f_1(x)=\frac{\delta\mathrm{e}^{\delta\gamma}\mathrm{e}^{\beta(x-\mu)}}{2\pi}\int_0^{\infty}z^{-2}\mathrm{e}^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}\mathrm{d}z \quad (23)$$

Therefore the Normal Weighted Inverse Gaussian formulation can be represented as

$$\therefore f(x)=\frac{\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{2\pi}\int_0^\infty \frac{w(z)}{E[w(Z)]}z^{-2}e^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}dz \tag{24}$$

Hence

$$\lambda=-\frac{1}{2}\Rightarrow\frac{w(z)}{E(w(Z))}=1 \tag{25}$$

$$\therefore f_1(x)=\frac{\alpha\delta e^{\delta\gamma+\beta(x-\mu)}K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}}$$

which is the Normal Inverse Gaussian (NIG) distribution with properties given in Table 1 below.
when

$$\lambda=\frac{1}{2}\Rightarrow\frac{w(z)}{E[w(Z)]}=\frac{\gamma}{\delta}z \tag{26}$$

$$f_2(x)=\frac{\gamma e^{\delta\gamma+\beta(x-\mu)}K_0\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi}$$

Which is the Normal Reciprocal Inverse Gaussian (NRIG) distribution with properties given in Table 2 below.

Similarly, when

$$\lambda=-\frac{3}{2}\Rightarrow\frac{w(z)}{E[w(Z)]}=\frac{\delta^2}{(1+\delta\gamma)z} \tag{27}$$

$$\therefore f_3(x)=\frac{\alpha^2\delta^3 e^{\delta\gamma+\beta(x-\mu)}K_2\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi(1+\delta\gamma)\left[\delta^2+(x-\mu)^2\right]}$$

which is the Normal-$GIG\left(\lambda=-\frac{3}{2},\delta,\gamma\right)$ with properties given in Table 3 below.

Table 1. Properties of NIG.

Item	Description	Expression
1	$E(X)$	$\mu+\beta\frac{\delta}{\gamma}$
2	$var(X)$	$\frac{\alpha^2\delta}{\gamma^3}$
3	Skewness, γ_1	$\frac{3\beta}{\alpha(\delta\gamma)^{\frac{1}{2}}}$
4	Excess Kurtosis, γ_2	$\frac{3\left(1+4\frac{\beta^2}{\alpha^2}\right)}{\delta\gamma}$

Table 2. Properties of NRIG.

Item	Description	Expression
1	$E(X)$	$\mu+\frac{\beta(1+\delta\gamma)}{\gamma^2}$
2	$var(X)$	$\frac{\alpha^2(1+\delta\gamma)+\beta^2}{\gamma^4}$
3	Skewness, γ_1	$\frac{3\beta\alpha^2(\delta\gamma+2)+2\beta^3}{\left(\alpha^2(1+\delta\gamma)+\beta^2\right)^{\frac{3}{2}}}$
4	Excess Kurtosis, γ_2	$\frac{(2+\delta\gamma)+\frac{\beta^2}{\alpha^2}(14+4\delta\gamma)}{\left(\beta^2+(1+\delta\gamma)\right)^2}$

Table 3. Properties for Normal- $GIG\left(-\frac{3}{2},\delta,\gamma\right)$.

Item	Description	Expression
1	$E(X)$	$\mu+\frac{\beta\delta^2}{1+\delta\gamma}$
2	$var(X)$	$\frac{\delta^2(\gamma+\alpha^2\delta)}{\gamma(1+\delta\gamma)^2}$
3	Skewness, γ_1	$\frac{3\beta\delta^3}{\sqrt{\gamma^2+\alpha^2\delta\gamma}}+\left(\frac{\beta\delta}{\sqrt{\gamma^2+\alpha^2\delta\gamma}}\right)^3$
4	Excess Kurtosis, γ_2	$\frac{3\left[\alpha^4(1+\delta\gamma)^2+2\delta\gamma\left(\alpha^4(1+2\delta\gamma)-\gamma^3\left(\gamma+2\alpha^2\delta\right)\right)\right]}{\delta\gamma^3\left(\gamma+\alpha^2\delta\right)^2}$

Finally, when

$$\lambda=\frac{3}{2}\Rightarrow\frac{w(z)}{E\left[w(Z)\right]}=\frac{\gamma^3z^2}{\delta(1+\delta\gamma)} \tag{28}$$

$$\therefore f_4(x)=\frac{\gamma^3\mathrm{e}^{\delta\gamma+\beta(x-\mu)}\left[\sqrt{\delta^2+(x-\mu)^2}\right]K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\alpha\pi(1+\delta\gamma)}$$

which is the Normal- $GIG\left(\lambda=\frac{3}{2},\delta,\gamma\right)$ with properties given in Table 4 below

7. Parameter Estimation via EM Algorithm

7.1. Theory/Concept

EM algorithm is a powerful technique for maximum likelihood estimation for data containing missing values or data that can be considered as containing missing values. It was introduced by Dempster *et al.* [10].

Table 4. Properties for Normal- $GIG\left(\frac{3}{2},\delta,\gamma\right)$.

Item	Description	Expression
1	$E(X)$	$\mu+\frac{\beta\left(\delta^2\gamma^2+3\delta\gamma+3\right)}{\gamma^2(1+\delta\gamma)}$
2	$var(X)$	$\frac{3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2}{\gamma^4(1+\delta\gamma)^2}$
3	Skewness, γ_1	$\beta\frac{6(1+\delta\gamma)^3\left(\beta^2+3\alpha^2\right)+\delta^3\gamma^3\left(3\alpha(1+\delta\gamma)-2\beta^2\right)}{\left(3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2\right)^{\frac{3}{2}}}$

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, Kostas [22].

This implies that the joint density of X and Z is given by

$$f(x,z)=f(x/z)g(z)$$

The likelihood function is

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

$$\begin{aligned}\therefore \log L &= \log\prod_{i=1}^{n} f(x_i/z_i)+\log\prod_{i=1}^{n} g(z_i)\\ &=\sum_{i=1}^{n}\log f(x_i/z_i)+\sum_{i=1}^{n}\log g(z_i)\\ &= l_1+l_2\end{aligned}$$

where

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

and

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

Karlis [23] applied EM algorithm to mixtures which he considered to consist of two parts; the conditional pdf is for observed data and the mixing distribution is based on an unobserved data, the missing values.

7.2. M-Step for Conditional pdf

In this study

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

Therefore

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

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

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

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

Similarly,

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

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

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

7.3. E-Step

Values of random variables Z_i and $\frac{1}{Z_i}$ are not known. So we estiamte them by considering posterior expectations $E\left(Z_i/X_i\right)$ and $E\left(\frac{1}{Z_i}\right)$.

The posterior distribution is defined as

$$f\left(z/x\right) = \frac{f\left(x/z\right)g\left(z\right)}{\int_0^\infty f\left(x/z\right)g\left(z\right)\mathrm{d}z}$$

Therefore

$$E\left(Z/X\right) = \frac{\int_0^\infty z f\left(x/z\right)g\left(z\right)\mathrm{d}z}{\int_0^\infty f\left(x/z\right)g\left(z\right)\mathrm{d}z}$$

$$E\left(Z/X\right) = \frac{\int_0^\infty \frac{1}{z} f\left(x/z\right)g\left(z\right)\mathrm{d}z}{\int_0^\infty f\left(x/z\right)g\left(z\right)\mathrm{d}z}$$

7.4. For Inverse Gaussian Mixing Distribution

7.4.1. M-Step

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

$$\therefore \frac{\partial}{\partial\gamma} l_2 = n\delta - \gamma\sum_{i=1}^{n} z_i$$

$$\therefore \frac{\partial}{\partial\gamma} l_2 = 0 \Rightarrow \hat{\gamma} = \frac{\hat{\delta}}{\overline{z}}$$

and

$$\frac{\partial}{\partial \delta} l_2 = 0 \Rightarrow \frac{n}{\delta} + n\gamma - \delta \sum_{i=1}^{n} \frac{1}{z_i} = 0$$

$$\therefore \hat{\delta} = \sqrt{\frac{n}{\sum_{i=1}^{n} \frac{1}{z_i} - \frac{n}{\bar{z}}}}$$

7.4.2. Posterior Expectations

For NIG

$$E(Z/X) = \frac{\delta\sqrt{\phi(x)}}{\alpha} \frac{K_0\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

and

$$E\left(\frac{1}{Z}\middle/X\right) = \frac{\alpha}{\delta\sqrt{\phi(x)}} \frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

7.4.3. Iterations

Let

$$s_i = E(Z_i/X_i) = \frac{\delta\sqrt{\phi(x_i)}}{\alpha} \frac{K_0\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}$$

where

$$\phi(x_i) = 1 + \frac{(x_i - \mu)^2}{\delta^2}$$

and

$$\alpha = \sqrt{\gamma^2 + \beta^2}$$

Therefore s_i is a function of $\alpha, \beta, \delta, \gamma, \mu$

Next, let

$$w_i = E\left(\frac{1}{Z_i}\middle/X_i\right) = \frac{\alpha}{\delta\sqrt{\phi(x_i)}} \frac{K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}$$

which is also a function of $\alpha, \beta, \delta, \gamma, \mu$

For computation, we have the following in the k-th iteration

$$s_i^{(k)} = \frac{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}{\alpha^{(k)}} \frac{K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \tag{29}$$

$$w_i^{(k)} = \frac{\alpha^{(k)}}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}} \frac{K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \tag{30}$$

$$\hat{\delta}^{(k+1)} = \sqrt{\frac{n}{\sum_{i=1}^{n} w_i^{(k)} - \frac{n^2}{\sum_{i=1}^{n} s_i^{(k)}}}} \tag{31}$$

$$\hat{\gamma}^{(k+1)} = \frac{\hat{\delta}^{(k+1)}}{\frac{1}{n}\sum_{i=1}^{n} s_i^{(k)}} \tag{32}$$

The log-likelihood function of NIG distribution is given by

$$\begin{aligned} \log L &= \sum_{i=1}^{n} \log \left\{ \frac{\alpha\delta e^{\delta\gamma+\beta(x-\mu)} K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}} \right\} \\ &= n\log\alpha + n(\delta\gamma - \beta\mu) + \beta\sum_{i=1}^{n} x_i - n\log\pi \\ &\quad -\frac{1}{2}\sum_{i=1}^{n}\phi(x_i) + \sum_{i=1}^{n}\log K_1\left(\alpha\delta\phi(x_i)^{\frac{1}{2}}\right) \end{aligned} \tag{33}$$

The k-th iteration of the loglikelihood function of the NIG distribution is

$$\begin{aligned} l^{(k)} &= n\log\alpha^{(k)} + n\left(\delta^{(k)}\gamma^{(k)} - \beta^{(k)}\mu^{(k)}\right) + \beta^{(k)}\sum_{i=1}^{n} x_i - n\log\pi \\ &\quad -\frac{1}{2}\sum_{i=1}^{n}\phi^{(k)}(x_i) + \sum_{i=1}^{n}\log K_1\left(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}}\right) \end{aligned} \tag{34}$$

7.5. For Length Biased (Reciprocal) Inverse Gaussian Distribution

7.5.1. M-Step

$$l_2(\delta,\gamma) = n\log\gamma - \frac{n}{2}\log(2\pi) + n\delta\gamma - \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} - \frac{1}{2}\sum_{i=1}^{n}\log z_i$$

$$\therefore \frac{\partial}{\partial\delta} l_2 = 0 \Rightarrow \hat{\delta} = \frac{n\hat{\gamma}}{\sum_{i=1}^{n}\frac{1}{z_i}}$$

$$\frac{\partial}{\partial\gamma} l_2 = 0 \Rightarrow \hat{\gamma} = \sqrt{\frac{n}{\sum_{i=1}^{n} z_i - n^2\left(\sum_{i=1}^{n}\frac{1}{z_i}\right)^{-1}}}$$

7.5.2. Posterior Expectation

$$E(Z/X) = \frac{\delta\sqrt{\phi(x)}}{\alpha}\frac{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_0\left(\alpha\delta\sqrt{\phi(x)}\right)} \tag{35}$$

and

$$E\left(\frac{1}{Z}\Big/X\right) = \frac{\alpha}{\delta\sqrt{\phi(x)}}\frac{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_0\left(\alpha\delta\sqrt{\phi(x)}\right)} \tag{36}$$

7.5.3. Iterations

Let

$$s_i = E\left(Z_i / X_i\right) = \frac{\delta\sqrt{\phi(x_i)}}{\alpha}\frac{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_0\left(\alpha\delta\sqrt{\phi(x_i)}\right)}$$

and

$$w_i = E\left(\frac{1}{Z_i}\middle/X_i\right) = \frac{\alpha}{\delta\sqrt{\phi(x_i)}}\frac{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_0\left(\alpha\delta\sqrt{\phi(x_i)}\right)}$$

For iterations we have

$$s_i^{(k)} = \frac{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}{\alpha^{(k)}}\frac{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \tag{37}$$

$$w_i^{(k)} = E\left(\frac{1}{Z_i}\middle/X_i\right) = \frac{\alpha^{(k)}}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}\frac{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \tag{38}$$

$$\hat{\gamma}^{(k+1)} = \sqrt{\frac{n}{\sum_{i=1}^n s_i^{(k)} - \frac{n^2}{\sum_{i=1}^n w_i^{(k)}}}} \tag{39}$$

$$\hat{\delta}^{(k+1)} = \frac{\gamma^{(k+1)}}{\frac{1}{n}\sum_{i=1}^n w_i^{(k)}} \tag{40}$$

The loglikelihood function of NRIG distribution is given by

$$\log L = n\log\gamma - n\log\pi + n(\delta\gamma - \beta\mu) + \beta\sum_{i=1}^n x_i + \sum_{i=1}^n \log K_0\left(\alpha\delta\phi(x_i)^{\frac{1}{2}}\right) \tag{41}$$

The k-th iteration becomes

$$\begin{aligned} l^{(k)} &= n\log\gamma^{(k)} - n\log\pi + n\left(\delta^{(k)}\gamma^{(k)} - \beta^{(k)}\mu^{(k)}\right) + \beta^{(k)}\sum_{i=1}^n x_i \\ &\quad + \sum_{i=1}^n \log K_0\left(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}}\right) \end{aligned} \tag{42}$$

7.6. Special Case When the Index Parameter Is −3/2

$$l_2 = 3n\log\delta - \frac{n}{2}\log(2\pi) - \frac{5}{2}\sum_{i=1}^n \log z_i - n\log(1+\delta\gamma) + n\delta\gamma - \frac{1}{2}\sum_{i=1}^n\left(\frac{\delta^2}{z_i} + \gamma^2 z_i\right)$$

$$\therefore \frac{\partial}{\partial\gamma} l_2 = 0 \Rightarrow \hat{\gamma} = \frac{\delta^2 - \overline{z}}{\delta\overline{z}}$$

$$\frac{\partial}{\partial\delta} l_2 = 0 \Rightarrow \frac{3n}{\delta} - n\left(\frac{n\delta^2 - \sum_{i=1}^n z_i}{\delta\sum_{i=1}^n z_i}\right)(1+\delta\gamma)^{-1} + n\left(\frac{n\delta^2 - \sum_{i=1}^n z_i}{\delta\sum_{i=1}^n z_i}\right) - \delta\sum_{i=1}^n \frac{1}{z_i} = 0$$

$$\therefore \delta^4\left(n^2 - \sum_{i=1}^n z_i \sum_{i=1}^n \frac{1}{z_i}\right) + n\delta^2\sum_{i=1}^n z_i + \left(\sum_{i=1}^n z_i\right)^2 = 0$$

$$\therefore\left(n^2-\sum_{i=1}^{n} z_i \sum_{i=1}^{n} \frac{1}{z_i}\right)t^2+\left(n\sum_{i=1}^{n} z_i\right)t+\left(\sum_{i=1}^{n} z_i\right)^2=0$$

i.e., $\left[\sum_{i=1}^{n}\frac{1}{z_i}-\frac{n}{\overline{z}}\right]t^2-nt-n\overline{z}=0$

where

$$t=\delta^2$$

Therefore

$$\hat{\delta}=\left[\frac{n\sum_{i=1}^{n} z_i+\sum_{i=1}^{n} z_i\left(4\sum_{i=1}^{n} z_i \sum_{i=1}^{n}\frac{1}{z_i}-3n^2\right)^{\frac{1}{2}}}{2\left[\sum_{i=1}^{n} z_i \sum_{i=1}^{n}\frac{1}{z_i}-n^2\right]}\right]^{\frac{1}{2}}$$

7.6.1. Posterior Expectation

$$E(Z/X)=\frac{\delta\sqrt{\phi(x)}}{\alpha}\frac{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)} \quad (43)$$

and

$$E\left(\frac{1}{Z}\Big/X\right)=\frac{\alpha}{\delta\sqrt{\phi(x)}}\frac{K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)} \quad (44)$$

7.6.2. Iterations

Let

$$s_i=E(Z_i/X_i)=\frac{\delta\sqrt{\phi(x_i)}}{\alpha}\frac{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)} \quad (45)$$

and

$$w_i=E\left(\frac{1}{Z_i}\Big/X_i\right)=\frac{\alpha}{\delta\sqrt{\phi(x_i)}}\frac{K_3\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)} \quad (46)$$

For iterations we have

$$s_i^{(k)}=\frac{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}{\alpha^{(k)}}\frac{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \quad (47)$$

$$w_i^{(k)}=E\left(\frac{1}{Z_i}\Big/X_i\right)=\frac{\alpha^{(k)}}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}\frac{K_3\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \quad (48)$$

$$\hat{\delta}^{(k+1)}=\left[\frac{n\sum_{i=1}^{n} s_i^{(k)}+\sum_{i=1}^{n} s_i^{(k)}\left(4\sum_{i=1}^{n} s_i^{(k)}\sum_{i=1}^{n} w_i^{(k)}-3n^2\right)^{\frac{1}{2}}}{2\left[\sum_{i=1}^{n} s_i^{(k)}\sum_{i=1}^{n} w_i^{(k)}-n^2\right]}\right]^{\frac{1}{2}} \quad (49)$$

$$\hat{\gamma}^{(k+1)} = \frac{\delta^{(k+1)}}{\frac{1}{n}\sum_{i=1}^{n} s_i^{(k)}} - \frac{1}{\delta^{(k+1)}} \tag{50}$$

The log-likelihood function of Normal- $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ is given by

$$\begin{aligned}\log L &= -n\log\pi - n\log(1+\delta\gamma) + 2n\log\alpha + n\log\delta + n(\delta\gamma - \beta\mu) \\ &+ \beta\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\phi(x_i) + \sum_{i=1}^{n}\log K_2\left(\alpha\delta\phi(x_i)^{\frac{1}{2}}\right)\end{aligned} \tag{51}$$

The k-th iteration becomes

$$\begin{aligned}l^{(k)} &= -n\log\pi - n\log\left(1+\delta^{(k)}\gamma^{(k)}\right) + 2n\log\alpha^{(k)} + n\log\delta^{(k)} + n\left(\delta^{(k)}\gamma^{(k)}\right. \\ &\left. - \beta^{(k)}\mu^{(k)}\right) + \beta^{(k)}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\phi^{(k)}(x_i) + \sum_{i=1}^{n}\log K_2\left(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}}\right)\end{aligned} \tag{52}$$

7.7. Special Case When the Index Parameter Is 3/2

7.7.1. M-Step

$$\begin{aligned}l_2 &= \sum_{i=1}^{n}\log g_4(z_i) \\ &= -\frac{n}{2} - n\log(1+\delta\gamma) + 3n\log\gamma + n\delta\gamma + \frac{1}{2}\log\sum_{i=1}^{n}(z_i) \\ &\quad -\frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} - \frac{\gamma^2}{2}\sum_{i=1}^{n} z_i\end{aligned}$$

$$\therefore \frac{\partial}{\partial\delta} l_2 = 0 \Rightarrow \frac{n\gamma^2\delta}{1+\delta\gamma} - \delta\sum_{i=1}^{n}\frac{1}{z_i} = 0$$

$$\therefore \hat{\delta} = \frac{n\hat{\gamma}^2 - \sum_{i=1}^{n}\frac{1}{z_i}}{\hat{\gamma}\sum_{i=1}^{n}\frac{1}{z_i}}$$

Similarly

$$\frac{\partial}{\partial\gamma} l_2 = 0 \Rightarrow \frac{\delta^2\gamma}{1+\delta\gamma} + \frac{3}{\gamma} - \gamma\overline{z} = 0$$

$$\therefore \gamma^2\delta^2 + \left(3\gamma - \gamma^3\overline{z}\right)\delta - \gamma^2\overline{z} + 3 = 0$$

Substituting for δ and letting $\gamma^2 = t$ we obtain

$$\left(n^2 - \sum_{i=1}^{n} z_i\sum_{i=1}^{n}\frac{1}{z_i}\right)t^2 + \left(n\sum_{i=1}^{n}\frac{1}{z_i}\right)t + \left(\sum_{i=1}^{n}\frac{1}{z_i}\right)^2 = 0$$

Therefore

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where

$$a = n^2 - \sum_{i=1}^{n} z_i\sum_{i=1}^{n}\frac{1}{z_i}$$

$$b = n\sum_{i=1}^{n}\frac{1}{z_i}$$

$$c = \left(\sum_{i=1}^{n}\frac{1}{z_i}\right)^2$$

$$\therefore \hat{\gamma} = \left[\frac{-b+\sqrt{b^2-4ac}}{2a}\right]^{\frac{1}{2}}$$

7.7.2. E-Step

Posterior Expectation

$$E(Z/X) = \frac{\delta\sqrt{\phi(x)}}{\alpha}\frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \quad (53)$$

$$E\left(\frac{1}{Z}\Big/X\right) = \frac{\alpha K_0\left(\alpha\delta\sqrt{\phi(x)}\right)}{\delta\sqrt{\phi(x)}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \quad (54)$$

7.8. Iterations

Let

$$s_i = E(Z_i/X_i) = \frac{\delta\sqrt{\phi(x_i)}}{\alpha}\frac{K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)}$$

and

$$w_i = E\left[\frac{1}{Z_i}\Big/X_i\right] = \frac{\alpha K_0\left(\alpha\delta\sqrt{\phi(x_i)}\right)}{\delta\sqrt{\phi(x)}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}$$

For iterations we have

$$s_i^{(k)} = \frac{\delta^{(k)}}{\alpha^{(k)}}\sqrt{\phi^{(k)}(x_i)}\frac{K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \quad (55)$$

$$w_i^{(k)} = \frac{\alpha^{(k)}K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)} \quad (56)$$

$$\hat{\gamma}^{(k+1)} = \sqrt{t^{(k)}} \quad (57)$$

$$\hat{\delta}^{(k+1)} = \frac{n\left(\hat{\gamma}^{(k+1)}\right)^2 - \sum_{i=1}^{n} w_i^{(k)}}{\hat{\gamma}^{(k+1)}\sum_{i=1}^{n} w_i^{(k)}} \quad (58)$$

The log-likelihood function for the Normal- $GIG\left(\frac{3}{2},\delta,\gamma\right)$ distribution is given by

$$\begin{aligned} l = & 3n\log\gamma + n(\delta\gamma - \beta\mu) - n\log(\alpha\pi(1+\delta\gamma)) + n\log\delta \\ & + \beta\sum_{i=1}^{n} x_i + \frac{1}{2}\sum_{i=1}^{n}\log\phi(x_i) + \sum_{i=1}^{n}\log K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \end{aligned} \tag{59}$$

For the k-th iteration we have

$$\begin{aligned} l^{(k)} = & 3n\log\gamma^{(k)} + n\left(\delta^{(k)}\gamma^{(k)} - \beta^{(k)}\mu^{(k)}\right) - n\log\left(\alpha^{(k)}\pi\left(1+\delta^{(k)}\gamma^{(k)}\right)\right) \\ & + n\log\delta^{(k)} + \beta^{(k)}\sum_{i=1}^{n} x_i + \frac{1}{2}\sum_{i=1}^{n}\log\phi^{(k)}(x_i) + \sum_{i=1}^{n}\log K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right) \end{aligned} \tag{60}$$

Remarks:

1) For all the proposed models, the β, μ and α parameters of the conditional distribution are updated as follows:

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})w_i^{(k)}}{n - \frac{1}{n}\sum_{i=1}^{n}s_i^{(k)}\sum_{i=1}^{n}w_i^{(k)}}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \hat{\beta}^{(k+1)}\sum_{i=1}^{n}\frac{s_i^{(k)}}{n}$$

$$\hat{\alpha}^{(k+1)} = \left[\left(\gamma^{(k+1)}\right)^2 + \left(\beta^{(k+1)}\right)^2\right]^{\frac{1}{2}}$$

2) The stopping criterion is when

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

where tol is the tolerance level chosen; e.g 10^{-6}.

3) Initial values used are moment estimates of NIG ditsribution as proposed by Karlis [23].

8. Application

8.1. Fitting of the Proposed Models

The data used in this research is the Range Resource Corporation weekly returns for the period 3/01/2000 to 1/07/2013 with 702 observations. The histogram for the weekly log-returns in **Figure 1** shows 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.

Table 5 provides descriptive statistics for the return series in consideration. We observe that the excess kurtosis of 2.768252 indicates the leptokurtic behaviour of the returns. The log-returns has a distributions with relatively heavier tails than the normal distribution. We observe skewness of −0.1886714 which indicates that the two tails of the returns behave slightly differently.

The proposed models are now fitted to RRC weekly log-returns. Using the sample estimates and the NIG estimators to the RRC data we obtain the following estimates as initial values for the EM algorithm (see Karlis [23]).

$$\hat{\alpha} = 0.3722511, \hat{\beta} = -0.02456226, \hat{\delta} = 2.950864, \hat{\mu} = 0.4284473\,.$$

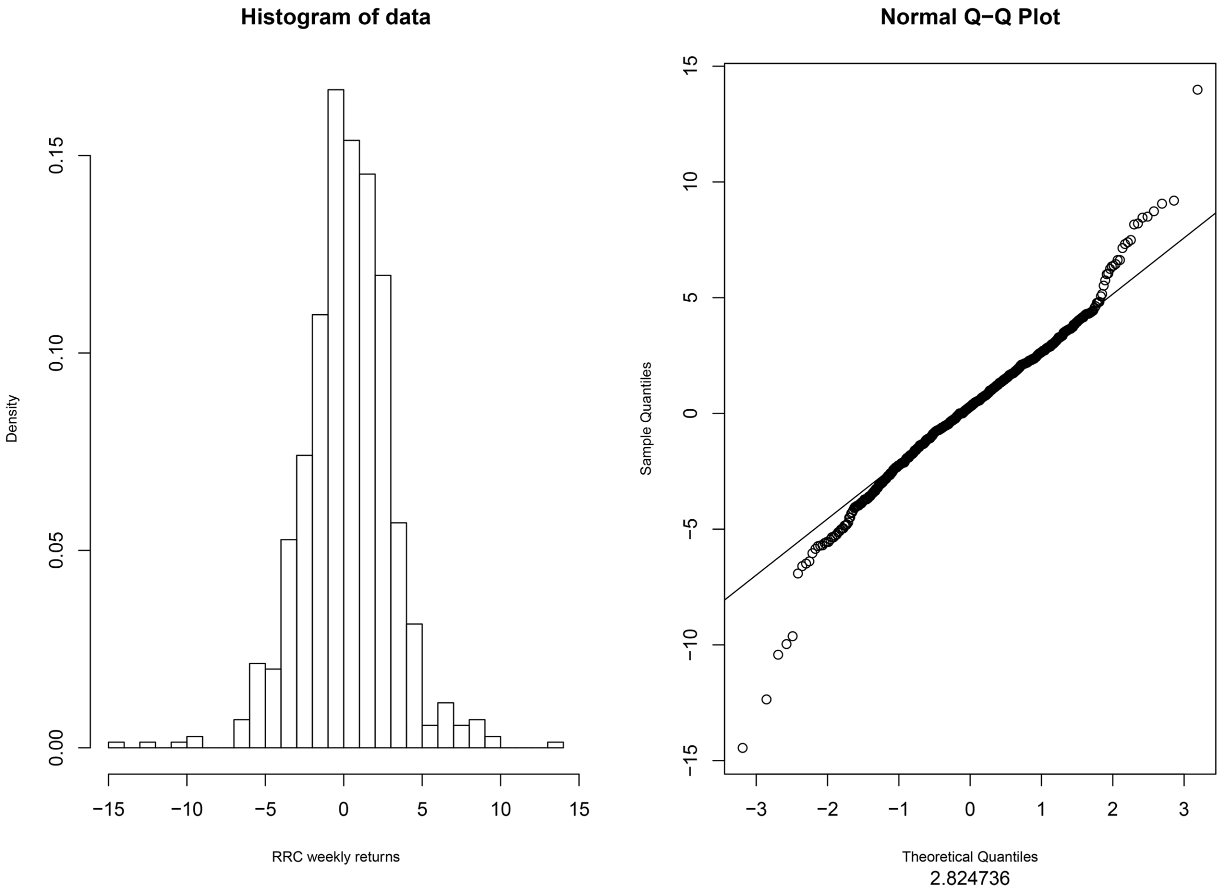

Figure 1. Fitting Model 1 to RRC weekly returns.

Table 5. Summary Statistics for RRC weekly log-returns.

Minimum	Standard.dev	skewness	exc.kurtosis	Maximum	Mean	N
−14.4465	2.824736	−0.03586155	2.768252	13.9830	0.2333	702

The initial values were used in all the proposed models to obtain the maximum likelihood estimates as shown in Table 6 below

The parameter estimates from Table 2 are now fitted to RRC weekly log-returns. Figures 2-5 show the histogram and Q-Q plots of the RRC returns fitted with the proposed models. Figure 2-5 show that the proposed model fit the data well.

Table 7 present results of Kolmogorv-Smirnov and Anderson-Darling test performed on the models. All models produce high p-values, a strong evidence that we can not reject the null hypothesis that the returns data follow the proposed models.

Table 8 presents values of Alkaike Information Criterion (AIC), Bayesian Information Creterion (BIC) and Log-likelihood. The values illustrate that the models are alternative to each other.

Model 3 has the lowest AIC and BIC with the highest log-likelihood. It is the best fit for the data.

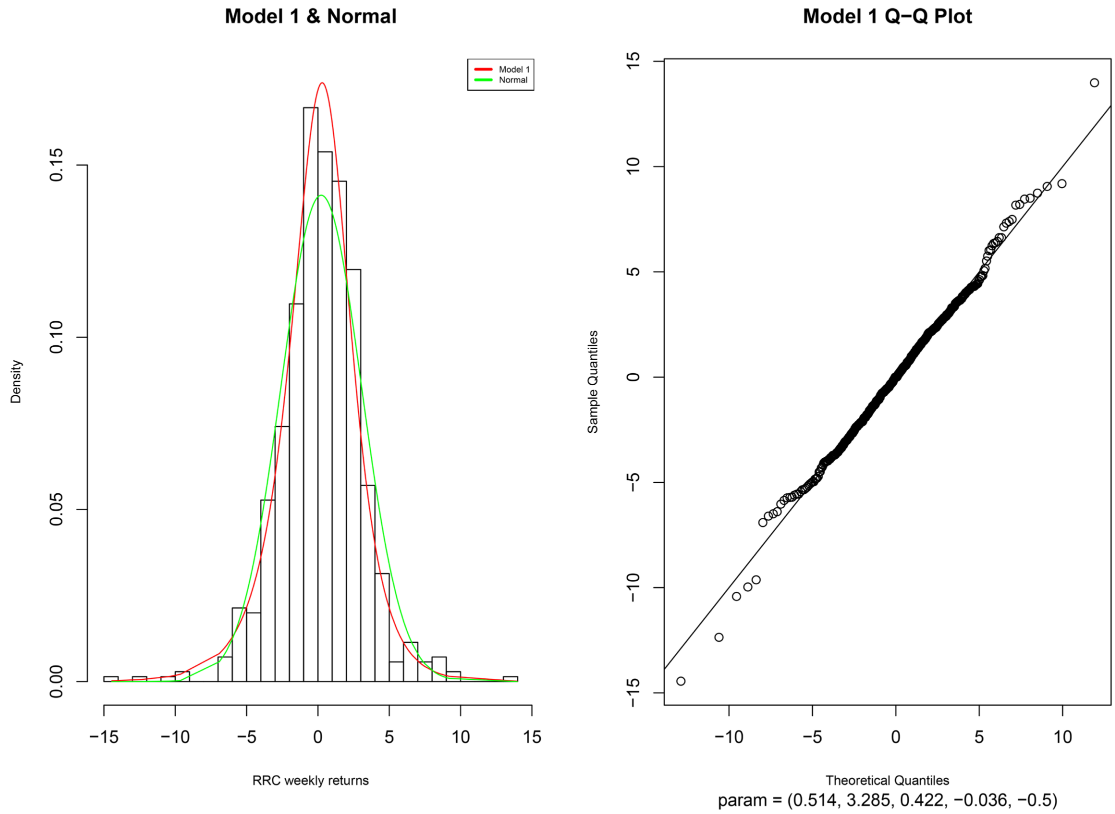

Figure 2. Fitting Model 1 to RRC weekly returns.

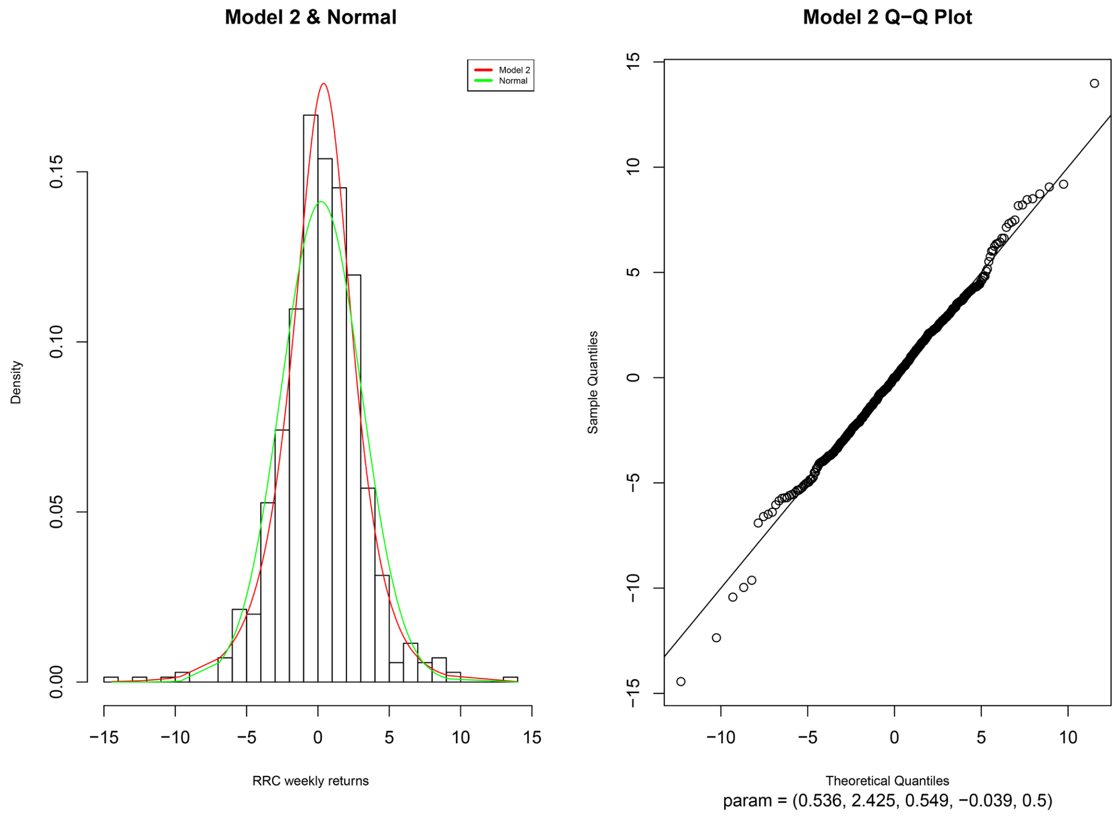

Figure 3. Fitting Model 2 to RRC weekly returns.

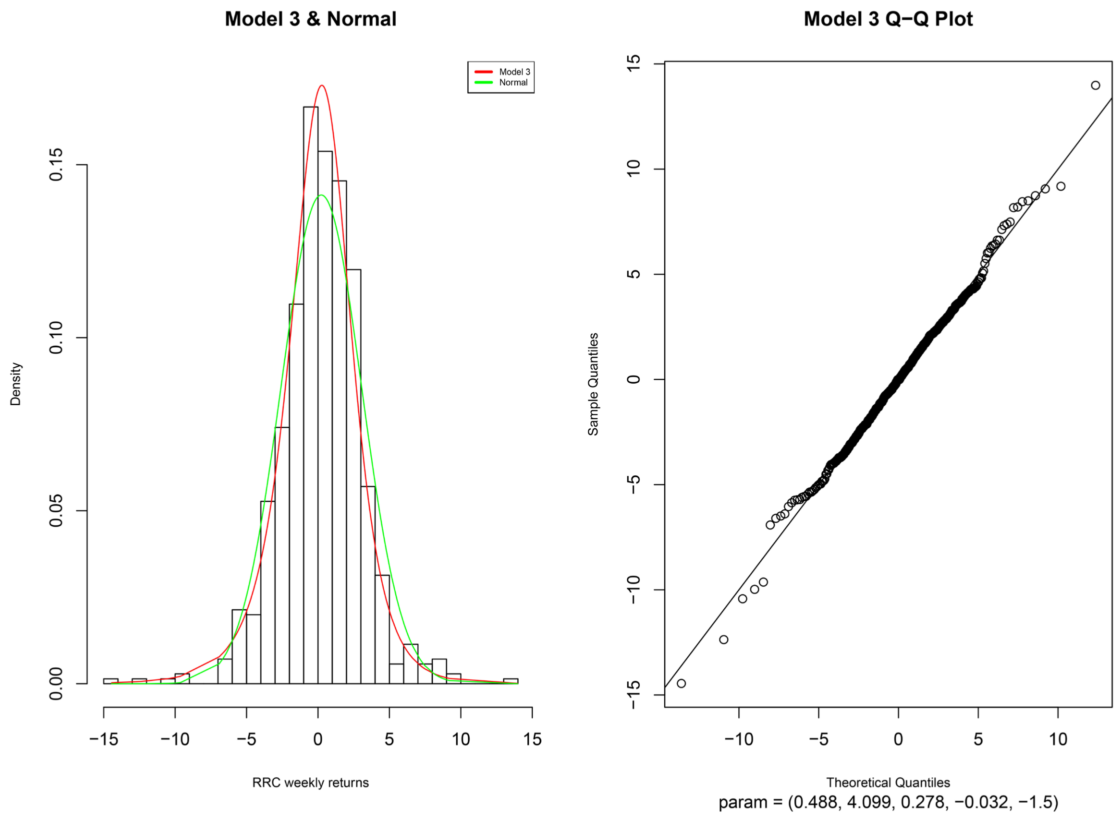

Figure 4. Fitting Model 3 to RRC weekly returns.

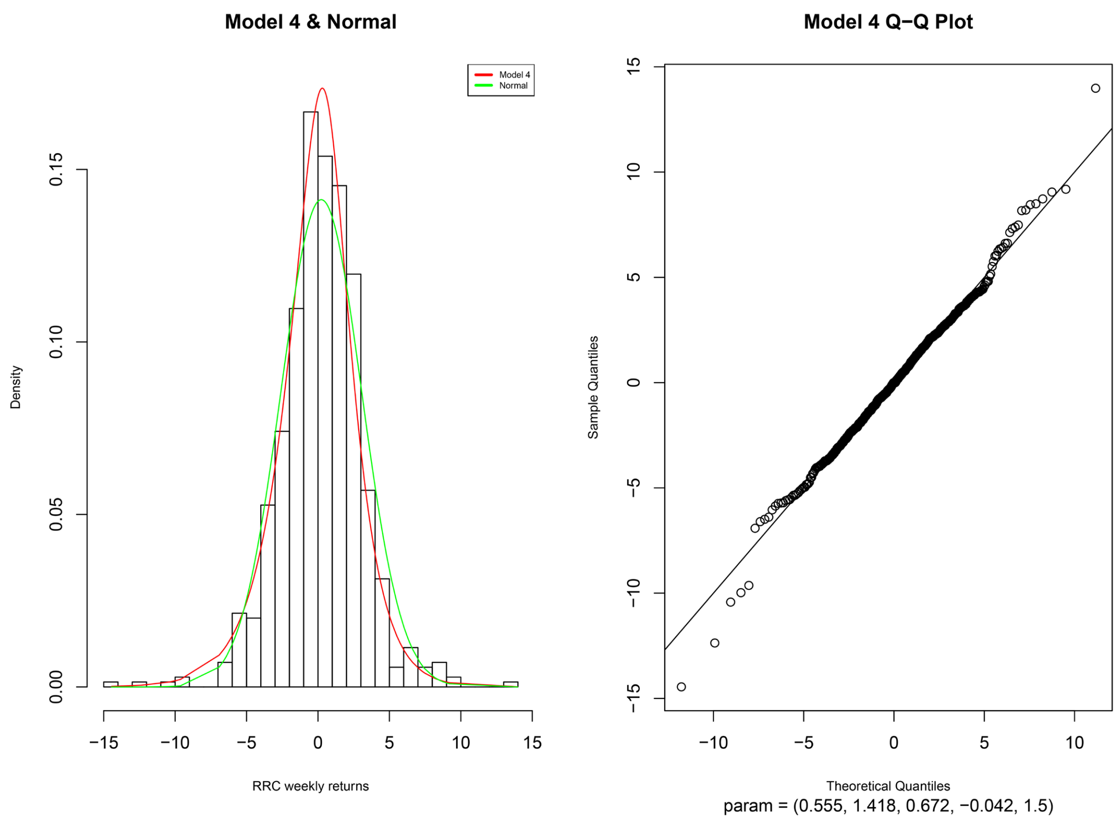

Figure 5. Fitting Model 4 to RRC weekly returns.

Table 6. Estimates for the proposed models.

Parameter	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{\mu}$
Model 1 ($\lambda = -\frac{1}{2}$)	0.4215579	3.285072	−0.03586155	0.5137899
Model 2 ($\lambda = +\frac{1}{2}$)	0.5491998	2.425010	−0.03904892	0.536296
Model 3 ($\lambda = -\frac{3}{2}$)	0.2778586	4.098694	−0.03234413	0.4882795
Model 4 ($\lambda = +\frac{3}{2}$)	0.6724609	1.418126	−0.04177948	0.5546103

Table 7. Results from Kolmogorov-Smirnov Test and Anderson-Darling Test.

Parameter	Kolmogorov-Smirnov statistic	p-value	Anderson-Darling statistic	p-value
Model 1 ($\lambda = -\frac{1}{2}$)	0.0168	0.9890	0.24765	0.9716
Model 2 ($\lambda = +\frac{1}{2}$)	0.0166	0.9904	0.23564	0.9775
Model 3 ($\lambda = -\frac{3}{2}$)	0.0165	0.9912	0.26102	0.9643
Model 4 ($\lambda = +\frac{3}{2}$)	0.0155	0.9958	0.27744	0.9541

Table 8. AIC, BIC and Log-likelihood Values.

Model	Model 1	Model 2	Model 3	Model 4
AIC	3399.776	3400.898	3398.976	3402.382
BIC	3417.992	3419.114	3417.192	3420.598
Log-likelihood	−1695.888	−1696.449	−1695.488	−1697.191

8.2. Risk Estimation and Backtesting

We use the parameter estimates for our proposed model to determine the VaR (Table 9) and ES (Table 10) at levels $\alpha \in \{0.001, 0.01, 0.05, 0.95, 0.99, 0.999\}$. The first three level are used to measure the risk of long position, while the last three levels are used to measure the risk of short positions. We apply the Kupiec Likelihood Ratio (LR) test given by Kupiec [16] which test the hypothesis that the expected proposition of violations is equal to α. The method consist of calculating $\tau(\alpha)$ the number of times the observed returns, x_t falls below (for long position) or above (for short position) the VaR_α estimates at level α; *i.e.*, $x_t < VaR_\alpha$ or $x_t < VaR_{1-\alpha}$, and compare the corresponding failure rate to α.

The likelihood ratio statistic is given by

$$2\log\left(\frac{\tau(\alpha)}{n}\right)^{\tau(\alpha)}\left(1-\frac{\tau(\alpha)}{n}\right)^{n-\tau(\alpha)}-2\log\left(\alpha^{\tau(\alpha)}-(1-\alpha)^{n-\tau(\alpha)}\right) \tag{61}$$

where $\tau(\alpha)$ is the number of violations. Under the null hypothesis this statistic is distributed as χ^2 distribution with one degree of freedom. Table 11 gives the number of violations for the models. Table 12 gives the p-values based on Kupiec test.

Model 3 has the highest VaR and ES value indicating that it perform well than the other models at the tails.

Table 9. VaR Values of RRC log-returns based on normal and proposed models.

	0.001	0.01	0.05	0.95	0.99	0.999
Normal	−8.495775	−6.33803	−4.412962	4.879592	6.804634	8.962406
Model 1	−12.175020	−7.483157	−4.387882	4.621687	7.300979	11.305172
Model 2	−11.676119	−7.396380	−4.414590	4.635737	7.248426	10.976183
Model 3	−12.770428	−7.524902	−4.344605	4.605084	7.328694	11.666360
Model 4	−11.206503	−7.271316	−4.422422	4.646686	7.176342	10.659890

Table 10. ES Values of RRC log-returns based on normal and proposed models.

	0.001	0.01	0.05
Model 1	−14.31580521	−9.51044987	−6.32267305
Model 2	−13.54898243	−9.25410370	−6.26915453
Model 3	−15.35943879	−9.77595177	−6.35304744
Model 4	−12.88318596	−8.98494206	−6.18936754

Table 11. Number of violations of VaR for each distribution at different levels.

	0.001	0.01	0.05	0.95	0.99	0.999
Normal	5	9	33	24	12	3
Model 1	2	5	33	28	11	1
Model 2	2	5	33	28	10	1
Model 3	2	5	33	27	11	1
Model 4	2	5	33	27	11	1

Table 12. P-value for the kupiec test for each distribution at different levels.

	0.001	0.01	0.05	0.95	0.99	0.999
Normal	8.8068×10^{-4}	0.471717	0.7134756	0.04196382	0.086239	0.0422255
Model 1	0.2067157	0.4191802	0.7134756	0.20316	0.1632629	0.7381375
Model 2	0.2067157	0.4181802	0.7134756	0.144112	0.1632629	0.7381375
Model 3	0.2067157	0.4191802	0.7134756	0.20316	0.287939	0.7381375
Model 4	0.2067157	0.4181802	0.7134756	0.1444112	0.1632629	0.7381375

Remark: At 5 percent level of significant, the Normal distribution is rejected at levels: 0.001, 0.95 and 0.999. In addition it is also rejected at level 0.99 at 10 percent level of significant. The Normal weighted Inverse Gaussian distributions were all effective and well specified on all levels of VaR. It can be noted model 3 (Normal- $GIG\left(-\frac{3}{2},\delta,\gamma\right)$) outperforms the other models at level 0.99.

9. Conclusions

In this paper we obtained VaR using NWIG distributions. We first constructed Normal Variance-Mean Mixture when the mixing distributions are $GIG\left(-\frac{1}{2},\delta,\gamma\right)$, $GIG\left(\frac{1}{2},\delta,\gamma\right)$, $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ and $GIG\left(\frac{3}{2},\delta,\gamma\right)$. We have shown that these mixing distribution are WIG distributions.

The parameters of the mixed models are estimated using EM-algorithm. The iterative schemes used are based on explicit solutions of normal equations. We used method of moment estimates of NIG as initial values and obtained monotonic convergence.

We used AIC, BIC and loglikelihood for model selection. Normal- $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ was found to be the best model. The results show that the three NWIG distributions are as good as NIG for VaR computation.

Further work can be done on Normal Mixtures when the mixing distributions are Finite mixtures of WIG distributions.

Conflicts of Interest

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

References

[1] Zhang, Y. and Nadarajah, S. (2017) A Review of Backtesting for Values at Risk. In: *Communications in Statistics— Theory and Methods*, **47**, 3616-3639.

[2] Wimmerstedt, L. (2015) Backtesting Expected Shortfall: The Design and Implementation of Different Backtest. Master's Thesis in Mathematical Statistics, Royal Institute of Technology, Stockholm.

[3] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1997) Thinking Coherently. *Risk Journal*, **10**, 68-71.

[4] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1997) Coherent Measures of Risk. *Mathematical Finance*, **9**, 203-228. https://doi.org/10.1111/1467-9965.00068

[5] Gneiting, T. (2011) Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, **106**, 746-762. https://doi.org/10.1198/jasa.2011.r10138

[6] Basel Committee (2013) Fundamental Review of the Trading Book: A Revised Market Risk Framework.

[7] 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

[8] Barndorff-Nielsen, O.E. (1977) Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Proceedings of the Royal Society of London. Series A*, **353**, 409-419. https://doi.org/10.1098/rspa.1977.0041

[9] 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

[10] Dempster, A.P., Laird, N.M. and Rubin, D. (1977) Maximum Likelihood from Incomplete Data Using 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

[11] 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

[12] 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

[13] Akman, H.O. and Gupta, R.C. (1992) A Comparison of Various Estimators of the Mean of an Inverse Gaussian Distribution. *Journal of Statistical Computation and Simulation*, **40**, 71-81. https://doi.org/10.1080/00949659208811366

[14] Gupta, R.C. and Akman, O. (1995) On Reliability Studies of a Weighted Inverse Gaussian Distribution Model. *Statistical Papers*, **38**, 445-452. https://doi.org/10.1007/BF02925999

[15] 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

[16] Kupiec, P. (1995) Techniques for Verifying the Accuracy of Risk Measurement Models. https://doi.org/10.3905/jod.1995.407942

[17] Rappoport, P. (1993) A New Approach: Average Shortfall. Technical Report, J.P. Morgan.

[18] Yamai, Y. and Yoshiba, T. (2002) On the Validity of Value at Risk Comparative Analyses with Expected Shortfall. Monetary and Economic Studies.

[19] Zhang, Y., Chu, J., Chan, S. and Chan, B. (2019) The Generalised Hyperbolic Distribution and Its Subclass in Analysis of a New Era of Cryptocurrencies: Ethereum and Its Financial Risk. *Physica A*, **526**, 1-12. https://doi.org/10.1016/j.physa.2019.04.136

[20] Bams, D. and Wilhouwer, J.L. (2001) Empirical Issues in Value at Risk. *ASTIN Bulletin*, **31**, 299-315. https://doi.org/10.2143/AST.31.2.1007

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

[22] Kostas, F. (2007) Tests of Fit for Symmetric Variance Gamma Distributions. UADPhilEcon, National and Kapodistrian University of Athens, Athens.

[23] 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