UNIVERSITY OF NAIROBI

NORMAL WEIGHTED INVERSE GAUSSIAN DISTRIBUTIONS AND EM ALGORITHM WITH APPLIACTIONS TO RISK MEASURES AND DEPENDENCE MODELLING

CALVIN BITANGE MAINA

(B.Sc. UNIVERSITY OF NAIROBI, KENYA;

M.Sc. UNIVERSITY OF NAIROBI, KENYA)

REG. NO. I80/58074/2021

ATHESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN ACTUARIAL SCIENCE IN THE DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NAIROBI

2022

Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

9.11.2022

Signature Date

CALVIN BITANGE MAINA
Reg No.I80/58074/2021

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

10.11.2022

Signature Date

PROF: P.G.O WEKE.
Department of Mathematics, University of Nairobi

10.11.2022

Signature Date

DR: C.A. OGUTU.
Department of Mathematics, University of Nairobi

10.11.2022

Signature Date

PROF: J.A.M OTTIENO.
Department of Mathematics, University of Nairobi

Acknowledgments

First, I thank our Almighty Father for the gift of life, health and provision which enabled me carry out this research. I experienced His love and blessing.
To Prof. P.G.O. Weke, you have guided me throughout my Actuarial Studies at the university. You have been a mentor any student would wish to have. At the commencement of this work, you happened to be the Director in the School of Mathematics (SoM), but despite the demands and busy schedule you spared time to input your invaluable intellect. Still, at all stages you afforded to work timely. All approvals especially signature online were done promptly.
To Dr. C. A. Ogutu, your invaluable input to this research is greatly appreciated. Thank you for ensuring that the application of the models were accurate and well done. I appreciate the timely approval required at different stages.
To Prof. J.A.M Ottieno, I am greatly indebted to you for your patience, intelligence and hard work. I will never forget the two months you fully accommodated me at your place. You made me understand it was a training. A period that I will forever treasure. Safe in your experience we proposed a new class of distribution.
I would like to thank Prof. J. I. Mwaniki for his computational advice in statistical Modelling. The concepts were of great importance for this work.
To Prof. P. Ngare, am grateful for the guidance accorded in writing of this Thesis. Your advice on publication was of great importance.
Special thanks goes to the Director, SoM, Prof. Luketero for always encouraging me to work hard and complete the research. I equally thank Dr. James Katende for for organising the seminar talks and for his encouragement and advice on how to improve on presentation.
A special thank to my family. The encouragement and financial support from parents always kept me moving. The constant push from my dear mother being what I needed most.
To my best friend and wife, Edwinah, thanks for understanding the dynamics we faced and being there for the kids. Michelle was born while I was away and you still encouraged me to work hard. You are a blessing and my God bless you.
To my twins Daniel and Sophie you became a source of strength for me. Thanks for understanding-even when the time spent with you was minimal.

Calvin Bitange Maina

Nairobi, 2022.

Dedication

For my daughter Michelle.

Abstract

High frequency financial data is characterized by non-normality, asymmetric, leptokurtic and fat-tailed behaviour. The normal distribution is inadequate in capturing these characteristics. To this end, various flexible distributions have been proposed. In this Thesis we introduced a new class of distributions known as Normal Weighted Inverse Gaussian distributions.

Weighted Inverse Gaussian distributions are special cases of Generalized Inverse Gaussian (GIG) distribution which are related to Inverse Gaussian (IG) distribution. Finite mixtures of these special cases are also weighted Inverse Gaussian (WIG) distributions. Using these WIG distributions as mixing distributions to the Normal Variance Mean Mixture (NVMM) we obtain a class of Normal Weighted Inverse Gaussian (NWIG) distributions.

The properties considered for these models are mean, variance, skewness and kurtosis. For data analysis we consider three data sets: 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. Estimation of parameters of these models are obtained using Expectation-Maximization (EM) algorithm. The 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. The mixing operation can be considered responsible for producing missing values.

Two important risk measures in literature are Value at Risk (VaR) and Expected Shortfall (ES). In this work we have obtained VaR and ES for the NWIG distributions. Backtesting of this measures is also performed.

We have also considered dependence modelling of financial returns using copulas. The marginal distributions are based on Normal Weighted Inverse Gaussian distributions.

We highlight the following contributions to this work

1. We have constructed a new class of Weighted Inverse Gaussian distributions.

2. We have used this class as a mixing distribution to the Normal Variance Mean Mixture to obtain a class of Normal Weighted Inverse Gaussian distributions.

3. All works on parameter estimation of EM algorithm at the maximization step is based on explicit solution to normal equations. Often this involves numerical techniques which are difficult to implement. In this work, we show that the iterative schemes are not necessarily based on explicit solutions. They can also be designed using a representation based on the normal equations. This subtle approach is easily programmable and preserves the monotonic convergence property of the EM algorithm with each iteration increasing the likelihood.

4. From the data sets used, this class of NWIG distributions is shown to be a good alternative to the Normal Inverse Gaussian distribution. However, one special case of the GIG distribution when used as a mixing distribution to NVMM outperforms the NIG and one finite case when used as a mixing distribution outperforms all models.

5. Using backtesting procedures it can be shown that this class of distributions, NWIG, which have heavy tailed, is an alternative candidate for financial risk management.

6. The models have also been used as marginals in dependence modelling using copulas approach.

Abbreviations and Notations

The general abbreviations and notations used in the thesis are given.

$k_\lambda(w)$	Modified Bessel function of the third kind
IG	Inverse Gaussian
RIG	Reciprocal Inverse Gaussian
GIG	Generalised Inverse Gaussian
WIG	Weighted Inverse Gaussian
NIG	Normal Inverse Gaussian
NRIG	Normal Reciprocal Inverse Gaussian
GHD	Generalised Hyperbolic Distribution
NVMM	Normal Variance Mean Mixture
NWIG	Normal Weighted Inverse Gaussian
cdf	Cummulative distribution function
pdf	Probability density function
BCBS	Basel Committee on Banking Supervision
VaR	Value at Risk
ES	Expected Shortfall
C(u,v)	Bivariate Copula function of margins u and v
MoM	Method of Moment
ML	Maximum Likelihood
EM	Expectation Maximization
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
$M_X(t)$	Moment generating function of X
$M_Z(t)$	Moment generating function of Z
RRC	Range Resource Corporation
CVX	Shares of Chevron
$s\&p500$	Standard and Poor's 500 index
$f(x\|z)$	Conditional distribution
$g(z)$	Mixing distribution

$f(z\|x)$	Posterior distribution
γ_1	Skewness
γ_2	Excess Kurtosis

List of Figures

List of Tables

Table of Contents

1 INTRODUCTION

1.1 Background Information

Managing risk for the purpose of economic capital allocation is one of the main objectives of financial institutions to guarantee solvency to their clients and counterparties. In general, Risk measures are used to quantify risk to determine appropriate capital to survive in extreme market conditions. Capital amounts therefore act as buffer against insolvency.

There are many possible interpretations and different ways of quantifying investment risk. In this work we focus on Value at Risk (VaR) and Expected Shortfall (ES). These are the two main risk measures used in financial institutions and regulation.

For the purpose of VaR and ES analysis (for a single asset or a portfolio), a model for the return distribution also known as profit and loss (P&L) distribution is important because it describes the potential behaviour of a financial security in the future.

1.2 Definitions, Terminologies and Notations

Risk Measure

Let $(\Omega, \mathscr{F}, P)$ be a probability space and V be a non-empty set of $\mathscr{F}$-measurable real-valued random variables. Then any mapping

$$\rho : V \longrightarrow \mathbb{R} \cup \{\infty\} \tag{1.1}$$

is called a risk measure.

Properties of Risk Measures

Let $\alpha \in (0,1]$ be fixed and $(\Omega, \mathscr{F}, P)$ be a probability space. Consider the risk measure ρ on the set V of all the $\mathscr{F} - measurable$ real-valued random variables. We have the following Axioms:

1 Translation invariance:

$$X \in V, a \in \mathbb{R}, X + a \in V \Longrightarrow \rho(X + a) = \rho(X) - a \tag{1.2}$$

This defines a risk measure as the buffer capital needed to maintain a certain level of risk.

2 Monotonicity:

$$X, Y \in V, X \leq Y \Longrightarrow \rho(X) \geq \rho(Y) \tag{1.3}$$

Portfolio with higher guaranteed value will always be less risky.

3 Positive homogeneity:

$$X \in V, h > 0, hX \in V \Longrightarrow \rho(hX) = h\rho(X) \tag{1.4}$$

to double the capital means to double the risk.

4 Law of invariance:

$$X, Y \in V, P[X \leq t] = P[Y \leq t] \text{for all} t \in \mathbb{R} \Longrightarrow \rho(X) = \rho(Y) \tag{1.5}$$

5 Subadditivity:

$$X, Y \in V \Longrightarrow \rho(X+Y) \leq \rho(X) + \rho(Y) \tag{1.6}$$

Two combined portfolios should never be more risky than the sum of the risk of the two portfolios separately.

The most popular downside risk measure is **Value at Risk** (VaR). It is generally defined as possible maximum loss over a given holding period within a fixed confidence level. In statistical terms, VaR is a quantile of distribution for financial asset returns. More formally, VaR is defined as

$$P\{X \leq -VaR^X_{1-\alpha}\} = \alpha \tag{1.7}$$

where X represents the Asset's returns for a symmetric distribution. In general, the integral form can be expressed as

$$\int_{-\infty}^{VaR^X_\alpha} f(x)dx = \alpha \tag{1.8}$$

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

Conditional Expectation

$$E[X|X < VaR_\alpha] = \int_{-\infty}^{VaR_\alpha} x\frac{f(x)}{\alpha}dx \tag{1.9}$$

is the Expected Shortfall denoted as ES_α.

Weighted Distribution

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

$$E[w(X)] = \int_{-\infty}^{\infty} w(x)f(x)\,dx \tag{1.10}$$

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

$$\implies f_W(x) \quad = \frac{w(x)}{E[w(X)]} f(x), -\infty < x < \infty \tag{1.11}$$

Normal Variance-Mean mixture

A Normal Variance-Mean mixture can be presented in integral form as:

$$f(x) = \int_0^\infty \frac{1}{\sqrt{2\pi z}} e^{-\frac{[x-(\mu+\beta z)]^2}{2z}} g(z)\, dz \tag{1.12}$$

with the conditional distribution,

$$X/Z = z \sim N(\mu + \beta z, z) \tag{1.13}$$

and g(z) being the mixing/averaging distribution. $f(x)$ is called a Normal Variance Mean Mixture.

1.3 Statement of Problem

Nadarajah et al., (2014) have given a detailed review of VaR and ES for various distributions. One of the distributions considered is the Generalised Hyperbolic Distribution (GHD). GHD is a Normal Variance Mean Mixture (NVMM) with the Generalised Inverse Gaussian (GIG) mixing distribution. The GIG is a three parameter distribution denoted as $GIG(\lambda, \delta, \gamma)$. The GIG embraces a number of special case distributions. When the index parameter take the value $\lambda = -\frac{1}{2}$ we obtain the Inverse Gaussian (IG) distribution. The Normal Inverse Gaussian (NIG) distribution is a NVMM with IG mixing distribution.

Other special cases of GIG can be obtained when $\lambda = \frac{1}{2}$, $\lambda = -\frac{3}{2}$ and $\lambda = \frac{3}{2}$. These special cases can be shown to be Weighted Inverse Gaussian (WIG) distribution. When used as mixing distribution in NVMM we obtain a class of distribution known as Normal Weighted Inverse Gaussian (NWIG) distributions. In constructing this new class of distributions a number of issues arose and needed to be addressed. Thus, some of these issues have formed part of the problem statement in terms of questions.

(i) A number of special cases of distributions can be obtained when the index parameter of the GIG assume specific values such as $\lambda = -\frac{1}{2}$, $\lambda = \frac{1}{2}$, $\lambda = -\frac{3}{2}$ and $\lambda = \frac{3}{2}$. The question therefore is; how are these special case distributions related to the IG distribution? What are their properties?

(ii) The **four** special cases can be combined to construct **Six** finite mixtures. Can these models be expressed in terms of IG distribution? What are their properties?

(iii) The **Ten** special cases can be used as mixing distribution in NVMM. What are the properties of these Mixtures?

(iv) The Expectation Maximization (EM) algorithm has been used to estimate parameters of the mixed models since the mixing operation is considered responsible for producing missing data. What iterative schemes can be designed to estimate the parameters of these models?

(v) How can these class of Normal Weighted Inverse Gaussian ($NWIG$) distributions be compared to the Normal Inverse Gaussian (NIG)?

(vi) For the purpose of Value at Risk (VaR) and Expected Shortfall (ES) computation, is the class of NWIG distributions of any economical value compared to the NIG distribution?

(vii) The NIG has commonly been used for marginals in dependence modelling using Copulas. What difference does it make using the "Best Model" identified for the class of NWIG distributions?

1.4 Objectives

1.4.1 General Objective

The main objective is to measure Value at Risk (VaR) and Expected Shortfall (ES) based on Normal Weighted Inverse Gaussian (NWIG) distributions for the purpose of economical capital allocation and dependence modelling.

1.4.2 Specific Objectives

(a) To obtain the class of Weighted Inverse Gaussian distribution.

  (i) Deduce the four special cases of GIG when the indexes are: $-\frac{1}{2}$, $\frac{1}{2}$, $-\frac{3}{2}$ and $\frac{3}{2}$.

  (ii) construct the six pairwise finite mixtures of the four special cases of the GIG.

  (iii) express the ten special cases as Weighted Inverse Gaussian distributions.

  (iv) Obtain the properties of the class of Weighted Inverse Gaussian distributions

(b) To use the class of Weighted Inverse Gaussian distributions as mixing distributions in Normal Variance Mean Mixture to obtain the class of Normal Weighted Inverse Gaussian distributions.

(c) To obtain the Maximum Likelihood parameter estimates for the class of Normal Weighted Inverse Gaussian distributions via the Expectation Maximization (EM) algorithm.

  (i) find the posterior estimates for the missing value which are functions of the random variable of the mixing distributions.

  (ii) to design an iterative scheme for each Normal Weighted Inverse Gaussian Mixtures based on the normal equations

(iii) identify initial values for the iteration

(d) To compute Value at Risk and Expected Shortfall of the fitted models for some financial data

(e) To perform model selection and identify the "Best Model" for dependence Modelling using Copulas.

1.5 Literature Review

Markowitz (1952) introduced the **Modern Portfolio Theory** based on **Variance** which became the dominating risk measure. However, MPT has a number of **Drawbacks:**

1 Risks are random variables with finite variance

2 It implicitly assumes that their distribution are approximately symmetric around the mean.

3 It does not necessarily corresponds to investors' perception of risk.

The most popular downside risk measure is **Value at Risk** (VaR). 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. It is generally defined as possible maximum loss over a given holding period within a fixed confidence level.In statistical terms, VaR is a quantile of distribution for financial asset returns. It has become the classic measure that financial executives use to quantify market risk. When RiskMetrics announced Value at Risk as its measure of risk in 1996, the Basel Committee on Banking Supervision enforced financial institutions to meet capital requirements based on VaR estimates.
However, the main shortcoming of VaR is that it fails to capture tail risk. It does not take into account what happens beyond the threshold level. VaR also lacks a mathematical property called subadditivity (Wimmerstedt, 2015). This implies that diversification could increase risk.
These limitations have prompted the implementation of an alternative, coherent measure of risk - **the Expected Shortfall (ES)**.
The concept of Expected Shortfall (ES) was first introduced in Rappoport (1993). Artzner et. al (1997, 1999) formally developed the concept.

1.5.1 Backtesting a Risk Measure

Backtesting a risk measure is the same as evaluating forecasting performance.
To achieve this directly, we exploit the mathematical property of elicitability introduced by Osband (1985) and further developed by Lambert et al. (2008).
A statistical functional, such as the mean or median, is called elicitable if there's a scoring function such that the correct forecast of the functional is the unique minimizer of the expected score. VaR can be shown to be elicitable. Although ES appears to be a more suitable risk measure than VaR because of its coherence and tail sensitivity, it has been shown that, in contrast to VaR, it lacks the property of elicitability (Gneiting, 2011). ES has been shown to be conditionally elicitable (Emmer et al., 2015) and jointly elicitable with VaR (Fissler and Ziegel, 2016). Therefore, VaR and ES can be backtested jointly. Deng

and Qui (2021) have conducted a comprehensive study of the performance of leading procedures for ES with more than a dozen variation. The Basel Committee (2013) proposed to replace Value at Risk with Expected Shortfall. The 2016 Basel IV framework for the first time openly advocated the use of expected shortfall (ES) in risk management on a daily basis (Basel Committee on Banking Supervision).
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 Wielhouwer, 2001). A Normal distribution supposedly underestimates the tail and hence VaR.

1.5.2 Normal Mixtures

Nadarajah et al., (2014) have given a detailed review of VaR and ES for various distributions. One of the distribution reviewed is the Generalized Hyperbolic Distribution (GHD) introduced by Barndorff-Nielsen (1977). The GHD nest a number of special and limiting case distribution. A number of researchers have studied these cases. They include: Eberlein and Keller(1995) studied the Hyperbolic distribution, Barndorff-Nielsen (1997) considered the special case of Normal Inverse Gaussian distribution, Madan and Seneta (2005) considered the limiting case of Variance Gamma and Aas and Haff (2005,2006) reviewed the NIG and studied the Generalised Hyperbolic Skew Student's t distribution.

1.5.3 Estimation

The Mixing operation is considered responsible for producing missing values. The Expectation Maximization algorithm introduced by Dempster et al. (1977) has been used to estimate parameters of mixed models. Karlis (1995) applied the algorithm to estimate parameters of mixed Poisson distribution, Karlis (2002) estimated the parameter of NIG, Aas and Haff (2005,2006) applied the algorithm to the Generalised Hyperbolic Skew Student's t distribution.

1.5.4 Dependence Modelling

Dependence modelling of financial data using copulas has gained attraction in finance. A number of researchers have used copulas to model dependent finacial data. Lo et al. (2013) used canonical vine (C-vine) copulas. Kraus and Czando (2017) used D-vine copula and Olechrin and Teteneva (2017) use hierachial copula to estimate VAR. Bynn and Sony (2021) sought the best copula calculating VAR of a portfolio with many assets. They used the vine copulas and hierachial copulas.The objective was to choose a proper copula function to reflect the variance dependence structures in a portfolio and the performance was computed by VAR of the portfolio.
As for the marginal distribution we can use different distributions to each asset in a portfolio. It has been shown that the normal distribution is inappropriate to model

the return distribution of financial assets. The return distributions of financial assets are slightly skewed and fat tailed. There has been extensive research on finding good alternatives to the normal distribution in literature.

For example, Venkiataraman (1997) used a quasi-bayesian maximum likelihood estimation procedure. Hull and White (1998) used a transform to multivariate normal distribution which is updating schemes such as GARCH. Eberlein and Keller (1995) used hyperbolic distribution. Modem et al. (1998) use Variance Gamma (VG) distribution. Bandorff-Neilsen (1997) and Mabitreda et al. (2015) used Normal Inverse Gaussian distribution. Byun and Sony (2021) Used NIG distribution as marginal distribution. The NIG is known to have better return portfolio than VG distribution. (Ericksen et al. 2009, Gencil and Yeng 2010) and the calculation of VAR using NIG is also better than other models such as GARCH or VG. (Welhelmson, 2009, Kim and Song, 2011, Doric and Doric, 2011.)

Bolvinken and Benth (2000), Godin et al. (2012) have also used the NIG distribution for modeling return distributions of financial assets.

Copulas are popular in modelling a joint distribution of several asset returns in finance. With copulas we can construct a multivariate distribution with different marginal distributions by separating the dependence from marginal distributions. Also, there are many copulas that can incorporate the proper dependence structure of the data. Embrechts et al. (2002) show that copulas are useful in identifying the dependence structure annually returns of assets. Bynn and Song (2021) used copulas to identify the dependence structures and to generate a multivariate distribution for returns of assets in a portfolio. The VAR of the portfolio was computed using the resulting multivariate distribution.

Empirical copula include Gaussian (normal) copula and Students t copula and they are commonly used in the field of finance. Archimedean copulas include Clayton, Gumbell, and Frank copulas and they have a generator function that is useful in expressing various dependence structures. Wu et al. (2007) use exchangeable Archimedean copulas to calculate VAR. When there are more than two assets in a portfolio, elliptical and exchangeable.

1.5.5 Value at Risk

According to Jorian (2007), Value at Risk (VAR) is defined as the worst loss over a target region that will not be exceeded given a certain level of confidence. VAR is one of the most common risk management tools in finance. It is simple and intuitive in the sense that it summarizes the change in a value of a portfolio into a single number. Calculating VAR of a portfolio of assets is important in finance because it has been widely accepted that the investment in a portfolio of several assets has advantages over investment in a single asset. (Markowitz, 1952). Different assets have different return distributions and they are correlated in various ways. The key element in calculating VAR is to capture the relationship among assets in the portfolio appropriately. In order to capture the relationship among assets and calculate VAR of the portfolio, multivariate distributions can be used.

We may want to work with a univariate distribution from the history of returns from the same portfolio because the returns of the portfolio itself is univariate. However, using a multivariate distribution for a collection of assets that make up the portfolio will give us more flexibility.

1.6 Significance of Study

Value at Risk and Expected Shortfall are the most popular measures of risk. For the purpose of their estimation a model for the return distribution is important because it describes the potential behaviour of a financial security in the future. Financial institutions allocate economic capital, based on the risk quantified, to guarantee solvency to their clients and counterparties. It has been shown that the normal distribution is inappropriate to model such returns. The returns distributions of financial assets are skewed, leptokurtic and fat tailed. There has been extensive research on finding good alternatives to the normal distribution in literature. In this work, the class of Normal Weighted Inverse Gaussian distributions has been proposed. The class nests some special cases of Generalised Hyperbolic Distribution (GHD) that have dominantly featured in finance. The Normal Inverse Gaussian (NIG) distribution has been applied extensively because of its analytical tractability property.
Eberlein and Keller (1995) used hyperbolic distribution. Madam et al. (1998) use Variance Gamma (VG) distribution. Bandorff Nielsen (1997) and Mabitseda et al. (2015) used Normal Inverse Gaussian distribution. Byun and Sony (2021) Used NIG distribution as marginal distribution. The NIG is known to have better return portfolio than VG distribution. (Erikssen et al. 2009, Gencil and Yang 2010) and the calculation of VAR using NIG is also better than other models such as GARCH or VG. (Welhelmson, 2009, Kim and Song, 2011, Doric and Doric, 2011.)
Bolvinken and Benth (2000), Godin et al. (2012) have also used the NIG distribution for modeling return distributions of financial assets. This work extends the NIG to NWIG to improve the accuracy of the estimates. The NWIG models offer good alternative models for NIG. For the purpose of Risk Managentment and financial modelling some NWIG models have outperformed the NIG. The models are used to accurately quantify the risk of financial assets to avoid underestimating or overestimating the economic capital.

1.7 Outline of the Thesis

The rest of the thesis is outlined as follows: In Chapter 2, Generalised Inverse Gaussian distribution (GIG) and its special cases have been constructed. The properties of the distributions have also been studied. In Chapter 3, the Weighted Inverse Gaussian (WIG) class of distribution has been constructed. Finite mixtures of the special cases of GIG have been constructed and their properties studied. Normal variance Mean mixtures with WIG class of distributions has been covered in Chapter 4 and Chapter 5. The Properties of the mixed model have also been studied. In Chapter 6, we have presented the Iterative schemes for the ten mixed models constructed in the previous two chapters. Parameter estimation has been done in chapter 7. Chapter 8 deals with Risk Measures. The fitted models are used to compute the Value at Risk and Expected Shortfall for the financial data sets considered. Dependence Modelling of financial data using Copulas is done in Chapter 9. Chapter 10 deals with Conclusion and Recommendations.

2 GENERALISED INVERSE GAUSSIAN DISTRIBUTION AND ITS SPECIAL CASES

2.1 Introduction

Generalized Inverse Gaussian Distribution is based on the modified Bessel function of the third kind which is the most important mathematical tool used for this work. We first consider the definitions and properties of the Bessel function. We then derive the Generalised Inverse Gaussian (GIG) distribution using Barndorff-Nielsen (1977) parametrization. Special cases of GIG distribution of interest are considered. These models will be fundamental construction weighted distributions discussed in our next chapter.

2.2 Modiffied Bessel Function of the Third Kind

2.2.1 Definition 1 and its properties

Integral Representation

$$K_\lambda(\omega) = \frac{1}{2}\int_0^\infty x^{\lambda-1}\exp\left\{-\frac{\omega}{2}\left(x+\frac{1}{x}\right)\right\}dx \tag{2.1}$$

An alternative form of definition 1 is given in the following:
Proposition 2.1

$$K_\lambda(\omega) = \frac{1}{2}\left(\frac{\omega}{2}\right)^\lambda \int_0^\infty t^{-\lambda-1} exp\left\{-t-\frac{\omega^2}{4t}\right\}dt \tag{2.2}$$

which is obtained when you let $x = \frac{\omega}{2t}$

Property 2.1 (**Symmetry**)

$$K_\lambda(\omega) = K_{-\lambda}(\omega)$$

Property 2.2 (**Derivative I**):

$$\frac{\partial}{\partial\omega}K_\lambda(\omega) = -\frac{1}{2}[K_{\lambda+1}(\omega)+K_{\lambda-1}(\omega)]$$

Property 2.3 (**Derivative II**):

$$\frac{\partial}{\partial\omega}K_\lambda(\omega)=\frac{\lambda}{\omega}K_\lambda(\omega)-K_{\lambda+1}(\omega)$$

Property 2.4 (**Recursive relation**):

$$\frac{2\lambda}{\omega}K_\lambda(\omega)=K_{\lambda+1}(\omega)-K_{\lambda-1}(\omega)$$

Proposition 2.2

$$K_\lambda(\omega)=\frac{1}{2}\int_{-\infty}^{\infty}\exp\{-w\cosh t\}\cosh\lambda t\,dt$$

2.2.2 Definition 2 and Its Properties

$$K_\lambda(\omega)=\left(\frac{\omega}{2}\right)^{\lambda}\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}\int_{1}^{\infty}(t^2-1)^{\lambda-\frac{1}{2}}e^{-\omega t}\,dt \tag{2.3}$$

in terms of hyperbolic functions we have
Proposition 2.3

$$K_\lambda(\omega)=\left(\frac{\omega}{2}\right)^{\lambda}\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)}\int_{0}^{\infty}(\sinh\theta)^{2\lambda}e^{-\omega\cosh\theta}\,d\theta$$

obtained by letting $t=\cosh\theta$
Definition 2 can also be expressed in summation form as given in the following
Proposition 2.4

$$\begin{aligned}(a)K_\lambda(\omega) &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\sum_{i=0}^{\infty}\frac{\Gamma(\lambda+\frac{1}{2})}{i!\Gamma(\lambda+\frac{1}{2}-i)}\frac{\Gamma(\lambda+\frac{1}{2}-i)}{\Gamma(\lambda+\frac{1}{2})}(2\omega)^{-i}\\ &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left(1+\sum_{i=1}^{\infty}\frac{\Gamma(\lambda+\frac{1}{2})}{i!\Gamma\lambda+\frac{1}{2}-i)}\frac{\Gamma(\lambda+\frac{1}{2}-i)}{\Gamma(\lambda+\frac{1}{2})}(2\omega)^{-i}\right)\end{aligned}$$

which can further be expressed as

$$(b)K_\lambda(\omega) = \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left(1+\sum_{i=1}^{\infty}\prod_{i=1}^{n}\frac{[4\gamma^2-(2i-1)^2]}{n!\Gamma(8\omega)^n}\right)$$

Proof

$$K_\lambda(\omega) \;=\; \left(\frac{\omega}{2}\right)^\lambda \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)} \int_1^\infty (t^2-1)^{\lambda-\frac{1}{2}} e^{-\omega t}\,dt$$

Let $y = \omega(t-1)$

$$\begin{aligned}
K_\lambda(\omega) &= \left(\frac{\omega}{2}\right)^\lambda \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)} \frac{e^{-\omega}}{\omega} 2^{2\lambda-1} \int_0^\infty \left(\frac{y}{2\omega}\right)^{\lambda-\frac{1}{2}} \left(1+\frac{y}{2\omega}\right)^{\lambda-\frac{1}{2}} e^{-y}\,dy \\
&= \frac{2^{2\lambda}}{2\omega}\frac{\omega^\lambda}{2^\lambda}\frac{\Gamma(\frac{1}{2})}{\Gamma(\lambda+\frac{1}{2})}\frac{(2\omega)^{\frac{1}{2}}}{2^\lambda\omega^\lambda} e^{-}\omega \sum_{i=0}^{\lambda-\frac{1}{2}} \binom{\lambda-\frac{1}{2}}{i} (2\omega)^{-i}\Gamma\left(\lambda+\frac{1}{2}+i\right) \\
&= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \sum_{i=0}^{\lambda-\frac{1}{2}} \binom{\lambda-\frac{1}{2}}{i} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{\Gamma(\lambda+\frac{1}{2})}(2\omega)^{-i} \\
&= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1+\sum_{i=1}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}\right)
\end{aligned}$$

Expanding the summation we obtain

$$\begin{aligned}
\sum_{i=1}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i} &= \frac{\Gamma(\lambda+\frac{1}{2}+1)}{\Gamma(\lambda+\frac{1}{2}-1)(2\omega)} + \sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i} \\
&= \frac{\Gamma(\lambda+\frac{3}{2})}{\Gamma(\lambda-\frac{1}{2})(2\omega)} + \sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i} \\
&= \frac{(\lambda+\frac{1}{2})(\lambda-\frac{1}{2})}{1!(2\omega)} + \sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i} \\
&= \frac{(4\lambda^2-1)}{1!(8\omega)} + \sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}
\end{aligned}$$

Hence,

$$K_\lambda(\omega) \;=\; \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1+\frac{(4\lambda^2-1)}{1!(8\omega)} + \sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}\right)$$

similarly, expanding the summation for the case when i=2, we obtain

$$\sum_{i=2}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)} = \frac{(\lambda+\frac{5}{2})}{3!\Gamma(\lambda-\frac{3}{2})(2\omega)^3} + \sum_{i=3}^{\lambda-\frac{1}{2}} \frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}$$

and therefore

$$K_\lambda(\omega) = \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left(1+\frac{(4\lambda^2-1)}{1!(8\omega)}+\frac{(4\lambda^2-1)(4\lambda^2-9)}{2!(8\omega^2)}+\sum_{i=3}^{\lambda-\frac{1}{2}}\frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}\right)$$

Similarly

$$\sum_{i=3}^{\lambda-\frac{1}{2}}\frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)} = \frac{(\lambda+\frac{7}{2})}{3!\Gamma(\lambda-\frac{5}{2})(2\omega)^3}+\sum_{i=4}^{\lambda-\frac{1}{2}}\frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}$$

hence

$$\begin{aligned}K_\lambda(\omega) &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\Bigg(1+\frac{(4\lambda^2-1)}{1!(8\omega)}+\frac{(4\lambda^2-1)(4\lambda^2-9)}{2!(8\omega^2)}+\\ &\quad \frac{(4\lambda^2-1)(4\lambda^2-9)(4\lambda^2-25)}{3!(8\omega^3)}+\sum_{i=4}^{\lambda-\frac{1}{2}}\frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}\Bigg)\\ &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\Bigg(1+\frac{(4\lambda^2-1)}{1!(8\omega)}+\prod_{k=1}^{2}\frac{(4\lambda^2-(2k-1)^2}{2!(8\omega^2)}+\prod_{k=1}^{3}\frac{(4\lambda^2-(2k-1)^2)}{3!(8\omega^3)}+\\ &\quad \prod_{k=1}^{4}\frac{(4\lambda^2-(2k-1)^2)}{4!(8\omega^4)}+\sum_{i=5}^{\lambda-\frac{1}{2}}\frac{\Gamma(\lambda+\frac{1}{2}+i)}{i!\Gamma(\lambda+\frac{1}{2}-i)}(2\omega)^{-i}\Bigg)\end{aligned}$$

therefore

$$K_\lambda(\omega) = \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left[1+\sum_{i=1}^{\lambda-\frac{1}{2}}\prod_{k=1}^{i}\frac{(4\lambda^2-(2k-1)^2}{i!(8\omega^i)}\right]$$

We state the following corollary
Corollary 2.4.1

$$\begin{aligned}a)K_{n+\frac{1}{2}}(\omega) &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left[1+\sum_{i=1}^{n}\frac{(n+i)!}{i!(n-i)!}(2\omega)^{-i}\right]\\ b)K_{n-\frac{1}{2}}(\omega) &= \sqrt{\frac{\pi}{2\omega}}e^{-\omega}\left[1+\sum_{i=1}^{n-1}\frac{(n+i-1)!}{i!(n-i-1)!}(2\omega)^{-i}\right]\end{aligned}$$

Proof
a) Where n is a positive integer

From proposition 2.4 put $\lambda = n + \frac{1}{2}$

$$\begin{aligned} K_{n+\frac{1}{2}}(\omega) &= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left[1 + \sum_{i=1}^{n+\frac{1}{2}-\frac{1}{2}} \frac{(n+\frac{1}{2}-\frac{1}{2}+i)!}{i!(n+\frac{1}{2}-\frac{1}{2}-i)!} (2\omega)^{-i} \right] \\ &= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left[1 + \sum_{i=1}^{n} \frac{(n+i)!}{i!(n-i)!} (2\omega)^{-i} \right] \end{aligned}$$

b) Where n is a positive integer
From proposition 2.4 put $\lambda = n - \frac{1}{2}$

$$\begin{aligned} K_{n-\frac{1}{2}}(\omega) &= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left[1 + \sum_{i=1}^{n-\frac{1}{2}-\frac{1}{2}} \frac{(n-\frac{1}{2}+\frac{1}{2}+i)!}{i!(n-\frac{1}{2}+\frac{1}{2}-i)!} (2\omega)^{-i} \right] \\ &= \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left[1 + \sum_{i=1}^{n-1} \frac{(n-1+i)!}{i!(n-1-i)!} (2\omega)^{-i} \right] \end{aligned}$$

we also state the following Corollary
Corollary 2.4.2

$$a) K_{\frac{1}{2}}(\omega) = K_{-\frac{1}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \quad (2.4)$$

$$b) K_{\frac{3}{2}}(\omega) = K_{-\frac{3}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{1}{\omega} \right) \quad (2.5)$$

$$c) K_{\frac{5}{2}}(\omega) = K_{-\frac{5}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{3}{\omega} + \frac{3}{\omega^2} \right) \quad (2.6)$$

$$d) K_{\frac{7}{2}}(\omega) = K_{-\frac{7}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{6}{\omega} + \frac{15}{\omega^2} + \frac{15}{\omega^3} \right) \quad (2.7)$$

$$e) K_{\frac{9}{2}}(\omega) = K_{-\frac{7}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{10}{\omega} + \frac{45}{\omega^2} + \frac{105}{\omega^3} + \frac{105}{\omega^4} \right) \quad (2.8)$$

$$f) K_{\frac{11}{2}}(\omega) = K_{-\frac{9}{2}}(\omega) = \sqrt{\frac{\pi}{2\omega}} e^{-\omega} \left(1 + \frac{15}{\omega} + \frac{105}{\omega^2} + \frac{420}{\omega^3} + \frac{945}{\omega^4} + \frac{945}{\omega^5} \right) \quad (2.9)$$

2.2.3 Definition 3

$$K_\lambda(\omega) = \frac{1}{2} \int_{-\infty}^{\infty} \exp\{-\omega \cosh t\} \cosh \lambda t \, dt \quad (2.10)$$

2.3 Generalised Inverse Gaussian Distribution

From equation (2.1)

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

Using the parametrization

$$\omega = \delta\gamma \tag{2.11}$$

and the transformation

$$x = \frac{\gamma}{\delta}z \tag{2.12}$$

we obtain

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

Thus

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

is a pdf known as Generalised Inverse Gaussian (GIG) distribution. The $r-th$ raw moment of the $GIG(\lambda,\delta,\gamma)$ distribution is given by

$$E(Z^r) = \left(\frac{\delta}{\gamma}\right)^r \frac{K_{\lambda+r}(\delta\gamma)}{K_\lambda(\delta\gamma)} \tag{2.13}$$

and this formula holds for positive and negative values of r.

2.4 Special Cases of Interest

2.4.1 Case 1: $GIG(-\frac{1}{2},\delta,\gamma)$

$$g_1(z) = \frac{\delta}{\sqrt{2\pi}}\exp(\delta\gamma)z^{-\frac{3}{2}}\exp\left\{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)\right\} \tag{2.14}$$

which is the Inverse Gaussian (IG) distribution with the following properties

$$E[Z] = \frac{\delta}{\gamma} \tag{2.15}$$

$$Var(Z) = \frac{\delta}{\gamma^3} \tag{2.16}$$

$$\mu_3(Z) = 3\frac{\delta}{\gamma^5} \tag{2.17}$$

$$\mu_4(Z) = \frac{3\delta^2}{\gamma^6}\left(1 + \frac{5}{\delta\gamma}\right) \tag{2.18}$$

2.4.2 Case 2: $GIG(\frac{1}{2}, \delta, \gamma)$

$$g_2(z) = \frac{\gamma}{\sqrt{2\pi}} \exp(\delta\gamma) z^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\frac{\delta^2}{z} + \gamma^2 z)\right\} \tag{2.19}$$

which is the reciprocal Inverse-Gaussian (RIG) distribution with the following properties

$$E(Z) = \frac{1+\delta\gamma}{\gamma^2} \tag{2.20}$$

$$Var(Z) = \frac{2+\delta\gamma}{\gamma^4} \tag{2.21}$$

$$\mu_3(Z) = \frac{3\delta\gamma + 8}{\gamma^6} \tag{2.22}$$

$$\mu_4(Z) = \frac{3\delta^2\gamma^2 + 27\delta\gamma + 60}{\gamma^8} \tag{2.23}$$

2.4.3 Case 3: $GIG(-\frac{3}{2}, \delta, \gamma)$

$$\begin{aligned} g_3(z) &= \left(\frac{\gamma}{\delta}\right)^{-\frac{3}{2}} \frac{z^{-\frac{3}{2}-1}}{2K_{-\frac{3}{2}}(\delta\gamma)} \exp\left\{ -\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \\ &= \left(\frac{\gamma}{\delta}\right)^{-\frac{3}{2}} \frac{z^{-\frac{5}{2}}}{2\sqrt{\frac{\pi}{2\delta\gamma}} e^{-\delta\gamma}(1+\frac{1}{\delta\gamma})} \exp\left\{ -\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \\ &= \frac{\delta^3}{\sqrt{2\pi}(1+\delta\gamma)} z^{-\frac{5}{2}} e^{\delta\gamma} \exp\left\{ -\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \end{aligned} \tag{2.24}$$

which is the $GIG(-\frac{3}{2},\delta,\gamma)$ with the following properties
Mean

$$\begin{aligned} E(Z) &= \frac{\delta}{\gamma}\frac{K_{\frac{1}{2}}(\delta\gamma)}{K_{\frac{3}{2}}(\delta\gamma)} \\ &= \frac{\delta^2}{1+\delta\gamma} \end{aligned} \tag{2.25}$$

Variance

$$\begin{aligned} var(Z) &= \frac{\delta^3}{\gamma(1+\delta\gamma)} - \frac{\delta^4}{(1+\delta\gamma)^2} \\ &= \frac{\delta^3}{\gamma(1+\delta\gamma)^2} \end{aligned} \tag{2.26}$$

Third Central Moment

$$\begin{aligned} \mu_3(Z) &= 3\frac{\delta^3}{\gamma^3} - \frac{3\delta^5}{\gamma(1+\delta\gamma)^2} + \frac{2\delta^6}{(1+\delta\gamma)^3} \\ &= \frac{\delta^3(1+\delta\gamma)^3 - 3\delta^5\gamma^2 - \delta^6\gamma^3}{\gamma^3(1+\delta\gamma)^3} \end{aligned} \tag{2.27}$$

Fourth Central Moment

$$\begin{aligned} \mu_4(Z) &= \frac{(\delta^5\gamma^2+3\delta^4\gamma+3\delta^3)(1+\delta\gamma)^3 - 4\delta^5\gamma^2(1+\delta\gamma)^3 + 6\delta^7\gamma^4(1+\delta\gamma) - 3\delta^8\gamma^5}{\gamma^5(1+\delta\gamma)^4} \\ &= \frac{15\delta^5\gamma^2 + 3\delta^6\gamma^3 + 12\delta^4\gamma + 3\delta^3}{\gamma^5(1+\delta\gamma)^4} \end{aligned} \tag{2.28}$$

2.4.4 Case 4: $GIG(\frac{3}{2},\delta,\gamma)$

$$g_4(z) = \frac{\gamma^3 e^{\delta\gamma}}{\sqrt{2\pi}(1+\delta\gamma)} z^{\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\frac{\delta^2}{z} + \gamma^2 z)\right\} \tag{2.29}$$

which is the $GIG(\frac{3}{2},\delta,\gamma)$ with the following properties

$$E(Z) = \frac{\delta^2\gamma^2+3\delta\gamma+3}{\gamma^2(1+\delta\gamma)} \tag{2.30}$$

$$Var(Z) = \frac{\delta^3\gamma^3+6\delta^2\gamma^2+12\delta\gamma+6}{\gamma^4(1+\delta\gamma)^2} \tag{2.31}$$

$$\mu_3(Z) = \frac{3\delta^4\gamma^4+25\delta^3\gamma^3+72\delta^2\gamma^2+72\delta\gamma+24}{\gamma^6(1+\delta\gamma)^3} \tag{2.32}$$

$$\mu_4(Z) = \frac{3\delta^6\gamma^6+51\delta^5\gamma^5+336\delta^4\gamma^4+1047\delta^3\gamma^3+1512\delta^2\gamma^2+1008\delta\gamma+252}{\gamma^8(1+\delta\gamma)^4} \tag{2.33}$$

3 Weighted Inverse Gaussian Distributions

The concept of weighted distribution was introduced by Fisher (1934) and elaborated by Patil and Rao (1978). Gupta and Kundu (2011) introduced the weighted inverse gaussian distribution. The objective of this chapter is to show that the special cases discussed in chapter 2 and their pairwise finite mixtures are Weighted Inverse Gaussian (WIG) distributions.

3.1 Definition

Let X be a random variable with pdf $f(x)$. Then a function $w(X)$ 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 we have

$$f_W(x) = \tfrac{w(x)}{E[w(X)]} f(x), -\infty < x < \infty \tag{3.1}$$

3.2 Weighted Inverse Gaussian Distributions

The following examples will be expressed in terms of inverse Gaussian distributions whose pdf is given by formula (2.14)

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

3.2.1 Reciprocal Inverse Gaussian (RIG)

The pdf of RIG is

$$\begin{aligned} g_2(z) &= \frac{\gamma e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \\ &= \frac{\gamma}{\delta} z \left[\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{3}{2}} \exp\left\{ -\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\}\right] \\ &= \frac{\gamma}{\delta} z g_1(z) \end{aligned} \tag{3.2}$$

Therefore

$$GIG\Big(\frac{1}{2},\delta,\gamma\Big) = \frac{\gamma}{\delta}ZGIG\Big(-\frac{1}{2},\delta,\gamma\Big) \tag{3.3}$$

that is, $GIG\Big(\frac{1}{2},\delta,\gamma\Big)$ is a weighted IG distribution with weight

$$w(Z) = Z \tag{3.4}$$

where

$$Z \sim GIG\Big(-\frac{1}{2},\delta,\gamma\Big) \tag{3.5}$$

Using formula (2.13)

$$\therefore E[w(Z)] = E(Z) = \frac{\delta}{\gamma} \tag{3.6}$$

The RIG distribution is called length biased Inverse Gaussian distribution.

3.2.2 $GIG(-\frac{3}{2},\delta,\gamma)$

The pdf is given by

$$\begin{aligned} g_3(z) &= \frac{\delta^3 e^{\delta\gamma}}{\sqrt{2\pi}(1+\delta\gamma)} z^{-\frac{5}{2}} \exp\Big\{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)\Big\} \\ &= \frac{\delta^2}{(1+\delta\gamma)} z^{-1}\Big[\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}} z^{-\frac{3}{2}} \exp\Big\{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)\Big\}\Big] \\ &= \frac{\delta^2}{(1+\delta\gamma)} z^{-1} g_1(z) \end{aligned} \tag{3.7}$$

Thus $GIG(-\frac{3}{2},\delta,\gamma)$ is a weighted inverse Gaussian distribution with weight

$$w(Z) = \frac{1}{Z} \tag{3.8}$$

and by formula (2.13)

$$\therefore E(\frac{1}{Z}) = \frac{1+\delta\gamma}{\delta^2} \tag{3.9}$$

3.2.3 $GIG(\frac{3}{2},\delta,\gamma)$

$$\begin{aligned} g_4(z) &= \frac{\gamma^3}{1+\delta\gamma}\frac{e^{\delta\gamma}}{\sqrt{2\pi}}z^{\frac{1}{2}}e^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)} \\ &= \frac{\gamma^3}{\delta(1+\delta\gamma)}z^2\left[\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}}z^{-\frac{3}{2}}\exp\left\{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)\right\}\right] \\ &= \frac{\gamma^3}{\delta(1+\delta\gamma)}z^2 g_1(z) \end{aligned} \tag{3.10}$$

Thus $GIG(\frac{3}{2},\delta,\gamma)$ is a weighted inverse Gaussian distribution with weight

$$w(Z) = Z^2 \tag{3.11}$$

and by formula (2.13)

$$E[w(Z)] = \frac{\delta(1+\delta\gamma)}{\gamma^3} \tag{3.12}$$

3.3 Cases of Finite Mixtures

Jorgensen, Seshadri and Whitmore (1991) introduced a finite mixture of Inverse Gaussian and Reciprocal Inverse Gaussian distribution. They studied some properties. Further characteristics were obtained by Akman and Gupta (1992), Gupta and Akman (1995); Gupta and Kundu (2011).

Lindley distribution introduced by Lindley (1958) is a finite mixture of exponential and gamma distribution. A detailed study of this one-parameter two component (finite) mixture was given by Gitany et al. (2008). Sankara (1970) had used it as a mixing distribution to Poisson distribution. Shanker, R. and Hogos, F. (2015) applied the Poisson-Lindley distribution to Biological Science data.

Generalizations of the one-parameter Lindley distribution has been studied by researchers. For example Zakerzadeh and Dolati (2010) generalized it to a three parameter Lindley distribution. Mahmoudi and Zakerzadeh (2010) used this three parameter Lindley as a mixing distribution to a Poisson distribution to come up with Poisson-Generalised Lindley distribution.

We have shown that pairwise finite mixtures of $GIG(-\frac{1}{2},\delta,\gamma)$, $GIG(\frac{1}{2},\delta,\gamma)$, $GIG(-\frac{3}{2},\delta,\gamma)$ and $GIG(\frac{3}{2},\delta,\gamma)$ are also weighted Inverse Gaussian distributions. We have given **SIX** finite mixtures including one case by Gupta and his colleagues. More cases can be constructed to form a class called Weighted Inverse Gaussian distributions.

In fact Lindley distribution and its generalizations form another class of weighted distributions, namely, a class of weighted Exponential distributions.

In the next section, we define $g_{ij}(z)$ for $i,j = 1,2,3,4$ but $i \neq j$ to be a finite mixture of $g_i(z)$ and $g_j(z)$.

3.3.1 Case 1

We consider finite Mixture of GIG for indexes $-\frac{1}{2}$ and $\frac{1}{2}$

Let

$$Z_1 \sim GIG\big(-\frac{1}{2},\delta,\gamma\big)$$

and

$$Z_2 \sim GIG\big(\frac{1}{2},\delta,\gamma\big)$$

Then the pdf of the finite mixture is

$$\begin{aligned} g_{12}(z) &= pg_1(z)+(1-p)g_2(z) \\ &= pg_1(z)+(1-p)\frac{\gamma}{\delta}g_1(z) \\ &= \big[p+(1-p)\frac{\gamma}{\delta}\big]g_1(z) \end{aligned} \tag{3.13}$$

Let

$$p=\frac{\gamma}{\gamma+\delta} \tag{3.14}$$

$$\therefore g_{12}(z)=\frac{\gamma}{\gamma+\delta}(1+z)g_1(z) \tag{3.15}$$

Thus the finite mixture of IG and RIG (length biased Inverse Gaussian) distribution is a weighted Inverse Gaussian distribution with weight

$$w(Z)=1+Z \tag{3.16}$$

and by formula (2.13)

$$E[w(Z)]=\frac{\gamma+\delta}{\gamma} \tag{3.17}$$

with the following properties

$$E[Z]=\frac{\delta+\delta\gamma(\gamma+\delta)}{\gamma^2(\gamma+\delta)} \tag{3.18}$$

and

$$var(Z)=\frac{3\delta\gamma+2\delta^2\gamma^2+\delta\gamma^3+2\delta^2+\delta^3\gamma}{\gamma^4(\delta+\gamma)^2} \tag{3.19}$$

3.3.2 Case 2

We consider finite Mixture of GIG for indexes $-\frac{1}{2}$ and $-\frac{3}{2}$.

Let

$$Z_3 \sim GIG\big(-\frac{3}{2},\delta,\gamma\big)$$

Then

$$\begin{aligned} g_{13}(z) &= pg_1(z)+(1-p)g_3(z) \\ &= \big[p+(1-p)\frac{\delta^2}{1+\delta\gamma}\big]g_1(z) \end{aligned}$$

Put

$$p=\frac{\delta^2}{(1+\delta^2)} \tag{3.20}$$

$$\therefore g_{13}(z)=\frac{\delta^2}{1+\delta^2}\big(1+\frac{1}{1+\delta\gamma}\frac{1}{z}\big)g_1(z) \tag{3.21}$$

Thus the finite mixture of $GIG\big(-\frac{1}{2},\delta,\gamma\big)$ and $GIG\big(-\frac{3}{2},\delta,\gamma\big)$ is a weighted Inverse Gaussian distribution with weights

$$w(Z)=1+\frac{1}{1+\delta\gamma}\frac{1}{Z} \tag{3.22}$$

and using formula (2.13)

$$E[w(Z)]=\frac{1+\delta^2}{\delta^2} \tag{3.23}$$

The mean and variance for the weighted distribution are

$$E[Z]=\frac{\delta^2}{1+\delta^2}\left[\frac{\delta(1+\delta\gamma)+\gamma}{\gamma(1+\delta\gamma)}\right] \tag{3.24}$$

$$\begin{aligned} var(Z) &= \frac{\delta^2}{1+\delta^2}\left[\frac{\delta\gamma^2+\delta(1+\delta\gamma)^2}{\gamma^3(1+\delta\gamma)}-\frac{\delta^2}{1+\delta^2}\frac{(\delta(1+\delta\gamma)+\gamma)^2}{\gamma^2(1+\delta\gamma)^2}\right] \\ &= \frac{\delta^3(\gamma^2+(1+\delta\gamma)^2)(1+\delta^2)(1+\delta\gamma)-\delta^4\gamma(\delta(1+\delta\gamma)+\gamma)^2}{\gamma^3(1+\delta\gamma)^2(1+\delta^2)^2} \end{aligned} \tag{3.25}$$

3.3.3 Case 3

We consider a finite mixture between $GIG\big(-\frac{1}{2},\delta,\gamma\big)$ and $GIG\big(\frac{3}{2},\delta,\gamma\big)$

$$\begin{aligned} g_{14}(z) &= pg_1(z)+(1-p)g_4(z) \\ &= \big[p+(1-p)\frac{\gamma^3}{\delta}\frac{z^2}{(1+\delta\gamma)}\big]g_1(z) \end{aligned}$$

Put

$$p = \frac{\gamma^3}{\gamma^3 + \delta} \tag{3.26}$$

$$\therefore g_{14}(z) = \frac{\gamma^3}{\gamma^3 + \delta}\left(1 + \frac{z^2}{1 + \delta\gamma}\right) g_1(z) \tag{3.27}$$

Thus the finite mixture of $GIG\big(-\frac{1}{2}, \delta, \gamma\big)$ and $GIG\big(\frac{3}{2}, \delta, \gamma\big)$ is a weighted Inverse Gaussian with weight

$$w(Z) = 1 + \frac{Z^2}{(1 + \delta\gamma)} \tag{3.28}$$

and using formula (2.13)

$$E[w(Z)] = \frac{\gamma^3 + \delta}{\gamma^3} \tag{3.29}$$

The mean and variance for the weighted distribution are

$$\begin{aligned} E[Z] &= \frac{\gamma^3}{\gamma^3 + \delta}\left[\frac{\delta\gamma^4(1 + \delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta}{\gamma^5(1 + \delta\gamma)}\right] \\ &= \frac{\delta\gamma^4(1 + \delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta}{\gamma^2(1 + \delta\gamma)(\gamma^3 + \delta)} \end{aligned} \tag{3.30}$$

$$\begin{aligned} var(Z) &= \frac{(\delta\gamma^6(1 + \delta\gamma)^2 + \delta^5\gamma^4 + 10\delta^4\gamma^3 + 45\delta^3\gamma^2 + 105\delta^2\gamma + 105\delta)(1 + \delta\gamma)(\gamma^3 + \delta)}{\gamma^6(1 + \delta\gamma)^2(\gamma^3 + \delta)} \\ &\quad - \frac{(\delta\gamma^4(1 + \delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta)^2}{\gamma^6(1 + \delta\gamma)^2(\gamma^3 + \delta)} \end{aligned} \tag{3.31}$$

3.3.4 Case 4

We consider a finite mixture between $GIG\big(\frac{1}{2}, \delta, \gamma\big)$ and $GIG\big(-\frac{3}{2}, \delta, \gamma\big)$.

$$\begin{aligned} g_{23}(z) &= p g_2(z) + (1 - p) g_3(z) \\ &= \left[p\frac{\gamma}{\delta}z + (1 - p)\frac{\delta^2}{1 + \delta\gamma}\frac{1}{z}\right] g_1(z) \end{aligned}$$

put

$$p = \frac{\delta^3}{\delta^3 + \gamma} \tag{3.32}$$

$$\begin{aligned} \therefore g_{23}(z) &= \left[\frac{\gamma\delta^2}{\delta^3 + \gamma}z + \frac{\gamma\delta^2}{\delta^3 + \gamma}\frac{1}{z}\right] g_1(z) \\ &= \frac{\gamma\delta^2}{\delta^3 + \gamma}\left(z + \frac{1}{1 + \delta\gamma}\frac{1}{z}\right) g_1(z) \end{aligned} \tag{3.33}$$

Thus a finite mixture of $GIG\big(\frac{1}{2},\delta,\gamma\big)$ and $GIG\big(-\frac{3}{2},\delta,\gamma\big)$ is a weighted inverse gaussian distribution with weight

$$w(Z) = Z + \frac{1}{1+\delta\gamma}\frac{1}{Z} \tag{3.34}$$

and using formula (2.13)

$$E[w(Z)] = \frac{\delta^3+\gamma}{\gamma\delta^2} \tag{3.35}$$

The mean and variance for the weighted distribution are

$$E[Z] = \frac{\delta^3(1+\delta\gamma)^2+\delta^2\gamma^3}{\gamma^2(1+\delta\gamma)(\delta^3+\gamma)} \tag{3.36}$$

and

$$var(Z) = \frac{\delta^2((\delta^3\gamma^2+3\delta^2\gamma+3\delta)(1+\delta\gamma)+\delta\gamma^4)(1+\delta\gamma)(\delta^3+\gamma)-\delta^4(\delta(1+\delta\gamma)^2+\gamma^3)^2}{\gamma^4(1+\delta\gamma)^2(\delta^3+\gamma)^2} \tag{3.37}$$

3.3.5 Case 5

We consider a finite mixture between $GIG\big(\frac{1}{2},\delta,\gamma\big)$ and $GIG\big(\frac{3}{2},\delta,\gamma\big)$.

$$\begin{aligned} g_{24}(z) &= pg_2(z)+(1-p)g_4(z) \\ &= [p\frac{\gamma}{\delta}z+(1-p)\frac{\gamma^3}{\delta(1+\delta\gamma)}z^2]g_1(z) \end{aligned}$$

Put

$$p = \frac{\gamma^2}{\gamma^2+1} \tag{3.38}$$

$$\begin{aligned} \therefore g_{24}(z) &= \left[\frac{\gamma^3}{\delta(\gamma^2+1)}z+\frac{\gamma^3}{\delta(\gamma^2+1)}\frac{z^2}{(1+\delta\gamma)}\right]g_1(z) \\ &= \frac{\gamma^3}{\delta(\gamma^2+1)}\left[z+\frac{z^2}{1+\delta\gamma}\right]g_1(z) \end{aligned} \tag{3.39}$$

Thus we have a finite mixture of $GIG\big(\frac{1}{2},\delta,\gamma\big)$ and $GIG\big(\frac{3}{2},\delta,\gamma\big)$ which is a weighted Inverse Gaussian distribution with weights

$$w(Z) = Z + \frac{Z^2}{1+\delta\gamma} \tag{3.40}$$

and using formula (2.13)

$$E[w(Z)] = \frac{\delta(\gamma^2+1)}{\gamma^3} \tag{3.41}$$

The mean and variance for the weighted distribution are

$$E[Z] = \frac{\gamma^2(1+\delta\gamma)^2+\delta^2\gamma^2+3\delta\gamma+3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)} \tag{3.42}$$

$$var(Z) = \frac{((\delta^2\gamma^2+3\delta\gamma+3)(1+\delta\gamma)\gamma^2+(\delta^3\gamma^3+6\delta^2\gamma^2+15\delta\gamma+15))(1+\delta\gamma)(\gamma^2+1)}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2} - \frac{(\gamma^2(1+\delta\gamma)^2+\delta^2\gamma^2+3\delta\gamma+3)^2}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2} \tag{3.43}$$

3.3.6 Case 6

We consider a finite mixture between $GIG\big(-\frac{3}{2},\delta,\gamma\big)$ and $GIG\big(\frac{3}{2},\delta,\gamma\big)$.

$$\begin{aligned} g_{34}(z) &= pg_3(z)+(1-p)g_4(z) \\ &= \left[p\frac{\delta^2}{1+\delta\gamma}\frac{1}{z}+(1-p)\frac{\gamma^3}{\delta}\left(\frac{1}{1+\delta\gamma}\right)z^2\right] \end{aligned}$$

Put

$$p = \frac{\gamma^3}{\gamma^3+\delta^3} \tag{3.44}$$

$$\therefore g_{34}(z) = \left[\frac{\gamma^3\delta^2}{\gamma^3+\delta^3}\frac{1}{1+\delta\gamma}\frac{1}{z}+\frac{\gamma^3\delta^2}{\gamma^3+\delta^3}\frac{1}{1+\delta\gamma}z^2\right]g_1(z)$$

$$\therefore g_{34}(z) = \frac{\gamma^3\delta^2}{\gamma^3+\delta^3}\left[\frac{1}{z}+z^2\right]g_1(z) \tag{3.45}$$

Thus a finite mixture of$GIG\big(-\frac{3}{2},\delta,\gamma\big)$ and $GIG\big(\frac{3}{2},\delta,\gamma\big)$ is a weighted Inverse Gaussian distriburtion with weight

$$w(Z) = \frac{1}{Z}+Z^2 \tag{3.46}$$

and using formula (2.13)

$$E[w(Z)] = \frac{\gamma^3+\delta^3}{\gamma^3\delta^2} \tag{3.47}$$

The mean and variance for the weighted distribution are

$$E[Z] = \delta^2\left(\frac{1}{1+\delta\gamma}+\frac{3\delta}{\gamma^2(\gamma^3+\delta^3)}\right) \tag{3.48}$$

and

$$var(Z) = \frac{3(1+\delta\gamma)^2\delta^3[5\gamma^3+2\delta^3]+\delta^3\gamma^9+2\delta^6\gamma^6+\delta^9\gamma^3}{\gamma^4(\gamma^3+\delta^3)^2(1+\delta\gamma)} \tag{3.49}$$

In this chapter we have proposed a new class of distribution known as Weighted Inverse Gaussian (WIG) distributions. The class will be used as mixing distributions in the next two chapters to a bew class of normal variance mean mixtures.

4 Normal Weighted Inverse Gaussian Distribution Part I

4.1 Introduction

A normal distribution has two parameters: the location parameter representing the mean and the scale parameter representing the variance. For a continuous mixture, we can fix the mean and vary the variance and vice-versa. Barndorff-Nielsen (1977) introduced a normal mixture where the mean is a linear function of a varying variance. This is called a Normal Variance-Mean mixture (NVMM).
The objectives of this chapter are:

i to give NVMM in a stochastic and hierarchial representation

ii to express the properties of NVMM in terms of properties of mixing distribution

iii to construct and obtain NVMMs with the following special cases of GIG distribution as mixing distributions: $GIG\left(-\frac{1}{2},\delta,\gamma\right)$ which is an Inverse Gaussian mixing distribution, $GIG\left(-\frac{1}{2},\delta,\gamma\right)$ which is the Reciprocal Inverse Gaussian mixing distribution, $GIG\left(-\frac{3}{2},\delta,\gamma\right)$ and $GIG\left(\frac{3}{2},\delta,\gamma\right)$

4.2 Hierarchial Representation of NVMM

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

$$X = \mu + \beta Z + \sqrt{Z}Y$$
$$\textit{where}$$
$$Y \sim N(0,1)$$

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

4.3 Stochastic Representation of NVMM

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

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

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

$$\begin{aligned} \therefore f(x) &= \int_0^\infty \frac{1}{\sqrt{z}} \phi\left(\frac{x-\mu-\beta z}{\sqrt{z}}\right) g(z)\, dz \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi z}} e^{-\frac{[x-(\mu+\beta z)]^2}{2z}} g(z)\, dz \end{aligned} \tag{4.1}$$

Thus we have a hierarchical representation as

$$X/Z = z \sim N(\mu + \beta z, z) \tag{4.2}$$

being the conditional pdf and g(z) the mixing distribution.
When the mixing distribution is Generalized Inverse Gaussian (GIG), then the mixture is called Generalized Hyperbolic Distribution (GHD) which nests a number of distributions obtained as special and limiting cases. Special cases are obtained when the index parameter, λ takes specific values. When $\lambda = 1$ we obtain the hyperbolic distribution which was the first special case used in financial modelling (Eberlein and Keller, 1995). Later on (Barndorff-Nielsen, 1999) introduced the case when $\lambda = -\frac{1}{2}$ which is the Normal Inverse Gaussian (NIG) distribution. The *NIG* has been extensively studied in finance because of its analytical tractability property. We extend this work by considering the other special cases mentioned in chapter 4. These models will be alternative distributions to NIG for application in finance.

4.4 Properties for Normal Variance Mean Mixtures

One of the features of constructing a distribution by mixing is that one can essentially read off the properties of the distribution given the properties of the chosen weight.

Proposition 4.4.1.

$$M_X(t) = e^{\mu t} M_Z(\beta t + \frac{t^2}{2}) \tag{4.3}$$

$$E(X) = \mu + \beta E(Z) \tag{4.4}$$

$$Var(X) = E(Z) + \beta^2 Var(Z) \tag{4.5}$$

$$\mu_3(X) = 3\beta var(Z) + \beta^3 \mu_3(Z) \tag{4.6}$$

$$\mu_4(X) = \beta^4 \mu_4(Z) + 6\beta^2 \mu_3(Z) + 6\beta^2 E[Z] var(Z) + 3E[Z^2] \tag{4.7}$$

Proof. **Moment Generating Function**

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\ &= \int_0^{\infty} \int_{-\infty}^{\infty} e^{tx} f_N(x; \mu + \beta z, z) dx f_Z(z) dz \\ &= e^{\mu t} \int_0^{\infty} exp\left[\left(\beta t + \frac{t^2}{2}\right) z\right] f_Z(z) dz \\ &= e^{\mu t} M_Z\left(\beta t + \frac{t^2}{2}\right) \end{aligned}$$

Mean

The mean of the mixed distribution, can be obtained by evaluating the first derivative of (1) at t=0.

$$M_X^I(t) = e^{\mu t} E\left[(\beta + t) Z e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right] + \mu e^{\mu t} E\left[e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right]$$

$$M_X'(0) = \mu + \beta E[Z]$$

Therefore

$$E[X] = \mu + \beta E[Z] \tag{4.8}$$

as obtained by (Paolella, 2007) using the conditional expectation approach.

Variance

$$\begin{aligned} M_X''(t) &= e^{\mu t} E\left[(\beta + t) Z e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right] + \mu e^{\mu t} E\left[(\beta + t) Z e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right] + \mu^2 e^{\mu t} E\left[e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right] + \\ &\quad \mu e^{\mu t} E\left[(\beta + t) Z e^{\left(\beta t + \frac{t^2}{2}\right) Z}\right] \end{aligned}$$

Therefore,

$$\begin{aligned} E[X^2] &= M_X''(0) \\ &= E[Z] + \mu^2 + \beta^2 E[Z^2] + 2\mu\beta E[Z] \end{aligned}$$

$$\begin{aligned} var(X) &= E[Z] + \mu^2 + \beta^2 E[Z^2] + 2\mu\beta E[Z] - (\mu + \beta E[X])^2 \\ &= E[Z] + \beta^2 \left[E(Z^2) - E(Z)^2 \right] \\ &= E[Z] + \beta^2 var(Z) \end{aligned}$$

as obtained by (Paolella, 2007) using the conditional expectation approach.

Third Central Moment
Note that

$$M_X^{III}(t) = 3\mu e^{\mu t} E\left[(\beta+t)^2 Z^2 e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right] + 3\mu e^{\mu t} E\left[Z e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right] + 3e^{\mu t} E\left[(\beta+t) Z^2 e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right] + 3\mu^2 e^{\mu t} E\left[(\beta+t) Z e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right] + e^{\mu t} E\left[(\beta+t)^3 Z^3 e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right] + \mu^3 e^{\mu t} E\left[e^{\left(\beta t + \frac{t^2}{2}\right)Z}\right]$$

Therefore

$$E[X^3] = \mu^3 + 3\mu\beta^2 E[Z^2] + 3\mu E[Z] + 3\beta E[Z^2] + 3\beta\mu^2 E[Z] + \beta^3 E[Z^3]$$

the terms,

$$\begin{aligned} E(X^2)E(X) &= (E[Z] + \mu^2 + \beta^2 E[Z^2] + 2\mu\beta E[Z])(\mu + \beta E[Z]) \\ &= \mu E[Z] + \mu^3 + \mu\beta^2 E[Z^2] + 2\mu^2\beta E[Z] + \beta E[Z]^2 + \mu^2\beta E[Z] + \beta^3 E[Z^2]E[Z] + \\ &\quad 2\mu\beta^2 E[Z]^2 \end{aligned}$$

$$E[X]^3 = \mu^3 + 3\mu^2\beta E[Z] + 3\mu\beta^2 E(Z)^2 + \beta^3 E[Z]^3$$

Since

$$\mu_3(X) = E(X^3) - 3E(X^2)E(X) + 2E(X)^3$$

We find that,

$$\begin{aligned}\mu_3(X) &= 3\beta(E[Z^2]-E[Z]^2)+\beta^3(E[Z^3]-3E[Z^2]E[Z]+2E[Z]^3) \\ &= 3\beta var(Z)+\beta^3\mu_3(Z)\end{aligned}$$

Fourth Central Moment
We start by stating the following theorem,

$$\mu_4(X)=\beta^4\mu_4(Z)+6\beta^2\mu_3(Z)+6\beta^2E(Z)var(Z)+3E(Z^2) \tag{4.9}$$

Proof

$$\begin{aligned}M_X^{IV}(t) = {} & 3\mu^2e^{\mu t}E\left[(\beta+t)^2Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+3\mu e^{\mu t}E\left[(\beta+t)^3Z^3e^{\left(\beta t+\frac{t^2}{2}\right)Z}+2(\beta+t)Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+ \\ & 3\mu^2e^{\mu t}E\left[Ze^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+3\mu e^{\mu t}E\left[(\beta+t)Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+3\mu e^{\mu t}E\left[(\beta+t)Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+ \\ & 3e^{\mu t}E\left[(\beta+t)^2Z^3e^{\left(\beta t+\frac{t^2}{2}\right)Z}+Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+3\mu^3e^{\mu t}E\left[(\beta+t)Ze^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+ \\ & 3\mu^2e^{\mu t}E\left[(\beta+t)^2Z^2e^{\left(\beta t+\frac{t^2}{2}\right)Z}+Ze^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+\mu e^{\mu t}E\left[(\beta+t)^3Z^3e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+ \\ & e^{\mu t}E\left[(\beta+t)^4Z^4e^{\left(\beta t+\frac{t^2}{2}\right)Z}+3(\beta+t)^2Z^3e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+\mu^4e^{\mu t}E\left[e^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]+ \\ & \mu^3e^{\mu t}E\left[(\beta+t)Ze^{\left(\beta t+\frac{t^2}{2}\right)Z}\right]\end{aligned}$$

Therefore

$$\begin{aligned}E[X^4] = {} & \mu^4+4\mu^3\beta E[Z]+6\mu^2\beta^2E[Z^2]+4\mu\beta^3E[Z^3]+12\mu\beta E[Z^2]+6\mu^2E[Z]+6\beta^2E[Z^3]+ \\ & 3E[Z^2]+\beta^4E[Z^4]\end{aligned}$$

Also note that

$$\begin{aligned}\mu_4(X) &= E[(X-E(x))^4] \\ &= E[X^4]-4E[X^3]E[X]+6E[X^2]E[X]^2-3E[X]^4\end{aligned}$$

Where the terms:

$$\begin{aligned}E(X)E(X^3) &= (\mu+\beta E[Z])(\mu^3+3\mu\beta^2E[Z^2]+3\mu E[Z]+3\beta E[Z^2]+3\beta\mu^2E[Z]+\beta^3E[Z^3]) \\ &= \mu^4+3\mu^2\beta^2E[Z^2]+3\mu^2E[Z]+3\mu\beta E[Z^2]+3\mu^3\beta E[Z]+\mu\beta^3E[Z^3]+\mu^3\beta E[Z]+ \\ & \quad 3\mu\beta^3E[Z^2]E[Z]+3\mu\beta E[Z]^2+3\beta^2E[Z^2]E[Z]+3\mu^2\beta^2E[Z]^2+\beta^4E[Z^3]E[Z]\end{aligned}$$

$$\begin{aligned} E(X)^2E(X^2) &= \mu^2E[Z]+\mu^4+\mu^2\beta^2E[Z^2]+2\mu^3\beta E[Z]+2\mu\beta E[Z]^2+2\mu^3\beta E[Z]+2\mu\beta E[Z]^2+ \\ &\quad 2\mu^3\beta E[Z]+2\mu\beta^3E[Z^2]E[Z]+4\mu^2\beta^2E[Z]^2+\beta^2E[Z]^3+\mu^2\beta^2E[Z]^2+ \\ &\quad \beta^4E[Z^2]E[Z]^2+2\mu\beta^3E[Z]^3 \end{aligned}$$

$$\begin{aligned} E[X]^4 &= (\mu+\beta E[Z])^4 \\ &= \mu^4+4\mu^3\beta E[Z]+6\mu^2+\beta^2E[Z]^2+4\mu\beta^3E[Z]^3+\beta^4E[Z]^4 \end{aligned}$$

Therefore,

$$\begin{aligned} \mu_4(X) &= \beta^4E[Z^4]+6\beta^2E[Z^3]+3E[Z^2]-12\beta^2E[Z^2]E[Z]-4\beta^4E[Z^3]E[Z]+6\beta^2E[Z]^3+ \\ &\quad 6\beta^4E[Z^2]E[Z]^2-3\beta^4E[Z]^4 \\ &= \beta^4E[Z^4]-4\beta^4E[Z^3]E[Z]+6\beta^4E[Z^2]E[Z]^2-3\beta^4E[Z]^4+6\beta^2E[Z^3]+12\beta^2E[Z]^3- \\ &\quad 18\beta^2E[Z^2]E[Z]+6\beta^2E[Z^2]E[Z]-6\beta^2E[Z]^3+3E[Z^2] \end{aligned}$$

noting that;

$$E[(Z-E[Z])^4]=E[Z^4]-4E[Z^3]E[Z]+6E[Z^2]E[Z]^2-3E[Z]^4$$

we get,

$$\begin{aligned} \mu_4(X) &= \beta^4[E[Z^4]-4E[Z^3]E[Z]+6E[Z^2]E[Z]^2-3E[Z]^4]+6\beta^2[E[Z^3]-3E[Z^2]E[Z]+ \\ &\quad 2E[Z]^3]+6\beta^2E[Z][E[Z^2]-E[Z]^2]+3E[Z^2] \end{aligned}$$

Finally then,

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

□

4.5 Normal Inverse Gaussian Distribution

Let the mixing distribution of z be Inverse Gaussian distribution. The mixed model is therefore,

$$\begin{aligned} f(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi z}}e^{-\frac{1}{2}\frac{(x-(\mu+\beta z))^2}{z}}g(z)\,dz \\ &= \frac{\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{2\pi}\int_0^\infty z^{-2}exp\left\{-\frac{1}{2}\left[\frac{(x-\mu)^2-\delta^2}{z}+(\beta^2+\gamma^2)z\right]\right\}dz \\ &= \frac{\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{2\pi}\int_0^\infty z^{-2}\exp\{-\frac{(\beta^2+\gamma^2)}{2}\left[z+\frac{(x-\mu)^2-\delta^2}{(\beta^2+\gamma^2)}\frac{1}{z}\right]\}\,dz \end{aligned}$$

Let

$$z = \sqrt{\frac{(x-\mu)^2 - \delta^2}{(\beta^2+\gamma^2)} t}$$

Therefore

$$\begin{aligned} f(x) &= \frac{\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{2\pi} \left(\sqrt{\frac{(x-\mu)^2 - \delta^2}{(\beta^2+\gamma^2)}} \right)^{-1} \\ &\quad \times \int_0^\infty t^{-1-1} \exp\left\{ -\frac{\sqrt{(\beta^2+\gamma^2)((x-\mu)^2 - \delta^2)}}{2} \left(t + \frac{1}{t} \right) \right\} dt \end{aligned}$$

denote $\alpha^2 = \beta^2 + \gamma^2$

$$f(x) = \frac{\alpha}{\pi} \exp\{\delta\sqrt{\alpha^2 - \beta^2} - \beta\mu\} \exp\{\beta x\} \phi(x)^{-\frac{1}{2}} K_1 \left(\alpha\delta\phi(x)^{\frac{1}{2}} \right) \tag{4.10}$$

which is the Normal Inverse-Gaussian distribution, where

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

4.5.1 Properties of the NIG

Mean

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

Since here $Z\ GIG(-\frac{1}{2}, \delta, \gamma)$

$$E(Z) = \frac{\delta}{\gamma} \frac{K_{\frac{1}{2}}(\delta\gamma)}{K_{\frac{1}{2}}(\delta\gamma)} = \frac{\delta}{\gamma}$$

therefore

$$E(X) = \mu + \beta \frac{\delta}{\gamma} \tag{4.11}$$

Variance

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

Where

$$\begin{aligned} var(Z) &= \frac{\delta^2}{\gamma^2}\left(1+\frac{1}{\delta\gamma}\right) - \frac{\delta^2}{\gamma^2} \\ &= \frac{\delta}{\gamma^3} \end{aligned}$$

and hence

$$\begin{aligned} var(X) &= \frac{\delta}{\gamma} + \beta^2 \frac{\delta}{\gamma^3} \\ &= \frac{\alpha^2\delta}{\gamma^3} \end{aligned} \tag{4.12}$$

Skewness

Note

$$\mu_X(3) = 3\beta var(Z) + \beta^3 \mu_Z(3)$$

Where

$$\begin{aligned} \mu_Z(3) &= E(Z^3) - 3E(Z^2)E(Z) + E(Z)^3 \\ &= 3\frac{\delta}{\gamma^3} \end{aligned}$$

hence

$$\mu_X(3) = \frac{3\beta\delta\alpha^2}{\gamma^5}$$

Therefore, denoting skewness by γ_1 we get

$$\begin{aligned} \gamma_1 &= \frac{\mu_3(X)}{(\mu_2(X))^{1.5}} \\ &= \frac{3\beta}{\alpha(\delta\gamma)^{\frac{1}{2}}} \end{aligned} \tag{4.13}$$

Excess Kurtosis

Using formula (4.7),

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

where for the IG distribution

$$\begin{aligned}\mu_4(Z) &= E(Z^4) + 4E(Z^3)E(Z) + 6E(Z^2)E(Z)^2 - E(Z)^4 \\ &= 3\frac{\delta^4}{\gamma^4}\frac{1}{\delta^2\gamma^2}\left(1+\frac{1}{\delta\gamma}\right)\end{aligned}$$

Therefore

$$\begin{aligned}\mu_4(X) &= 3\beta^4\frac{\delta^4}{\gamma^4}\frac{1}{\delta^2\gamma^2}\left(1+\frac{1}{\delta\gamma}\right) + 18\beta^2\frac{\delta}{\gamma^5} + 6\beta^2\frac{\delta}{\gamma}\left(\frac{\delta}{\gamma^3}\right) + 3\frac{\delta^2}{\gamma^2}(1+\frac{1}{\delta\gamma}) \\ &= \frac{6\beta^2\delta^2\gamma^3 + 18\beta^2\delta\gamma^5 + 3\delta\gamma^4 + 3\beta^4\delta^2\gamma + 15\beta^4\delta}{\gamma^7}\end{aligned}$$

then working out kurtosis gives

$$\begin{aligned}\frac{\mu_4(X)}{(\mu_2(X))^2} &= \frac{6\beta^2\delta^2\gamma^3 + 18\beta^2\delta\gamma^5 + 3\delta\gamma^4 + 3\beta^4\delta^2\gamma + 15\beta^4\delta}{\gamma^7} \times \frac{\gamma^6}{\alpha^4\delta^2} \\ &= \frac{6\beta^2\delta\gamma^3 + 18\beta^2\gamma^5 + 3\gamma^4 + 3\beta^4\delta\gamma + 15\beta^4}{\alpha^4\delta\gamma}\end{aligned}$$

Denoting excess kurtosis by γ_2 we obtain

$$\begin{aligned}\gamma_2 &= \frac{\mu_4(X)}{(\mu_2(X))^2} - 3 \\ &= \frac{6\beta^2\delta\gamma^3 + 18\beta^2\gamma^5 + 3\gamma^4 + 3\beta^4\delta\gamma + 15\beta^4 - 3\alpha^4\delta\gamma}{\alpha^4\delta\gamma}\end{aligned}$$

note $\alpha^4 = \beta^4 + 2\beta^2\gamma^2 + \gamma^4$

$$\begin{aligned}\gamma_2 &= \frac{6\beta^2\delta\gamma^3 + 18\beta^2\gamma^2 + 3\delta\gamma^5 + 3\gamma^4 + 3\beta^4\delta\gamma + 15\beta^4 - 3(\beta^4 + 2\beta^2\gamma^2 + \gamma^4)\delta\gamma}{\alpha^4\delta\gamma} \\ &= \frac{3\left[1+4\frac{\beta^2}{\alpha^2}\right]}{\delta\gamma} \qquad (4.14)\end{aligned}$$

as obtained by (Karlis, 2000). Therefore in summary

Table 4.1: Properties of the NIG: GHD($\lambda = -\frac{1}{2}, \delta, \gamma$)

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	$\frac{3\left(1+4\frac{\beta^2}{\alpha^2}\right)}{\delta\gamma}$

4.6 Normal Reciprocal Inverse Gaussian Distribution

Now suppose that the mixing distribution is the Reciprocal inverse Gaussian (RIG) Distribution, the mixed model becomes,

$$f(x) = \frac{\gamma \exp(\delta\gamma) e^{\beta(x-\mu)}}{2\pi} \int_0^\infty z^{-1} \exp - \frac{(\beta^2+\gamma^2)}{2} [z + \frac{(x-\mu)^2+\delta^2}{(\beta^2+\gamma^2)} \frac{1}{z}] \, dz$$

let

$$z = \sqrt{\frac{(x-\mu)^2+\delta^2}{(\beta^2+\gamma^2)}} t$$

$$\begin{aligned} f(x) &= \frac{\gamma \exp(\delta\gamma) e^{\beta(x-\mu)}}{2\pi} \int_0^\infty t^{0-1} exp\left\{ - \frac{\sqrt{(\beta^2+\gamma^2)(x-\mu)^2+\delta^2}}{2} \left(t + \frac{1}{t}\right) \right\} dt \\ &= \frac{\gamma \exp(\delta\gamma - \beta\mu) \exp(\beta x)}{\pi} K_0(\alpha\delta\phi(x)^{\frac{1}{2}}) \end{aligned} \quad (4.15)$$

which is the Normal Reciprocal Inverse-Gaussian (NRIG) distribution.

4.6.1 Properties of the *NRIG* distribution

Mean

For the *NRIG* distribution, $Z \sim GIG(\frac{1}{2}, \delta, \gamma)$
therefore

$$\begin{aligned} E(Z) &= \frac{\delta}{\gamma}\frac{K_{\frac{3}{2}}(\delta\gamma)}{K_{\frac{1}{2}}(\delta\gamma)} \\ &= \frac{1+\delta\gamma}{\gamma^2} \end{aligned}$$

hence

$$E(X) = \mu + \frac{\beta(1+\delta\gamma)}{\gamma^2} \tag{4.16}$$

Variance

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

where

$$\begin{aligned} var(Z) &= \frac{\delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^4} - \frac{(1+\delta\gamma)^2}{\gamma^4} \\ &= \frac{2+\delta\gamma}{\gamma^4} \end{aligned}$$

hence

$$\begin{aligned} var(X) &= \frac{\alpha^2 + \beta^2 + \delta\gamma(\gamma^2 + \beta^2)}{\gamma^4} \\ &= \frac{\alpha^2(1+\delta\gamma) + \beta^2}{\gamma^4} \end{aligned} \tag{4.17}$$

Skewness

Note

$$\begin{aligned} E(Z^3) &= \left(\frac{\delta}{\gamma}\right)^3 \frac{K_{\frac{7}{2}}(\delta\gamma)}{K_{\frac{1}{2}}(\delta\gamma)} \\ &= \frac{\delta^3\gamma^3 + 6\delta^2\gamma^2 + 15\delta\gamma + 15}{\gamma^6} \end{aligned}$$

and hence

$$\begin{aligned}\mu_3(Z) &= \frac{\delta^3\gamma^3+6\delta^2\gamma^2+15\delta\gamma+15-3(1+\delta\gamma)(\delta^2\gamma^2+3\delta\gamma+3)+2(1+\delta\gamma)^3}{\gamma^6} \\ &= \frac{3\delta\gamma+8}{\gamma^6}\end{aligned}$$

therefore

$$\begin{aligned}\mu_3(X) &= \frac{3\beta(\delta\gamma+2)}{\gamma^4}+\frac{\beta^3(3\delta\gamma+8)}{\gamma^6} \\ &= \frac{3\beta\alpha^2(\delta\gamma+2)+2\beta^3}{\gamma^6}\end{aligned}$$

We obtain skewness to be

$$\gamma_1 = \frac{3\beta\alpha^2(\delta\gamma+2)+2\beta^3}{(\alpha^2(1+\delta\gamma)+\beta^2)^{\frac{3}{2}}} \tag{4.18}$$

Excess Kurtosis

Using formula (4.7),

$$\begin{aligned}\mu_4(Z) &= \frac{\delta^4\gamma^4+10\delta^3\gamma^3+45\delta^2\gamma^2+105\delta\gamma+105}{\gamma^8}-\frac{4(\delta^3\gamma^3+6\delta^2\gamma^2+15\delta\gamma+15)}{\gamma^6}\frac{(1+\delta\gamma)}{\gamma^2}+ \\ &\quad \frac{6(\delta^2\gamma^2+3\delta\gamma+3)}{\gamma^4}-\frac{3(1+\delta\gamma)^4}{\gamma^8} \\ &= \frac{3\delta^2\gamma^2+27\delta\gamma+60}{\gamma^8}\end{aligned}$$

Using the theorem

$$\mu_4(X) = 3\frac{[\beta^4\delta^2\gamma^2+9\beta^4\delta\gamma+20\beta^4+12\beta^2\delta\gamma^3+20\beta^2\gamma^2+2\beta^2\delta^2\gamma^4+\delta^2\gamma^6+3\delta\gamma^5+3\gamma^4]}{\gamma^8}$$

Therefore

$$\gamma_2 = \frac{3[5\beta^4\delta\gamma+16\beta^4+6\beta^2\delta\gamma^3+16\beta^2\gamma^2+\delta\gamma^5+2\gamma^4]}{(\alpha^2(1+\delta\gamma)+\beta^2)^2} \tag{4.19}$$

Therefore in summary

Table 4.2: Properties of the NRIG: GHD($\lambda = \frac{1}{2}, \delta, \gamma$)

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	$\frac{3\beta\alpha^2(\delta\gamma+2)+2\beta^3}{(\alpha^2(1+\delta\gamma)+\beta^2)^{\frac{3}{2}}}$
4	Excess Kurtosis, γ_2	$\frac{3[5\beta^4\delta\gamma+16\beta^4+6\beta^2\delta\gamma^3+16\beta^2\gamma^2+\delta\gamma^5+2\gamma^4]}{(\alpha^2(1+\delta\gamma)+\beta^2)^2}$

4.7 Normal Variance Mean Mixture with a $GIG(-\frac{3}{2},\delta,\gamma)$ Mixing Distribution

Let the mixing distribution of z be $GIG(-\frac{3}{2},\delta,\gamma)$. The mixed model is therefore,

$$\begin{aligned} f(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi z}} e^{-\frac{1}{2}\frac{(x-(\mu+\beta z))^2}{z}} g(z)\,dz \\ &= \frac{\delta^3 \exp(\delta\gamma) e^{\beta(x-\mu)}}{2\pi(1+\delta\gamma)} \int_0^\infty z^{-3} \exp\left\{-\frac{(\beta^2+\gamma^2)}{2}\left[z+\frac{(x-\mu)^2+\delta^2}{(\beta^2+\gamma^2)}\frac{1}{z}\right]\right\} dz \end{aligned}$$

Let

$$z=\sqrt{\frac{(x-\mu)^2-\delta^2}{(\beta^2+\gamma^2)}t}$$

Therefore

$$\begin{aligned} f(x) &= \frac{\delta^3 \exp(\delta\gamma) e^{\beta(x-\mu)}}{2\pi(1+\delta\gamma)} \frac{(\beta^2+\gamma^2)}{(x-\mu)^2+\delta^2} \int_0^\infty t^{-3} \exp\left\{\frac{\sqrt{(\beta^2+\gamma^2)(x-\mu)^2+\delta^2}}{2} \times \right. \\ &\left. \left[t+\frac{1}{t}\right]\right\} dt \\ &= \frac{\delta^3 \exp(\delta\gamma) e^{\beta(x-\mu)}}{\pi(1+\delta\gamma)} \frac{(\beta^2+\gamma^2)}{(x-\mu)^2+\delta^2} K_{-2}(\sqrt{(\beta^2+\gamma^2)(x-\mu)^2+\delta^2}) \end{aligned}$$

denote $\alpha^2=\beta^2+\gamma^2$

$$f(x)=\frac{\alpha^2\delta\exp(\delta\sqrt{\alpha^2-\beta^2}-\beta\mu)\exp(\beta x)\phi(x)^{-1}K_2(\alpha\delta\phi(x)^{\frac{1}{2}})}{\pi(1+\delta\sqrt{\alpha^2-\beta^2})} \tag{4.20}$$

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

4.7.1 Properties

Mean

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

Since here $Z\ GIG(-\frac{3}{2}, \delta, \gamma)$

$$\begin{aligned} E(Z) &= \frac{\delta}{\gamma} \frac{K_{\frac{1}{2}}(\delta\gamma)}{K_{\frac{3}{2}}(\delta\gamma)} \\ &= \frac{\delta^2}{1+\delta\gamma} \end{aligned}$$

therefore

$$E(X) = \mu + \beta \frac{\delta^2}{1+\delta\gamma} \tag{4.21}$$

Variance

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

Where

$$\begin{aligned} var(Z) &= \frac{\delta^3}{\gamma(1+\delta\gamma)} - \frac{\delta^4}{(1+\delta\gamma)^2} \\ &= \frac{\delta^3}{\gamma(1+\delta\gamma)^2} \end{aligned}$$

and hence

$$\begin{aligned} var(X) &= \frac{\delta^2}{(1+\delta\gamma)} + \beta^2 \frac{\delta^3}{\gamma(1+\delta\gamma)^2} \\ &= \frac{\delta^2\gamma + \delta^3(\gamma^2+\beta^2)}{\gamma(1+\delta\gamma)^2} \\ &= \frac{\delta^2(\gamma+\alpha^2\delta)}{\gamma(1+\delta\gamma)^2} \end{aligned} \tag{4.22}$$

Skewness

Note

$$\mu_X(3) = 3\beta var(Z) + \beta^3 \mu_Z(3)$$

Where

$$\begin{aligned}\mu_Z(3) &= E(Z^3) - 3E(Z^2)E(Z) + E(Z)^3 \\ &= 3\frac{\delta^3}{\gamma^3} - \frac{3\delta^5}{\gamma(1+\delta\gamma)^2} + \frac{2\delta^6}{(1+\delta\gamma)^3} \\ &= \frac{\delta^3(1+\delta\gamma)^3 - 3\delta^5\gamma^2 - \delta^6\gamma^3}{\gamma^3(1+\delta\gamma)^3}\end{aligned}$$

hence

$$\begin{aligned}\mu_X(3) &= \frac{3\beta\delta^3}{\gamma(1+\delta\gamma)^2} + \frac{\beta^3(\delta^3(1+\delta\gamma)^3 - 3\delta^5\gamma^2 - \delta^6\gamma^3)}{\gamma^3(1+\delta\gamma)^3} \\ &= \frac{3\beta\delta^3\gamma^2 + 3\beta\delta^4\gamma^3 + \beta^3\delta^3 + 3\beta^3\delta^4\gamma}{\gamma^3(1+\delta\gamma)^3} \\ &= \frac{\beta^3\delta^3 + 3\beta\delta^3\gamma(\gamma+\alpha^2\delta)}{\gamma^3(1+\delta\gamma)^3}\end{aligned}$$

Therefore, denoting skewness by γ_1 we get

$$\begin{aligned}\gamma_1 &= \frac{\mu_3(X)}{(\mu_2(X))^{\frac{3}{2}}} \\ &= \frac{3\beta\delta^3\gamma^2 + 3\beta\delta^4\gamma^3 + \beta^3\delta^3 + 3\beta^3\delta^4\gamma}{\gamma^3(1+\delta\gamma)^3} \times \frac{\gamma^{\frac{3}{2}}(1+\delta\gamma)^3}{\delta^3(\gamma+\alpha^2\delta)^{\frac{3}{2}}} \\ &= \frac{3\beta\delta^3(\gamma^2+\alpha^2\delta\gamma) + \beta^3\delta^3}{(\gamma^2+\alpha^2\delta\gamma)^{\frac{3}{2}}} \\ &= \frac{3\beta\delta^3}{\sqrt{(\gamma^2+\alpha^2\delta\gamma)}} + \left(\frac{\beta\delta}{\sqrt{(\gamma^2+\alpha^2\delta\gamma)}}\right)^3 \qquad (4.23)\end{aligned}$$

Excess Kurtosis

Using formula (4.7),

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

where for the $GIG(-\frac{3}{2},\delta,\gamma)$ distribution

$$
\begin{aligned}
E(Z^4) &= \frac{\delta^4}{\gamma^4}\frac{K_{\frac{5}{2}}(\delta\gamma)}{K_{\frac{3}{2}}(\delta\gamma)} \\
&= \frac{\delta^4}{\gamma^4}\frac{\delta^2\gamma^2+3\delta\gamma+3}{\delta^2\gamma^2}\frac{\delta\gamma}{(1+\delta\gamma)} \\
&= \frac{\delta^5\gamma^2+3\delta^4\gamma+3\delta^3}{\gamma^5(1+\delta\gamma)}
\end{aligned}
$$

Therefore

$$
\mu_4(Z) = \frac{(\delta^5\gamma^2+3\delta^4\gamma+3\delta^3)(1+\delta\gamma)^3-4\delta^5\gamma^2(1+\delta\gamma)^3+6\delta^7\gamma^4(1+\delta\gamma)-3\delta^8\gamma^5}{\gamma^5(1+\delta\gamma)^4}
$$

note;

$$
\begin{aligned}
(\delta^5\gamma^2+3\delta^4\gamma+3\delta^3)(1+\delta\gamma)^3 &= 19\delta^5\gamma^2+15\delta^6\gamma^3+6\delta^7\gamma 4+\delta^8\gamma^5+12\delta^4\gamma+3\delta^3 \\
4\delta^5\gamma^2(1+\delta\gamma)^3 &= 4\delta^5\gamma^2+3\delta^6\gamma^3+12\delta^7\gamma^4+4\delta^8\gamma^5
\end{aligned}
$$

hence

$$
\mu_4(Z) = \frac{15\delta^5\gamma^2+3\delta^6\gamma^3+12\delta^4\gamma+3\delta^3}{\gamma^5(1+\delta\gamma)^4}
$$

Therefore

$$
\begin{aligned}
\mu_4(X) &= \frac{\beta^4 15\delta^5\gamma^2+3\delta^6\gamma^3+12\delta^4\gamma+3\delta^3}{\gamma^5(1+\delta\gamma)^4}+6\beta^2\frac{\delta^3(1+\delta\gamma)^3-3\delta^5\gamma^2-\delta^6\gamma^3}{\gamma^3(1+\delta\gamma)^3}+ \\
&\quad 6\beta^2\frac{\delta^5}{\gamma(1+\delta\gamma)^3}+3\frac{\delta^3}{\gamma(1+\delta\gamma)} \\
&= \frac{3\beta^4\delta^3(5\delta^3\gamma^2+\delta^3\gamma^3+4\delta\gamma+1)+6\beta^2\gamma^2\delta^3(1+4\delta\gamma+3\delta^2\gamma^2+\delta^2\gamma^2+\delta^3\gamma^3)}{\gamma^5(1+\delta\gamma)^4}+ \\
&\quad \frac{3\delta^3\gamma^4(1+3\delta\gamma+3\delta^2\gamma^2+\delta^3\gamma^3)}{\gamma^5(1+\delta\gamma)^4}
\end{aligned}
$$

then working out kurtosis gives

$$
\begin{aligned}
\frac{\mu_4(X)}{(\mu_2(X))^2} &= \Bigg(\frac{3\beta^4\delta^3(5\delta^3\gamma^2+\delta^3\gamma^3+4\delta\gamma+1)+6\beta^2\gamma^2\delta^3(1+4\delta\gamma+4\delta^2\gamma^2+\delta^3\gamma^3)}{\gamma^5(1+\delta\gamma)^4}+ \\
&\quad \frac{3\delta^3\gamma^4(1+3\delta\gamma+3\delta^2\gamma^2+\delta^3\gamma^3)}{\gamma^5(1+\delta\gamma)^4}\Bigg)\times\frac{\gamma^2(1+\delta\gamma)^4}{\delta^4(\gamma+\alpha^2\delta)^2}
\end{aligned}
$$

noting that

$$\begin{aligned}\alpha^2 &= \delta^2+\gamma^2\\ \alpha^4 &= \beta^4+2\beta^2\gamma^2+\gamma^4\end{aligned}$$

then,

$$\delta\gamma^3(\gamma+\alpha^2\delta)^2=\delta\gamma^5+2\beta^2\delta^2\gamma^4+2\beta^2\gamma^6+\beta^4\delta^3\gamma^3+2\beta^2\delta^3\gamma^5+\delta^3\gamma^7$$

Denoting excess kurtosis by γ_2 we obtain

$$\begin{aligned}\gamma_2 &= \frac{3[5\beta^4\delta^2\gamma^2+4\beta^4\delta\gamma+\beta^4+2\beta^2\gamma^2+8\beta^2\delta\gamma^3+6\beta^2\delta^2\gamma^4+\gamma^4+2\delta\gamma^5+\delta^2\gamma^6]}{\delta\gamma^3(\gamma+\alpha^2\delta)^2}\\ &= \frac{3[\alpha^4(1+\delta\gamma)^2+4\beta^4\delta^2\gamma^2+2\beta^4\delta\gamma+4\beta^2\delta\gamma^3+4\beta^2\delta^2\gamma^4]}{\delta\gamma^3(\gamma+\alpha^2\delta)^2}\\ &= \frac{3[\alpha^4(1+\delta\gamma)^2+2\delta\gamma(\alpha^4(1+2\delta\gamma)-\gamma^3(\gamma+2\alpha^2\delta))]}{\delta\gamma^3(\gamma+\alpha^2\delta)^2} \qquad (4.24)\end{aligned}$$

Therefore in summary

Table 4.3: Properties of the GHD($\lambda=-\frac{3}{2}$)

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	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	γ_2	$\frac{3[\alpha^4(1+\delta\gamma)^2+2\delta\gamma(\alpha^4(1+2\delta\gamma)-\gamma^3(\gamma+2\alpha^2\delta))]}{\delta\gamma^3(\gamma+\alpha^2\delta)^2}$

4.8 Normal Variance Mean Mixture with a $GIG(\frac{3}{2},\delta,\gamma)$ Mixing Distribution

Now suppose that the mixing distribution is $GIG(\frac{3}{2},\delta,\gamma)$, the special case of the GIG becomes

$$g(z)=\frac{\gamma^3}{\sqrt{2\pi}(1+\delta\gamma)}z^{\frac{1}{2}}\exp(\delta\gamma)\exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2z\right)\right\} \qquad (4.25)$$

4.8.1 Mixed model

The mixed model becomes

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

let

$$z=\sqrt{\frac{(x-\mu)^2+\delta^2}{(\beta^2+\gamma^2)}}t$$

therefore

$$\begin{aligned} f(x) &= \frac{\gamma^3\exp(\delta\gamma-\beta\mu)e^{\beta x}}{\pi(1+\delta\gamma)}\sqrt{\frac{(x-\mu)^2+\delta^2}{(\beta^2+\gamma^2)}}K_1(\sqrt{(\beta^2+\gamma^2)(x-\mu)^2+\delta^2}) \\ &= \frac{\delta\gamma^3\exp(\delta\gamma-\beta\mu)e^{\beta x}\phi(x)^{\frac{1}{2}}K_1(\alpha\delta\phi(x)^{\frac{1}{2}})}{\alpha\pi(1+\delta\gamma)} \end{aligned} \tag{4.26}$$

Which is the GH distribution for $\lambda=\frac{3}{2}$

4.8.2 Properties

Mean

Since

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

$$E(X)=\mu+\frac{\beta(\delta^2\gamma^2+3\delta\gamma+3)}{\gamma^2(1+\delta\gamma)} \tag{4.27}$$

Variance

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

where

$$\begin{aligned} var(Z) &= E(Z^2)-(E(Z))^2 \\ &= \frac{\delta^3\gamma^3+6\delta^2\gamma^2+12\delta\gamma+6}{\gamma^4(1+\delta\gamma)^2} \end{aligned}$$

hence

$$\begin{aligned} var(X) &= \frac{(\delta^2\gamma^2+3\delta\gamma+3)(\gamma^2+\delta\gamma^3)+\beta^2(\delta^3\gamma^3+6\delta^2\gamma^2+12\delta\gamma+6)}{\gamma^4(1+\delta\gamma)^2} \\ &= \frac{4\alpha^2\delta^2\gamma^2+2\beta^2\delta^2\gamma^2+6\alpha^2\delta\gamma+6\beta^2\delta\gamma+3\alpha^2+3\beta^2+\alpha^2\delta^3\gamma^3}{\gamma^4(1+\delta\gamma)^2} \\ &= \frac{3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2}{\gamma^4(1+\delta\gamma)^2} \end{aligned} \tag{4.28}$$

Skewness

Note

$$\begin{aligned} E(Z^3) &= \left(\frac{\delta}{\gamma}\right)^3 \frac{K_{\frac{9}{2}}(\delta\gamma)}{K_{\frac{3}{2}}(\delta\gamma)} \\ &= \frac{\delta^4\gamma^4+10\delta^3\gamma^3+45\delta^2\gamma^2+105\delta\gamma+105}{\gamma^6(1+\delta\gamma)} \end{aligned}$$

It can easily be shown that

$$\mu_3(Z) = \frac{3\delta^4\gamma^4+25\delta^3\gamma^3+72\delta^2\gamma^2+72\delta\gamma+24}{\gamma^6(1+\delta\gamma)^3}$$

therefore

$$\begin{aligned} \mu_3(X) &= \frac{3\beta\alpha^2(1+\delta\gamma)^4+3\beta(1+\delta\gamma)^3(3\alpha^2+2\beta^2)+6\alpha^2\beta(1+\delta\gamma)^2}{\gamma^6(1+\delta\gamma)^3} \\ &\quad + \frac{3\alpha^2\beta\delta\gamma(1+\delta\gamma)-2\beta\delta^3\gamma^3}{\gamma^6(1+\delta\gamma)^3} \end{aligned}$$

we obtain skewness to be

$$\begin{aligned} \gamma_1 &= \frac{3\beta\alpha^2(1+\delta\gamma)^4+3\beta(1+\delta\gamma)^3(3\alpha^2+2\beta^2)+6\alpha^2\beta(1+\delta\gamma)^2}{3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2)^{\frac{3}{2}}} \\ &\quad + \frac{3\alpha^2\beta\delta\gamma(1+\delta\gamma)-2\beta\delta^3\gamma^3}{(3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2)^{\frac{3}{2}}} \end{aligned} \tag{4.29}$$

Excess Kurtosis

Note

$$E(Z^4) = \frac{\delta^5\gamma^5+15\delta^4\gamma^4+105\delta^3\gamma^3+420\delta^2\gamma^2+945\delta\gamma+945}{\gamma^8(1+\delta\gamma)}$$

and after some simplification it can be shown that

$$\mu_4(Z) = \frac{3\delta^6\gamma^6+51\delta^5\gamma^5+336\delta^4\gamma^4+1047\delta^3\gamma^3 1512\delta^2\gamma^2+1008\delta\gamma+252}{\gamma^8(1+\delta\gamma)^4}$$

Using the theorem, excess kurtosis for the model can be shown to be

$$\begin{aligned}\gamma_2 = {} & \frac{3[5\beta^4\delta^5\gamma^5+52\beta^4\delta^4\gamma^4+193\beta^4\delta^3\gamma^3+288\beta^4\delta^2\gamma^2+192\beta^4\delta\gamma+48\beta^4+56\beta^2\delta^4\gamma^6]}{(3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2)^2} \\ & +\frac{3[192\beta^2\delta^3\gamma^5+288\beta^2\delta^2\gamma^4+228\beta^2\delta\gamma^3+6\beta^2\delta^5\gamma^7+48\beta^2\gamma^2+\delta^5\gamma^9+8\delta^4\gamma^8]}{(3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2)^2} \\ & +\frac{3[25\delta^3\gamma^7+36\delta^2\gamma^6+24\delta\gamma^5+6\gamma^4]}{(3\alpha^2(1+\delta\gamma)^2+\alpha^2\delta^2\gamma^2(1+\delta\gamma)+2\beta^2\delta\gamma(3+\delta\gamma)+3\beta^2)^2} \qquad (4.30)\end{aligned}$$

In this capture we have worked on Normal Variance Mean Mixtures when four special cases of GIG are used as mixing distribution. The properties of these models have also been derived. In the next chapter, we utilize the finite cases of the four special cases as mixing distributions to extend the work on normal mixtures.

5 Normal Weighted Inverse Gaussian Distribution Part II

5.1 Introduction

In this chapter we extend the work on Normal Variance Mean Mixture by using finite mixtures of the special cases of Generalised Inverse Gaussian as mixing distribution. There are few works on finite mixtures as mixing distribution.

Lindley distribution introduced by Lindley (1958) is a finite mixture of exponential and gamma distribution. A detailed study of this one-parameter two component (finite) mixture was given by Gitany et al. (2008). Sankara (1970) had used it as a mixing distribution to Poisson distribution. Shanker, R. and Hogos, F. (2015) applied the Poisson-Lindley distribution to Biological Science data.

Generalizations of the one-parameter Lindley distribution has been studied by researchers. For example Zakerzadeh and Dolati (2010) generalized it to a three parameter Lindley distribution. Mahmoudi and Zakerzadeh (2010) used this three parameter Lindley as a mixing distribution to a Poisson distribution to come up with Poisson-Generalised Lindley distribution.

Bakouch et al (2012) have extended the Lindley distribution by exponentiated approach. However, it has not been used as a mixing distribution. A transmuted exponential distribution is a finite mixture of exponential distribution with parameter θ and another exponential with parameter 2θ.

Bhati et al (2015) used it as a mixing distribution to a Poisson distribution. The mixed Poisson Transmuted Exponential distribution was applied to Health care data.

The skew normal distribution of Azzalini (1985) is one of the widely used probability distribution for modeling skewed data. Independently Azzalini (1986) and Henze (1986) showed that the skew normal distribution is a normal mixture with varying mean taking the half-normal distribution as mixing distribution.

Negarestani et al (2018) generalised the mixing to be any positive random variable. They specifically used the standard exponential distribution distribution and a finite mixture of an exponential and half-normal distribution as mixing distribution.

The Skew normal (SN) distribution possess limited ranges of skewness and excess Kurtosis coefficients given by (-0.995, 0.995) and (0, 0.869) respectively (Azzalini, 1985).

When the exponential distribution is the mixing distribution, the mean mixture of the normal exponential distribution has intervals for the skewness and kurtosis coefficents as (-2,2) and (0,6) respectively. These ranges are considerable wider than those for SN. The ranges of the skewness and (excess) kurtosis coefficients for the finite mixture as a mixing distribution are (0.995, 3.916) and (0.869,31.980) respectively.

The point of message we are conveying is that finite mixtures are more flexible (robust) than single mixing distributions; a result that has been stated a lot in statistical literature.
In chapter 4 We have further shown that the pairwise finite mixtures of $GIG(-\frac{1}{2},\delta,\gamma)$, $GIG(\frac{1}{2},\delta,\gamma)$, $GIG(-\frac{3}{2},\delta,\gamma)$ and $GIG(\frac{3}{2},\delta,\gamma)$ are also weighted Inverse Gaussian distributions. We have given **SIX** finite mixtures including one case by Gupta and his colleagues. More cases can be constructed to form a class called Weighted Inverse Gaussian distributions.
In fact Lindley distribution and its generalizations form another class of weighted distributions, namely, a class of weighted Exponential distributions.
The objective of this chapter is to model flexible distributions that can handle skewed and heavy-tailed data by using a weighted inverse Gaussian finite mixture as a mixing distribution to the normal variance mean mixture.
This idea is motivated by the fact that finite mixtures are more flexible than single distributions. Nadarajah, Zhang and Chan (2014) 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 distribution and its generalization in Poisson mixtures (e.g Sankaran, 1970; Mohamud and Zakerzadeh, 2010). Lindley distribution and its generalizations are basically finite gamma mixtures.

5.2 The Mixing Mechanism

From chapter 4,the normal variance mean mixture is given by formula (4.1) expressed as

$$\begin{aligned} f(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi z}} e^{-\frac{(x-\mu-\beta z)^2}{2z}} g(z)\,dz \\ &= \frac{e^{\beta(x-\mu)}}{\sqrt{2\pi}} \int_0^\infty z^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\beta^2 z+\frac{(x-\mu)^2}{z}\right)} g(z)\,dz \end{aligned}$$

For the weighted inverse Gaussian distribution

$$g(z) = \frac{w(z)}{E[w(Z)]} g_1(z)$$

where

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

Hence

$$f(x) = \frac{\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{\pi} \frac{1}{2} \int_0^\infty \frac{w(z)}{E[w(Z)]} z^{-2} \exp\Big\{-\frac{1}{2}\Big((\gamma^2+\beta^2)z + \frac{\delta^2+(x-\mu)^2}{z}\Big)\Big\}\,dz \quad \textbf{(5.1)}$$

Therefore a weighted Inverse Gaussian Mixture of a Normal Variance-Mean distribution is

$$\therefore f(x) = \frac{\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{2\pi} \int_0^\infty \frac{w(z)}{E[w(Z)]} z^{-2} e^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}\, dz \tag{5.2}$$

We now construct and obtain properties for the six cases of finite Inverse Gaussian distribution derived in chapter 3. The posterior expectations which will be used in parameter estimation in the next chapter are also given.

Note

The pdf of a posterior distribution is given by

$$\begin{aligned} f(z/x) &= \frac{f(x,z)}{f(x)} \\ &= \frac{f(x/z)g(z)}{f(x)} \\ &= \frac{f(x/z)g(z)}{\int_0^\infty f(x/z)g(z)\, dz} \end{aligned} \tag{5.3}$$

Let define $f_{ij}(z)$ for $i,j = 1,2,3,4$ but $i \neq j$ to be a mixed model for a mixing distribution which is a finite mixture between $g_i(z)$ and $g_j(z)$.

5.3 Model 1

5.3.1 construction

Suppose the mixing distribution follows formula (3.15). That is, the finite mixture of IG and RIG distribution as a mixing distribution we have the NVM mixture given by

$$\begin{aligned} f_{12}(x) &= \frac{\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{2\pi} \int_0^\infty \frac{\gamma}{\gamma+\delta}(1+z) z^{-2} e^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}\, dz \\ &= \frac{\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{2\pi} \frac{\gamma}{\gamma+\delta} \int_0^\infty (z^{-2}+z^{-1}) z^{-2} e^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}\, dz \\ &= \frac{\delta\gamma e^{\delta\gamma} e^{\beta(x-\mu)}}{2\pi(\gamma+\delta)} \int_0^\infty \left[z^{-1-1}+z^{0-1}\right] e^{-\frac{1}{2}\left\{\frac{\delta^2\phi_x}{z}+\alpha^2 z\right\}}\, dz \\ &= \frac{\delta\gamma e^{\delta\gamma} e^{\beta(x-\mu)}}{\pi(\gamma+\delta)} \left\{ \left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1} K_1(\alpha\delta\sqrt{\phi(x)}) + K_0(\alpha\delta\sqrt{\phi(x)}) \right\} \\ &= \frac{\delta\gamma e^{\delta\gamma} e^{\beta(x-\mu)}}{\pi(\gamma+\delta)} \left\{ \frac{\alpha}{\delta\sqrt{\phi(x)}} K_1(\alpha\delta\sqrt{\phi(x)}) + K_0(\alpha\delta\sqrt{\phi(x)}) \right\} \end{aligned} \tag{5.4}$$

5.3.2 The log-likelihood function of the proposed mixed model

Let

$$
\begin{aligned}
l &= log\prod_{i=1}^{n} f(x_i) \\
&= \sum_{i=1}^{n} \log f(x_i) \\
&= \sum_{i=1}^{n} \log\left\{\frac{\delta\gamma e^{\delta\gamma}e^{\beta(x_i-\mu)}}{\pi(\gamma+\delta)}\left[\frac{\alpha}{\delta\sqrt{\phi(x_i)}}K_1(\alpha\delta\sqrt{\phi(x_i)})+K_0(\alpha\delta\sqrt{\phi(x_i)})\right]\right\} \\
&= \sum_{i=1}^{n}\left\{\log(\delta\gamma)+\delta\gamma+\beta(x_i-\mu)-\log(\pi(\gamma+\delta))+\log\left[\frac{\alpha}{\delta\sqrt{\phi(x_i)}}K_1(\alpha\delta\sqrt{\phi(x_i)})+\right.\right. \\
&\qquad \left.\left. K_0(\alpha\delta\sqrt{\phi(x_i)})\right]\right\} \\
&= n\log(\delta\gamma)+n\delta\gamma+\sum_{i=1}^{n}\beta(x_i-\mu)-n\log(\pi(\gamma+\delta))+\sum_{i=1}^{n}\log\left[\frac{\alpha}{\delta\sqrt{\phi(x_i)}}K_1(\alpha\delta\sqrt{\phi(x_i)})+\right. \\
&\qquad \left. K_0(\alpha\delta\sqrt{\phi(x_i)})\right]
\end{aligned}
$$

Therefore

$$
\begin{aligned}
l &= n\log(\delta\gamma)+n\delta\gamma+n\beta\bar{x}-n\beta\mu-n\log(\pi(\gamma+\delta))+\sum_{i=1}^{n}\log\left[\frac{\alpha}{\delta\sqrt{\phi(x_i)}}K_1(\alpha\delta\sqrt{\phi(x_i)})+\right. \\
&\qquad \left. K_0(\alpha\delta\sqrt{\phi(x_i)})\right] \qquad (5.5)
\end{aligned}
$$

5.3.3 Properties of Model 1

Since

$$g(z) = \frac{\gamma}{\gamma+\delta}(1+z)GIG(-\frac{1}{2},\delta,\gamma)$$

then

$$E[Z] = \frac{\delta+\delta\gamma(\gamma+\delta)}{\gamma^2(\gamma+\delta)} \qquad (5.6)$$

and

$$var(Z) = \frac{3\delta\gamma+2\delta^2\gamma^2+\delta\gamma^3+2\delta^2+\delta^3\gamma}{\gamma^4(\delta+\gamma)^2} \qquad (5.7)$$

Therefore the mean and variance of the proposed model are

$$E(X) = \mu+\frac{\beta(\delta+\delta\gamma(\gamma+\delta))}{\gamma^2(\gamma+\delta)} \qquad (5.8)$$

and

$$var(X) = \frac{\delta^2\gamma^2(1+2\alpha^2)+\delta\gamma^3(1+\alpha^2)+\alpha^2\delta^3\gamma+\beta^2\delta(3\gamma+2\delta)}{\gamma^4(\delta+\gamma)^2} \qquad (5.9)$$

5.3.4 Posterior Expectations

$$
\begin{aligned}
E[Z/X=x] &= \frac{\int_0^\infty z f(x/z) g(z)\, dz}{\int_0^\infty f(x/z) g(z)\, dz} \\
&= \frac{\frac{1}{2}\int_0^\infty z(z^{-2}+z^{-1}) e^{-\frac{1}{2}(\alpha^2 z+\frac{\delta^2}{z}\phi(x))}\, dz}{\frac{1}{2}\int_0^\infty (z^{-2}+z^{-1}) e^{-\frac{1}{2}(\alpha^2 z+\frac{\delta^2}{z}\phi(x))}\, dz} \\
&= \frac{\frac{1}{2}\int_0^\infty (z^{0-1}+z^{1-1}) e^{-\frac{1}{2}(\alpha^2 z+\frac{\delta^2}{z}\phi(x))}\, dz}{\frac{1}{2}\int_0^\infty (z^{-1-1}+z^{0-1}) e^{-\frac{1}{2}(\alpha^2 z+\frac{\delta^2}{z}\phi(x))}\, dz}
\end{aligned}
$$

and in terms of modiffied Bessel function of the third kind we can express the expectations as:

$$
\begin{aligned}
E[Z/X=x] &= \frac{K_0(\alpha\delta\sqrt{\phi(x)}) + \frac{\delta\sqrt{\phi(x)}}{\alpha} K_1(\alpha\delta\sqrt{\phi(x)})}{\frac{\alpha}{\delta\sqrt{\phi(x)}} K_1(\alpha\delta\sqrt{\phi(x)}) + K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\delta\sqrt{\phi(x))} K_0(\alpha\delta\sqrt{\phi(x)}) + \frac{\delta^2\phi(x)}{\alpha} K_1(\alpha\delta\sqrt{\phi(x)})}{\alpha K_1(\alpha\delta\sqrt{\phi(x)}) + \delta\sqrt{\phi(x)} K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta\sqrt{\phi(x))} K_0(\alpha\delta\sqrt{\phi(x)}) + \delta^2\phi(x) K_1(\alpha\delta\sqrt{\phi(x)})}{\alpha^2 K_1(\alpha\delta\sqrt{\phi(x)}) + \alpha\delta\sqrt{\phi(x)} K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta\sqrt{\phi(x))} K_0(\alpha\delta\sqrt{\phi(x)}) + \delta^2\phi(x) K_1(\alpha\delta\sqrt{\phi(x)})}{\alpha\delta\sqrt{\phi(x)} K_0(\alpha\delta\sqrt{\phi(x)}) + \alpha^2 K_1(\alpha\delta\sqrt{\phi(x)})} \qquad (5.10)
\end{aligned}
$$

Next

$$
\begin{aligned}
E\left(\frac{1}{Z}/X=x\right) &= \frac{\frac{1}{2}\int_0^\infty \frac{1}{z}(z^{-2}+z^{-1})e^{-\frac{1}{2}(\alpha^2+\frac{\delta^2}{z}\phi(x))}\,dz}{\frac{1}{2}\int_0^\infty \frac{1}{z}(z^{-2}+z^{-1})e^{-\frac{1}{2}(\alpha^2+\frac{\delta^2}{z}\phi(x))}} \\
&= \frac{\frac{1}{2}\int_0^\infty (z^{-2-1}+z^{-1-1})e^{-\frac{1}{2}(\alpha^2+\frac{\delta^2}{z}\phi(x))}\,dz}{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-2}K_2(\alpha\delta\sqrt{\phi(x)})+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1}K_1(\alpha\delta\sqrt{\phi(x)})}{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\left(\frac{\alpha}{\delta\sqrt{\phi(x)}}\right)^{2}K_2(\alpha\delta\sqrt{\phi(x)})+\left(\frac{\alpha}{\delta\sqrt{\phi(x)}}\right)^{1}K_1(\alpha\delta\sqrt{\phi(x)})}{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^2K_2(\alpha\delta\sqrt{\phi(x)})+\alpha\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})}{\alpha\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})+\delta^2\phi(x)K_0(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})+\alpha^2K_2(\alpha\delta\sqrt{\phi(x)})}{\alpha\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})+\delta^2\phi(x)K_0(\alpha\delta\sqrt{\phi(x)})} \qquad \textbf{(5.11)}
\end{aligned}
$$

5.4 Model 2

5.4.1 Construction

Suppose the mixing distribution follows formula (3.21), the mixed model becomes

$$
\begin{aligned}
f_{13}(x) &= \frac{\delta^3e^{\delta\gamma}}{2\pi(1+\delta^2)}e^{\beta(x-\mu)}\int_0^\infty (1+\frac{z^{-1}}{1+\delta\gamma})z^{-2}e^{-\frac{1}{2}\left(\alpha^2z+\frac{\delta^2\phi(x)}{z}\right)}\,dz \\
&= \frac{\delta^3e^{\delta\gamma}}{2\pi(1+\delta^2)}e^{\beta(x-\mu)}\int_0^\infty (z^{-1-1}+\frac{z^{-2-1}}{1+\delta\gamma})e^{-\frac{1}{2}\left(\alpha^2z+\frac{\delta^2\phi(x)}{z}\right)}\,dz
\end{aligned}
$$

therefore

$$
\begin{aligned}
f_{13}(x) &= \frac{\delta^3e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(1+\delta^2)}\left\{\left[\frac{\delta\sqrt{\phi(x)}}{\alpha}\right]^{-1}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\left[\frac{\delta\sqrt{\phi(x)}}{\alpha}\right]^{-2}\frac{K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{1+\delta\gamma}\right\} \\
&= \frac{\delta^3e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(1+\delta^2)}\left\{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^2}{\delta^2\phi(x)}\frac{K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{1+\delta\gamma}\right\} \\
&= \frac{\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(1+\delta^2)\phi(x)}\left\{\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^2}{1+\delta\gamma}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)\right\} \qquad \textbf{(5.12)}
\end{aligned}
$$

where

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

5.4.2 The log-likelihood function

$$
\begin{aligned}
l &= \sum_{i=1}^{n} \log f(x_i) \\
&= \sum_{i=1}^{n} \Bigg\{ \log\delta + \delta\gamma + \beta(x_i-\mu) - \log(\pi(1+\delta^2)) - \log(\phi(x_i)) + \log\Bigg\{ \alpha\delta\sqrt{\phi(x)} \times \\
&\quad K_1\big(\alpha\delta\sqrt{\phi(x)}\big) + \frac{\alpha^2}{1+\delta\gamma} K_2\big(\alpha\delta\sqrt{\phi(x)}\big) \Bigg\}\Bigg\} \\
&= n\log\delta + n\delta\gamma + \beta\sum_{i=1}^{n} x_i - n\beta\mu - n\log(\pi(1+\delta^2)) - \sum_{i=1}^{n}\log(\phi(x_i)) + \sum_{i=1}^{n}\log\Bigg\{ \alpha\delta\sqrt{\phi(x)} \times \\
&\quad K_1\big(\alpha\delta\sqrt{\phi(x)}\big) + \frac{\alpha^2}{1+\delta\gamma} K_2\big(\alpha\delta\sqrt{\phi(x)}\big) \Bigg\}
\end{aligned} \tag{5.13}
$$

5.4.3 Properties of Model 2

$$
E(X) = \mu + \beta\frac{\delta^2}{1+\delta^2}\left[\frac{\delta(1+\delta\gamma)+\gamma}{\gamma(1+\delta\gamma)}\right] \tag{5.14}
$$

$$
\begin{aligned}
var(X) &= \frac{\delta^2}{1+\delta^2}\left[\frac{\delta(1+\delta\gamma)+\gamma}{\gamma(1+\delta\gamma)}\right] + \\
&\quad \frac{\beta^2[\delta^3(\gamma^2+(1+\delta\gamma)^2)(1+\delta^2)(1+\delta\gamma) - \delta^4\gamma(\delta(1+\delta\gamma)+\gamma)^2]}{\gamma^3(1+\delta\gamma)^2(1+\delta^2)^2}
\end{aligned} \tag{5.15}
$$

5.4.4 Posterior Expectation

$$
\begin{aligned}
E(Z/X=x) &= \frac{\int_0^\infty z f(x/z) g(z)\, dz}{\int_0^\infty f(x/z) g(z)\, dz} \\
&= \frac{\int_0^\infty z(1+\frac{z^{-1}}{1+\delta\gamma}) z^{-2} e^{-\frac{1}{2}\left(\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right)}\, dz}{\int_0^\infty (1+\frac{z^{-1}}{1+\delta\gamma}) z^{-2} e^{-\frac{1}{2}\left(\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right)}\, dz}
\end{aligned}
$$

$$
\begin{aligned}
E(Z/X=x) &= \frac{\int_0^\infty (z^{0-1}+\frac{z^{-1-1}}{1+\delta\gamma})e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty (z^{-1-1}+\frac{z^{-2-1}}{1+\delta\gamma})e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1}\frac{K_{-1}\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1}K_{-1}\big(\alpha\delta\sqrt{\phi(x)}\big)+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-2}\frac{K_{-2}\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}} \\
&= \frac{K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\big(\frac{\alpha}{\delta\sqrt{\phi(x)}}\big)\frac{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{\big(\frac{\alpha}{\delta\sqrt{\phi(x)}}\big)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\big(\frac{\alpha^2}{\delta^2\phi(x)}\big)\frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}} \\
&= \frac{\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\big(\frac{\alpha\delta\sqrt{\phi(x)}}{1+\delta\gamma}\big)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)}{\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\big(\frac{\alpha^2}{1+\delta\gamma}\big)K_2\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{(1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)}{(1+\delta\gamma)\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha^2K_2\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+(1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)}{(1+\delta\gamma)\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha^2K_2\big(\alpha\delta\sqrt{\phi(x)}\big)} \qquad \textbf{(5.16)}
\end{aligned}
$$

Similarly

$$
\begin{aligned}
E\left(\frac{1}{Z}/X=x\right) &= \frac{\int_0^\infty z^{-1}f(x/z)g(z)\,dz}{\int_0^\infty f(x/z)g(z)\,dz} \\
&= \frac{\int_0^\infty z^{-1}(1+\frac{z^{-1}}{1+\delta\gamma})z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty (1+\frac{z^{-1}}{1+\delta\gamma})z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\int_0^\infty (z^{-2-1}+\frac{z^{-3-1}}{1+\delta\gamma})e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty (z^{-1-1}+\frac{z^{-2-1}}{1+\delta\gamma})e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-2}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-3}\frac{K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-2}\frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}} \\
&= \frac{\frac{\alpha^2}{\delta^2\phi(x)}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^3}{\delta^3(\sqrt{\phi(x)})^3}\frac{K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{\frac{\alpha}{\delta\sqrt{\phi(x)}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}+\frac{\alpha^2}{\delta^2\phi(x)}\frac{K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}} \\
&= \frac{\alpha^2\delta\sqrt{\phi(x)}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^3K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{\alpha\delta^2\phi(x)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha^2\delta\sqrt{\phi(x)}\frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
E\left(\frac{1}{Z}/X=x\right) &= \frac{(1+\delta\gamma)\alpha^2\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+\alpha^3K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{(1+\delta\gamma)\alpha\delta^2\phi(x)K_1\left(\alpha\delta\sqrt{\phi(x)}\right)+\alpha^2\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)} \\
&= \frac{(1+\delta\gamma)\alpha\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+\alpha^2K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{(1+\delta\gamma)\delta^2\phi(x)K_1\left(\alpha\delta\sqrt{\phi(x)}\right)+\alpha\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)} \\
&= \frac{(1+\delta\gamma)\alpha\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+\alpha^2K_3\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha\delta\sqrt{\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+(1+\delta\gamma)\delta^2\phi(x)K_1\left(\alpha\delta\sqrt{\phi(x)}\right)} \quad \textbf{(5.17)}
\end{aligned}
$$

5.5 Model 3

5.5.1 Construction

Suppose the mixing distribution follows formula (3.27), the mixed model becomes

$$
\begin{aligned}
f_{14}(x) &= \frac{\delta\gamma^3e^{\delta\gamma}}{2\pi(\gamma^3+\delta)}e^{\beta(x-\mu)}\int_0^\infty\left(1+\frac{z^2}{1+\delta\gamma}\right)z^{-2}e^{-\frac{1}{2}\left(\alpha^2z+\frac{\delta^2\phi(x)}{z}\right)}dz \\
&= \frac{\delta\gamma^3e^{\delta\gamma}}{2\pi(\gamma^3+\delta)}e^{\beta(x-\mu)}\int_0^\infty\left(z^{-1-1}+\frac{z^{-1-1}}{1+\delta\gamma}\right)e^{-\frac{1}{2}\left(\alpha^2z+\frac{\delta^2\phi(x)}{z}\right)}dz
\end{aligned}
$$

$$
\begin{aligned}
f_{14}(x) &= \frac{\delta\gamma^3e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\gamma^3+\delta)}\left\{\left[\frac{\delta\sqrt{\phi(x)}}{\alpha}\right]^{-1}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)+\left[\frac{\delta\sqrt{\phi(x)}}{\alpha}\right]\frac{K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}\right\} \\
&= \frac{\delta\gamma^3e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\gamma^3+\delta)}\left\{\frac{\alpha}{\delta\sqrt{\phi(x)}}+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\right\}K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \\
&= \frac{\delta\gamma^3e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\gamma^3+\delta)\alpha\delta(1+\delta\gamma)\sqrt{\phi(x)}}\left\{\alpha^2(1+\delta\gamma)+\delta^2\phi(x)\right\}K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \\
&= \frac{\gamma^3e^{\delta\gamma}e^{\beta(x-\mu)}\left(\sqrt{\phi(x)}\right)^{-1}}{\alpha\pi(\gamma^3+\delta)(1+\delta\gamma)}\left\{\alpha^2(1+\delta\gamma)+\delta^2\phi(x)\right\}K_1\left(\alpha\delta\sqrt{\phi(x)}\right) \quad \textbf{(5.18)}
\end{aligned}
$$

5.5.2 The log-likelihood function

$$
\begin{aligned}
l = \log L &= \sum_{i=1}^{n} \log f(x_i) \\
&= \sum_{i=1}^{n} \left\{ 3\log\gamma + \delta\gamma + \beta(x_i - \mu) - \log(\alpha\pi(\gamma^3 + \delta))(1+\delta\gamma) - \frac{1}{2}\log\phi(x_i) + \right. \\
&\quad \left. \log\left[\alpha^2(1+\delta\gamma) + \delta^2\phi(x_i)\right] + \log K_1(\alpha\delta\sqrt{\phi(x_i)}) \right\} \\
&= \left\{ 3n\log\gamma + n\delta\gamma + \beta\sum_{i=1}^{n} x_i - n\beta\mu - n\log(\alpha\pi(\gamma^3+\delta))(1+\delta\gamma) - \frac{1}{2}\sum_{i=1}^{n}\log\phi(x_i) + \right. \\
&\quad \left. \sum_{i=1}^{n}\log\left[\alpha^2(1+\delta\gamma) + \delta^2\phi(x_i)\right] + \sum_{i=1}^{n}\log K_1(\alpha\delta\sqrt{\phi(x_i)}) \right\} \qquad (5.19)
\end{aligned}
$$

5.5.3 Properties of Model 3

$$
E(X) = \mu + \beta\frac{\delta\gamma^4(1+\delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta}{\gamma^2(1+\delta\gamma)(\gamma^3+\delta)} \qquad (5.20)
$$

$$
\begin{aligned}
var(X) &= \frac{\delta\gamma^4(1+\delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta}{\gamma^2(1+\delta\gamma)(\gamma^3+\delta)} \\
&\quad -\beta^2\left[\frac{(\delta\gamma^6(1+\delta\gamma)^2 + \delta^5\gamma^4 + 10\delta^4\gamma^3 + 45\delta^3\gamma^2 + 105\delta^2\gamma + 105\delta)(1+\delta\gamma)(\gamma^3+\delta)}{\gamma^6(1+\delta\gamma)^2(\gamma^3+\delta)}\right. \\
&\quad \left. -\frac{(\delta\gamma^4(1+\delta\gamma) + \delta^3\gamma^2 + 3\delta^2\gamma + 3\delta)^2}{\gamma^6(1+\delta\gamma)^2(\gamma^3+\delta)}\right] \qquad (5.21)
\end{aligned}
$$

5.5.4 Posterior Expectation

$$
\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z\big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty \big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{2-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{K_0(\alpha\delta\sqrt{\phi(x)})+\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^2\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)}}{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-1}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^2\phi(x)}{\alpha^2}\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)}}{\big[\frac{\alpha}{\delta\sqrt{\phi(x)}}+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta(1+\delta\gamma)\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^3(\phi(x))^{\frac{3}{2}}}{\alpha}K_2(\alpha\delta\sqrt{\phi(x)})}{[\alpha^2(1+\delta\gamma)+\delta^2\phi(x)]K_1(\alpha\delta\sqrt{\phi(x)})} \qquad \textbf{(5.22)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}\big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty \big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{-2-1}+\frac{z^{0-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-2}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{\big[\big(\frac{\delta\sqrt{\phi(x)}}{\alpha}\big)^{-1}+\big(\frac{\delta\sqrt{\phi(x)}}{\alpha}\big)\frac{1}{1+\delta\gamma}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\frac{\alpha^2}{\delta^2\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{\big[\frac{\alpha}{\delta\sqrt{\phi(x)}}+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^2K_2(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^2\phi(x)}{1+\delta\gamma}K_0(\alpha\delta\sqrt{\phi(x)})}{\big[\alpha\delta\sqrt{\phi(x)}+\frac{\delta^3\left(\phi(x)\right)^{\frac{3}{2}}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^3(1+\delta\gamma)K_2(\alpha\delta\sqrt{\phi(x)})+\alpha\delta^2\phi(x)K_0(\alpha\delta\sqrt{\phi(x)})}{\big[\alpha^2\delta(1+\delta\gamma)\sqrt{\phi(x)}+\delta^3\left(\phi(x)\right)^{\frac{3}{2}}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \qquad \textbf{(5.23)}
\end{aligned}
$$

Similarly

$$E(z^2/X=x)=\frac{(1+\delta\gamma)\alpha^2\delta^2\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})+\delta^4(\phi(x))^2K_3(\alpha\delta\sqrt{\phi(x)})}{[(1+\delta\gamma)\alpha^4+\alpha^2\delta^2\phi(x)]K_1(\alpha\delta\sqrt{\phi(x)})} \tag{5.24}$$

5.6 Model 4

5.6.1 Construction

Suppose the mixing distribution follows formula (3.33), the mixed model becomes

$$\begin{aligned}
f_{23}(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi z}}e^{-\frac{[(x-\mu)-\beta z]^2}{2z}}g_{23}(z)\,dz \\
&= \frac{\gamma\delta^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{2\pi(\delta^3+\gamma)}\int_0^\infty \left(z+\frac{1}{1+\delta\gamma}z^{-1}\right)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz \\
&= \frac{\gamma\delta^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{2\pi(\delta^3+\gamma)}\int_0^\infty \left(z^{0-1}+\frac{z^{-2-1}}{1+\delta\gamma}\right)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz \\
&= \frac{\gamma\delta^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\delta^3+\gamma)}\left\{K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^2}{\delta^2\phi(x)}\frac{K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{1+\delta\gamma}\right\}
\end{aligned}$$

$$\begin{aligned}
f_{23}(x) &= \frac{\gamma\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi\phi(x)(\delta^3+\gamma)}\left\{\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\alpha^2K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{1+\delta\gamma}\right\} \\
&= \frac{\gamma\delta e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi\phi(x)(\delta^3+\gamma)(1+\delta\gamma)}\Big\{(1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+ \\
&\quad \alpha^2K_2\big(\alpha\delta\sqrt{\phi(x)}\big)\Big\}
\end{aligned} \tag{5.25}$$

where

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

and

$$\alpha^2=\beta^2+\gamma^2$$

5.6.2 The log-likelihood function

$$
\begin{aligned}
l = \log L &= \sum_{i=1}^{n} \log f(x_i) \\
&= \sum_{i=1}^{n} \log \left\{ \frac{\gamma\delta e^{\delta\gamma} e^{\beta(x-\mu)}}{\pi\phi(x)(\delta^3+\gamma)(1+\delta\gamma)} \left\{ (1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big) + \alpha^2 K_2\big(\alpha\delta\sqrt{\phi(x)}\big) \right\} \right\} \\
&= \sum_{i=1}^{n} \big\{ \log(\delta\gamma) + \delta\gamma + \beta x_i - \beta\mu - \log((1+\delta\gamma)\pi(\delta^3+\gamma)) - \log\phi(x_i) + \\
&\quad \log\{(1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big) + \alpha^2 K_2\big(\alpha\delta\sqrt{\phi(x)}\big)\}\big\} \\
&= n\log(\delta\gamma) + n\delta\gamma + \beta\sum_{i=1}^{n} x_i - n\beta\mu - n\log((1+\delta\gamma)\pi(\delta^3+\gamma)) - \sum_{i=1}^{n}\log\phi(x_i) + \\
&\quad \sum_{i=1}^{n} \log\big\{(1+\delta\gamma)\delta^2\phi(x)K_0\big(\alpha\delta\sqrt{\phi(x)}\big) + \alpha^2 K_2\big(\alpha\delta\sqrt{\phi(x)}\big)\big\} \qquad \textbf{(5.26)}
\end{aligned}
$$

5.6.3 Properties of Model 4

$$
E(X) = \mu + \beta \frac{\delta^3(1+\delta\gamma)^2 + \delta^2\gamma^3}{\gamma^2(1+\delta\gamma)(\delta^3+\gamma)} \qquad \textbf{(5.27)}
$$

$$
\begin{aligned}
var(X) &= \frac{\delta^3(1+\delta\gamma)^2 + \delta^2\gamma^3}{\gamma^2(1+\delta\gamma)(\delta^3+\gamma)} \\
&\quad -\beta^2 \frac{\delta^2((\delta^3\gamma^2 + 3\delta^2\gamma + 3\delta)(1+\delta\gamma) + \delta\gamma^4)(1+\delta\gamma)(\delta^3+\gamma) - \delta^4(\delta(1+\delta\gamma)^2 + \gamma^3)^2}{\gamma^4(1+\delta\gamma)^2(\delta^3+\gamma)^2} \qquad \textbf{(5.28)}
\end{aligned}
$$

5.6.4 Posterior Expectations

$$
\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z\big(z+\frac{z^{-1}}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^{-1}}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{1-1}+\frac{z^{-1-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{-2-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\left[\frac{\delta\sqrt{\phi(x)}}{\alpha}+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-1}\right]K_1(\alpha\delta\sqrt{\phi(x)})}{K_0(\alpha\delta\sqrt{\phi(x)})+\left(\frac{\delta\sqrt{\phi(x)}}{\alpha}\right)^{-2}\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}} \\
&= \frac{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}+\big(\frac{\alpha}{\delta\sqrt{\phi(x)}}\big)\big]K_1(\alpha\delta\sqrt{\phi(x)})}{K_0(\alpha\delta\sqrt{\phi(x)})+\big(\frac{\alpha^2}{\delta^2\phi(x)}\big)\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}} \\
&= \frac{[(1+\delta\gamma)\delta^3(\sqrt{\phi(x)})^3+\alpha^2\delta\sqrt{\phi(x)}]K_1(\alpha\delta\sqrt{\phi(x)})}{\alpha(1+\delta\gamma)\delta^2\phi(x)K_0(\alpha\delta\sqrt{\phi(x)})+\alpha^3K_2(\alpha\delta\sqrt{\phi(x)})} \qquad \textbf{(5.29)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}\big(z+\frac{z^{-1}}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^{-1}}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{-3-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{-2-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}
\end{aligned}
$$

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+\big(\frac{\alpha}{\delta\sqrt{\phi(x)}}\big)^3\frac{K_3(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{K_0(\alpha\delta\sqrt{\phi(x)})+\big(\frac{\alpha}{\delta\sqrt{\phi(x)}}\big)^2\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}} \\
&= \frac{\alpha\delta^2\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})+\frac{\alpha^3}{1+\delta\gamma}K_3(\alpha\delta\sqrt{\phi(x)})}{(\delta\sqrt{\phi(x)})^3K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\alpha^2\delta\sqrt{\phi(x)}}{1+\delta\gamma}K_2(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta^2(1+\delta\gamma)\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})+\alpha^3K_3(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)(\delta\sqrt{\phi(x)})^3K_0(\alpha\delta\sqrt{\phi(x)})+\alpha^2\delta\sqrt{\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)})} \qquad \textbf{(5.30)}
\end{aligned}
$$

$$
\begin{aligned}
E\left(Z^2/X\right) &= \frac{\int_0^\infty z^2\left(z+\frac{z^{-1}}{1+\delta\gamma}\right)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \left(z+\frac{z^{-1}}{1+\delta\gamma}\right)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \left((1+\delta\gamma)z^{2-1}+z^{0-1}\right)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \left((1+\delta\gamma)z^{0-1}+z^{-2-1}\right)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{(1+\delta\gamma)\frac{\delta^2\phi(x)}{\alpha^2}K_2(\alpha\delta\sqrt{\phi(x)})+K_0(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)K_0(\alpha\delta\sqrt{\phi(x)})+K_2(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{(1+\delta\gamma)\delta^2\phi(x)K_2(\alpha\delta\sqrt{\phi(x)})+\alpha^2K_0(\alpha\delta\sqrt{\phi(x)})}{\alpha^2(1+\delta\gamma)K_0(\alpha\delta\sqrt{\phi(x)})+\alpha^2K_2(\alpha\delta\sqrt{\phi(x)})} \qquad (5.31)
\end{aligned}
$$

5.7 Model 5

5.7.1 Construction

Suppose the mixing distribution follows formula (3.39), the mixed model becomes

$$
\begin{aligned}
f_{24}(x) &= \frac{\delta e^{\delta\gamma}}{2\pi}e^{\beta(x-\mu)}\int_0^\infty \frac{w(z)}{E[w(Z)]}z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz \\
&= \frac{\delta e^{\delta\gamma}}{2\pi}e^{\beta(x-\mu)}\frac{\gamma^3}{\delta(\gamma^2+1)}\int_0^\infty \left(z+\frac{z^2}{1+\delta\gamma}\right)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz \\
&= \frac{\gamma^3 e^{\delta\gamma}}{2\pi(\gamma^2+1)}e^{\beta(x-\mu)}\int_0^\infty \left(z^{0-1}+\frac{z^{1-1}}{1+\delta\gamma}\right)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz
\end{aligned}
$$

$$
\begin{aligned}
f_{24}(x) &= \frac{\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\gamma^2+1)}\left\{K_0\left(\alpha\delta\sqrt{\phi(x)}\right)+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)\right\} \\
&= \frac{\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\alpha\pi(1+\delta\gamma)(\gamma^2+1)}\left\{\alpha(1+\delta\gamma)K_0\left(\alpha\delta\sqrt{\phi(x)}\right)+\delta\sqrt{\phi(x)}\times\right. \\
&\quad \left. K_1\left(\alpha\delta\sqrt{\phi(x)}\right)\right\} \qquad (5.32)
\end{aligned}
$$

5.7.2 The log-likelihood

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

$$\begin{aligned}
l = \log L &= \sum_{i=1}^{n} \log f(x_i) \\
&= \sum_{i=1}^{n} \Big\{ 3\log\gamma + \delta\gamma + \beta(x_i - \mu) - \log(\alpha\pi(\gamma^2+1)(1+\delta\gamma)) + \\
&\quad \log\Big[\alpha(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big) + \delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)\Big]\Big\} \\
&= 3n\log\gamma + n\delta\gamma + \beta\sum_{i=1}^{n} x_i - n\beta\mu - n\log(\alpha\pi(\gamma^2+1)(1+\delta\gamma)) + \\
&\quad \sum_{i=1}^{n} \log\Big[\alpha(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big) + \delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)\Big] \qquad (5.33)
\end{aligned}$$

5.7.3 Properties of Model 5

$$E(X) = \mu + \beta\frac{\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)} \qquad (5.34)$$

$$\begin{aligned}
var(X) &= \frac{\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3}{\gamma^2(1+\delta\gamma)(\gamma^2+1)} \\
&\quad -\beta^2\frac{((\delta^2\gamma^2+3\delta\gamma+3)(1+\delta\gamma)\gamma^2 + (\delta^3\gamma^3+6\delta^2\gamma^2+15\delta\gamma+15))(1+\delta\gamma)(\gamma^2+1)}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2} \\
&\quad -\frac{(\gamma^2(1+\delta\gamma)^2 + \delta^2\gamma^2 + 3\delta\gamma + 3)^2}{\gamma^4(1+\delta\gamma)^2(\gamma^2+1)^2} \qquad (5.35)
\end{aligned}$$

5.7.4 Posterior Expectation

$$
\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z\big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\int_0^\infty \big(z^{1-1}+\frac{z^{2-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z^{0-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{\delta\sqrt{\phi(x)}}{\alpha}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^2\frac{K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{1+\delta\gamma}}{K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{\delta\sqrt{\phi(x)}(1+\delta\gamma)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\delta^2\phi(x)(1+\delta\gamma)}{\alpha}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{\alpha(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{\alpha\delta\sqrt{\phi(x)}(1+\delta\gamma)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\delta\phi(x)(1+\delta\gamma)K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{\alpha^2(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)}
\end{aligned}
\tag{5.36}
$$

Similarly,

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}\big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{0-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-1}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})}
\end{aligned}
$$

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\frac{\alpha^2(1+\delta\gamma)}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+\alpha K_0(\alpha\delta\sqrt{\phi(x)})}{\alpha(1+\delta\gamma)K_0(\alpha\delta\sqrt{\phi(x)})+\alpha\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^2(1+\delta\gamma)K_1(\alpha\delta\sqrt{\phi(x)})+\alpha\delta\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})}{\alpha\delta\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})+\delta^2\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})}
\end{aligned}
\tag{5.37}
$$

Similarly

$$
E(Z^2/X=x)=\frac{\alpha(1+\delta\gamma)\delta^2\phi(x)K_2(\alpha\delta\sqrt{\phi(x)})+(\delta\sqrt{\phi(x)})^3K_3(\alpha\delta\sqrt{\phi(x)})}{\alpha^3(1+\delta\gamma)K_0(\alpha\delta\sqrt{\phi(x)})+\alpha^2\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})}
\tag{5.38}
$$

5.8 Model 6

5.8.1 Construction

Suppose the mixing distribution follows formula (3.45), the mixed model becomes

$$
\begin{aligned}
f_{34}(x) &= \frac{\delta^3\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\delta^3+\gamma^3)(1+\delta\gamma)}\frac{1}{2}\int_0^\infty (z^{-1}+z^2)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz \\
&= \frac{\delta^3\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\delta^3+\gamma^3)(1+\delta\gamma)}\frac{1}{2}\int_0^\infty (z^{-2-1}+z^{1-1})e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz \\
&= \frac{\delta^3\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\delta^3+\gamma^3)(1+\delta\gamma)}\left\{\frac{\alpha^2}{\delta^2\phi(x)}K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+\frac{\delta\sqrt{\phi(x)}}{\alpha}K_1\left(\alpha\delta\sqrt{\phi(x)}\right)\right\} \\
&= \frac{\delta^3\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\pi(\delta^3+\gamma^3)(1+\delta\gamma)}\left\{\frac{\alpha^3K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+(\delta\sqrt{\phi(x)})^3K_1\left(\alpha\delta\sqrt{\phi(x)}\right)}{\alpha\delta^2\phi(x)}\right\} \\
&= \frac{\delta\gamma^3 e^{\delta\gamma}e^{\beta(x-\mu)}}{\alpha\pi\phi(x)(\delta^3+\gamma^3)(1+\delta\gamma)}\left\{\alpha^3K_2\left(\alpha\delta\sqrt{\phi(x)}\right)+(\delta\sqrt{\phi(x)})^3K_1\left(\alpha\delta\sqrt{\phi(x)}\right)\right\}
\end{aligned}
\tag{5.39}
$$

5.8.2 The log-likelihood function

$$
\begin{aligned}
l=\log L &= \sum_{i=1}^n \log f(x_i) \\
&= \sum_{i=1}^n \log\left\{\frac{\delta\gamma^3 e^{\delta\gamma}e^{\beta(x_i-\mu)}}{\pi(\delta^3+\gamma^3)(1+\delta\gamma)}\left\{\alpha^3K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)+(\delta\sqrt{\phi(x_i)})^3K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)\right\}\right\} \\
&= \sum_{i=1}^n \left\{3\log\gamma+\log\delta+\delta\gamma+\beta(x_i-\mu)-\log(\alpha\pi(\delta^3+\gamma^3)(1+\delta\gamma))-\log\phi(x_i)+\right. \\
&\qquad \left.\log\left\{\alpha^3K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)+(\delta\sqrt{\phi(x_i)})^3K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)\right\}\right\} \\
&= 3n\log\gamma+n\log\delta+n\delta\gamma+\beta\sum_{i=1}^n(x_i-\mu)-n\log(\alpha\pi(\delta^3+\gamma^3)(1+\delta\gamma))-\sum_{i=1}^n\log\phi(x_i)+ \\
&\qquad \sum_{i=1}^n\log\left\{\alpha^3K_2\left(\alpha\delta\sqrt{\phi(x_i)}\right)+(\delta\sqrt{\phi(x_i)})^3K_1\left(\alpha\delta\sqrt{\phi(x_i)}\right)\right\}
\end{aligned}
\tag{5.40}
$$

5.8.3 Properties of Model 6

$$\begin{aligned}
E(X) &= \mu+\beta E(Z) \\
&= \mu+\beta\delta^2\left(\frac{1}{1+\delta\gamma}+\frac{3\delta}{\gamma^2(\gamma^3+\delta^3)}\right) \qquad (5.41)\\
var(X) &= E(Z)+\beta^2 Var(Z) \\
&= \frac{3(1+\delta\gamma)^2\delta^3[\alpha^2(\gamma^3+\delta^3)+\beta^2(4\gamma^3+\delta^3)]+a(\delta,\gamma)}{\gamma^4(\gamma^3+\delta^3)^2(1+\delta\gamma)^2} \qquad (5.42)
\end{aligned}$$

where

$$a(\delta,\gamma)=\delta^2\gamma^9(\alpha^2\delta+\gamma)+(2\gamma^3+\delta^3)\delta^5\gamma^2[\alpha^2(1+\delta\gamma)-\beta^2] \qquad (5.43)$$

5.8.4 Posterior Expectation

$$\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z(z^{-1}+z^2)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty (z^{-1}+z^2)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\int_0^\infty (z^{-1-1}+z^{2-1})e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty (z^{-2-1}+z^{1-1})e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{\alpha}{\delta\sqrt{\phi(x)}}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^2\phi(x)}{\alpha^2}K_2(\alpha\delta\sqrt{\phi(x)})}{\frac{\alpha^2}{\delta^2\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\phi(x)}{\alpha}K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^2\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{(\delta^2\phi(x))^2}{\alpha}K_2(\alpha\delta\sqrt{\phi(x)})}{\alpha^3K_2(\alpha\delta\sqrt{\phi(x)})+\delta^3(\phi(x))^{\frac{3}{2}}K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^3\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})+(\delta^2\phi(x))^2K_2(\alpha\delta\sqrt{\phi(x)})}{\alpha^4K_2(\alpha\delta\sqrt{\phi(x)})+\alpha\delta^3(\phi(x))^{\frac{3}{2}}K_1(\alpha\delta\sqrt{\phi(x)})} \qquad (5.44)
\end{aligned}$$

and

$$\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}(z^{-1}+z^2)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty (z^{-1}+z^2)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\int_0^\infty (z^{-3-1}+z^{0-1})e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty (z^{-2-1}+z^{1-1})e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{\alpha^3}{(\delta\sqrt{\phi(x)})^3}K_3(\alpha\delta\sqrt{\phi(x)})+K_0(\alpha\delta\sqrt{\phi(x)})}{\frac{\alpha^2}{\delta^2\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\phi(x)}{\alpha}K_1(\alpha\delta\sqrt{\phi(x)})}
\end{aligned}$$

$$E\left(\frac{1}{Z}/X=x\right) = \frac{\alpha^3 K_3(\alpha\delta\sqrt{\phi(x)}) + (\delta\sqrt{\phi(x)})^3 K_0(\alpha\delta\sqrt{\phi(x)})}{\delta\alpha^2\sqrt{\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)}) + \frac{\delta^4[\phi(x)]^2}{\alpha}K_1(\alpha\delta\sqrt{\phi(x)})}$$

$$= \frac{\alpha^4 K_3(\alpha\delta\sqrt{\phi(x)}) + \alpha(\delta\sqrt{\phi(x)})^3 K_0(\alpha\delta\sqrt{\phi(x)})}{\delta\alpha^3\sqrt{\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)}) + (\delta^2\phi(x))^2 K_1(\alpha\delta\sqrt{\phi(x)})} \quad \textbf{(5.45)}$$

The loglikelihood and posterior expectations derived in this chapter, will be used in maximum likelihood estimation of parameters via the Expectation-Maximization (EM) Algorithm discussed in chapter 7.

6 Iterative Schemes Designs for NWIG Distributions Based on EM Algorithm

6.1 Introduction

Parametric methods commonly used in parameter estimation are Method of Moments (MoM) and Maximum Likelihood (ML) method. However, these methods have some limitations. Equations obtained by these methods require complex numerical techniques to solve in cases where the parameters are hard to separate.
Alternative simple methods have been sought. One such method is the Expectation-Maximization $EM-algorithm$. In this chapter we apply this concept to the NWIG distributions. In the next section we briefly define the theory behind this algorithm.

6.2 The Expectation-Maximization (EM) Algorithm

The EM-algorithm was introduced by Dempster et. al (1977) for ML estimation for data containing missing values or data that can be considered as producing missing values. Karlis (2002) pointed out that the mixing operation can be considered responsible for producing missing data. The statistical beauty of the $EM-algorithm$ is that it estimates the unobserved values using the posterior expectations.
It becomes easier if we exploit the normal variance mean structure of the mixtures through this algorithm. Assume that true data are made of observed part X and unobserved part Z. This then ensures the log likelihood of the complete data $(x_i, z_i), i = 1, 2, 3, ..., n$ factorises into two parts (Kostas, 2007). The $EM-algorithm$ consists of two main steps: the maximization step which optimises the loglikelihoods with respect to the parameters and the Expectation step which estimates the unobserved values using the posterior expectations. This implies that the joint density of X and Z is given by:

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

Therefore the likelihood function for the joint data becomes;

$$L(\alpha, \beta, \delta, \mu) = \prod_{i=1}^{n} f(x_i/z_i) \prod_{i=1}^{n} f(z_i)$$

and the log likelihood

$$\begin{aligned} l(\alpha, \beta, \delta, \mu) &= \sum_{i=1}^{n} \log f(x_i/z_i) + \sum_{i=1}^{n} \log f(z_i) \\ &= l_1(\mu, \beta) + l_2(\delta, \gamma) \end{aligned}$$

For these models,

$$l_1(\mu,\beta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log z_i - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu-\beta z)^2}{z_i} \tag{6.1}$$

6.2.1 M-step

In this step the log likelihood of the conditional and mixing distributions are optimised with respect to their parameters: μ, β, δ and γ. Since the conditional distribution is common for all these models, we proceed as follows:

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

Equating these equations to zero and solving simultaneously, we obtain

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

and hence

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

Therefore at the k-th iteration of the algorithm, the estimate for β and μ are:

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n}\frac{x_i}{z_i} - \bar{x}\sum_{i=1}^{n}\frac{1}{z_i}}{n-\bar{z}\sum_{i=1}^{n}\frac{1}{z_i}} \tag{6.2}$$

$$\mu^{(k+1)} = \bar{x} - \hat{\beta}^{(k+1)}\bar{z} \tag{6.3}$$

In general the estimates are in terms of Values of random variables Z_i, $\frac{1}{Z_i}$, z_i^2, etc which are unknown. So we estimate them by considering posterior expectations $E(Z_i/X_i)$, $E\left(\frac{1}{Z_i}\right)$, $E(z_i^2)$, and this amounts to the **E-step**.

6.3 Maximization of the mixing distribution for the Mixed Models

In this section we design the iterative schemes for the mixing distributions considered before.

6.3.1 Inverse Gaussian Distribution

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

Therefore

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

optimizing w.r.t δ and γ we get

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

and

$$\hat{\gamma} = \frac{\hat{\delta}}{\bar{z}}$$

E-step

We estimate the missing values z_i and $\frac{1}{z_i}$ by considering posterior expectations $E\left(Z_i/X_i\right)$ and $E\left(Z_i^{-1}/X_i\right)$ respectively. For the NIG distribution the posterior distribution of Z/X is

$$\begin{aligned} E[Z/X=x] &= \frac{\int_0^\infty z^{-1} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}dz}{\int_0^\infty z^{-1-1} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}dz} \\ &= \frac{\delta\sqrt{\phi(x)}}{\alpha}\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{K_1(\alpha\delta\sqrt{\phi(x)})} \end{aligned}$$

Next

$$\begin{aligned} E[Z^{-1}/X=x] &= \frac{\int_0^\infty z^{-3} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}dz}{\int_0^\infty z^{-2} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}dz} \\ &= \frac{\alpha}{\delta\sqrt{\phi(x)}}\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{K_1(\alpha\delta\sqrt{\phi(x)})} \end{aligned}$$

Iteration

Defining $s_i = E(Z_i/X_i)$ and $w_i = E(Z_i^{-1}/X_i))$ which are all functions of α, β, δ and μ for computation we proceed as follows: for the k-th iteration we have

$$s_i^{(k)} = \frac{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}}{\alpha^{(k)}} \frac{K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.4}$$

and

$$w_i^{(k)} = \frac{\alpha^{(k)}}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}} \frac{K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.5}$$

We then compute

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^n (x_i - \bar{x}) w_i^{(k)}}{n - \frac{1}{n}\sum_{i=1}^n s_i^{(k)} \sum_{i=1}^n w_i^{(k)}} \tag{6.6}$$

$$\hat{\delta}^{(k+1)} = \sqrt{\frac{n}{\sum_{i=1}^n w_i^{(k)} - \frac{n^2}{\sum_{i=1}^n s_i^{(k)}}}} \tag{6.7}$$

$$\hat{\gamma}^{(k+1)} = \frac{\delta^{(k+1)}}{\frac{1}{n}\sum_{i=1}^n s_i^{(k)}} \tag{6.8}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \beta^{(k+1)} \frac{1}{n} \sum_{i=1}^n s_i^{(k)} \tag{6.9}$$

$$\hat{\alpha}^{(k+1)} = [(\hat{\gamma}^{(k+1)})^2 + (\hat{\beta}^{(k+1)})^2]^{\frac{1}{2}} \tag{6.10}$$

Remark: When $k = 0$, we have the initial values.

6.3.2 Length Biased (Reciprocal) Inverse Gaussian Distribution

M-step

$$\begin{aligned} GIG(\frac{1}{2}, \delta, \gamma) &= \frac{\gamma}{\delta} z GIG(-\frac{1}{2}, \delta, \gamma) \\ g_2(z) &= \frac{\gamma}{\sqrt{2\pi}} \exp(\delta\gamma) z^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\frac{\delta^2}{z} + \gamma^2 z) \right\} \end{aligned}$$

$$\therefore l_2(\delta, \gamma) = n \log \gamma - \frac{n}{2} \log(2\pi) + n\delta\gamma - \frac{\delta^2}{2} \sum_{i=1}^n \frac{1}{z_i} - \frac{1}{2} \sum_{i=1}^n \log z_i \tag{6.11}$$

Differentiating w.r.t δ and γ and optimizing we obtain

$$\hat{\delta} = \frac{n\gamma}{\sum_{i=1}^{n} \frac{1}{z_i}} \tag{6.12}$$

$$\hat{\gamma} = \sqrt{\frac{n}{\sum_{i=1}^{n} z_i - n^2 (\sum_{i=1}^{n} \frac{1}{z_i})^{-1}}} \tag{6.13}$$

E-step

We now wish to estimate z_i and z_i^{-1} using posterior expectation

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

Using the transformation $z = \frac{\delta}{\alpha}\sqrt{\phi(x)}t$ we obtain

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

and

$$\begin{aligned} E[Z^{-1}/X=x] &= \frac{\int_0^\infty z^{-2} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}\,dz}{\int_0^\infty z^{-1} e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2}{\alpha^2}\frac{\phi(x)}{z}\right)}\,dz} \\ &= \frac{\alpha}{\delta\sqrt{\phi(x)}}\frac{K_1(\alpha\delta\sqrt{\phi(x)})}{K_0(\alpha\delta\sqrt{\phi(x)})} \end{aligned}$$

Let $s_i = E(Z_i/X_i)$ and $w_i = E(Z_i^{-1}/X_i)$. For iteration we have

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})w_i^{(k)}}{n - \frac{1}{n}\sum_{i=1}^{n} s_i^{(k)} \sum_{i=1}^{n} w_i^{(k)}} \tag{6.14}$$

$$\hat{\gamma}^{(k+1)} = \sqrt{\frac{n}{\sum_{i=1}^{n} s_i^{(k)} - \frac{n^2}{\sum_{i=1}^{n} w_i^{(k)}}}} \tag{6.15}$$

$$\hat{\delta}^{(k+1)} = \frac{\gamma^{(k+1)}}{\frac{1}{n}\sum_{i=1}^{n} w_i^{(k)}} \tag{6.16}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \beta^{(k+1)}\frac{1}{n}\sum_{i=1}^{n} s_i^{(k)} \tag{6.17}$$

$$\hat{\alpha}^{(k+1)} = [(\hat{\gamma}^{(k+1)})^2 + (\hat{\beta}^{(k+1)})^2]^{\frac{1}{2}} \tag{6.18}$$

6.3.3 GIG$(\frac{3}{2}, \delta, \gamma)$

$$g_4(z) = \frac{\gamma^3}{\sqrt{2\pi}} \frac{e^{\delta\gamma} z^{\frac{1}{2}}}{(1+\delta\gamma)} e^{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)}$$

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

$$\begin{aligned}
\frac{\partial}{\partial\gamma} l_2 &= -\frac{n\delta}{(1+\delta\gamma)} + \frac{3n}{\gamma} + n\delta - \gamma\sum_{i=1}^{n} z_i \\
&= n\delta\left(1 - \frac{1}{(1+\delta\gamma)}\right) + \frac{3n}{\gamma} - \gamma\sum_{i=1}^{n} z_i \\
&= \frac{n\delta^2\gamma}{(1+\delta\gamma)} + \frac{3n}{\gamma} - \gamma\sum_{i=1}^{n} z_i \\
\frac{\partial}{\partial\gamma} l_2 = 0 &\Longrightarrow \frac{n\delta^2\gamma}{(1+\delta\gamma)} + \frac{3n}{\gamma} - \gamma\sum_{i=1}^{n} z_i = 0 \qquad (6.19) \\
\frac{\partial}{\partial\delta} l_2 &= -\frac{n\gamma}{1+\delta\gamma} + n\gamma - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
\frac{\partial}{\partial\delta} l_2 = 0 &\Longrightarrow \frac{n\gamma^2\delta}{1+\delta\gamma} - \delta\sum_{i=1}^{n}\frac{1}{z_i} = 0 \qquad (6.20)
\end{aligned}$$

6.3.4 E-STEP: Posterior Expectations

Using Formular (5.3) we have

$$\begin{aligned}
E(Z/X=x) &= \frac{\frac{1}{2}\int_0^\infty z z^2 z^{-2} e^{-\frac{1}{2}\left[\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right]} dz}{\frac{1}{2}\int_0^\infty z^2 z^{-2} e^{-\frac{1}{2}\left[\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right]} dz} \\
&= \frac{\frac{1}{2}\int_0^\infty z^{2-1} e^{-\frac{1}{2}\left[\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right]} dz}{\frac{1}{2}\int_0^\infty z^{1-1} e^{-\frac{1}{2}\left[\alpha^2 z + \frac{\delta^2\phi(x)}{z}\right]} dz} \\
&= \frac{\delta\sqrt{\phi(x)}}{\alpha}\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{K_1(\alpha\delta\sqrt{\phi(x)})} \qquad (6.21)
\end{aligned}$$

Next

$$E(\frac{1}{Z}/X=x)=\frac{\frac{1}{2}\int_0^\infty z^{-1}z^2z^{-2}e^{-\frac{1}{2}\left[\alpha^2z+\frac{\delta^2\phi(x)}{z}\right]}\,dz}{\frac{1}{2}\int_0^\infty z^2z^{-2}e^{-\frac{1}{2}\left[\alpha^2z+\frac{\delta^2\phi(x)}{z}\right]}\,dz}$$
$$=\frac{\frac{1}{2}\int_0^\infty z^{0-1}e^{-\frac{1}{2}\left[\alpha^2z+\frac{\delta^2\phi(x)}{z}\right]}\,dz}{\frac{1}{2}\int_0^\infty z^{1-1}e^{-\frac{1}{2}\left[\alpha^2z+\frac{\delta^2\phi(x)}{z}\right]}\,dz}$$
$$=\frac{\alpha K_0(\alpha\delta\sqrt{\phi(x)})}{\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})} \tag{6.22}$$

6.3.5 Iterative Schemes

From posterior expectations let

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

and

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

The k-th iterations are

$$s_i^{(k)}=\frac{\delta^{(k)}}{\alpha^{(k)}}\sqrt{\phi^{(k)}(x_i)}\frac{K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.23}$$

and

$$w_i^{(k)}=\frac{\alpha^{(k)}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.24}$$

We also define

$$\bar{s}^{(k)}=\frac{\sum_{i=1}^n s_i^{(k)}}{n}$$

We now want to present various iterative schemes.

Scheme 1

This scheme depends on explicit solution to the simultaneous equations as described

below.
From equation (6.19)

$$\hat{\delta} = \frac{n\gamma^2 - \sum_{i=1}^{n}\frac{1}{z_i}}{\hat{\gamma}\sum_{i=1}^{n}\frac{1}{z_i}} \tag{6.25}$$

and equation (6.20)

$$1 + \hat{\delta}\hat{\gamma} = \frac{n\hat{\gamma}^2}{\sum_{i=1}^{n}\frac{1}{z_i}} \tag{6.26}$$

Substituting (6.25) and (6.26) in (6.19), we get the following quadratic equation

$$\left(n^2 - \sum_{i=1}^{n} z_i \sum_{i=1}^{n}\frac{1}{z_i}\right)t^2 + \left(n\sum_{i=1}^{n}\frac{1}{z_i}\right)t + \left(\sum_{i=1}^{n}\frac{1}{z_i}\right)^2 = 0 \tag{6.27}$$

where

$$t = \gamma^2 \tag{6.28}$$

$$= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{6.29}$$

where

$$\begin{aligned} a &= n^2 - \sum_{i=1}^{n} z_i \sum_{i=1}^{n}\frac{1}{z_i} \\ b &= n\sum_{i=1}^{n}\frac{1}{z_i} \\ c &= \left(\sum_{i=1}^{n}\frac{1}{z_i}\right)^2 \\ \therefore \hat{\gamma} &= \sqrt{t} \end{aligned} \tag{6.30}$$

So we have the following iterative formulae

$$\hat{\gamma}^{(k+1)} = \sqrt{t^{(k)}} \tag{6.31}$$

$$\hat{\delta}^{(k+1)} = \frac{n(\hat{\gamma}^{(k+1)})^2 - \sum_{i=1}^{n} w_i^{(k)}}{\hat{\gamma}^{(k+1)}\sum_{i=1}^{n} w_i^{(k)}} \tag{6.32}$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n} x_i w_i^{(k)} - \bar{x}\sum_{i=1}^{n} w_i^{(k)}}{n - \bar{s^{(k)}}\sum_{i=1}^{n} w_i^{(k)}} \tag{6.33}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \hat{\beta}^{(k+1)}\bar{s^{(k)}} \tag{6.34}$$

The other schemes depend on the the following representations of δ and γ
Let us express equation (6.19) as

$$\hat{\gamma} = \left(\frac{\hat{\delta}^2\hat{\gamma}}{1+\hat{\delta}\hat{\gamma}} + \frac{3}{\hat{\gamma}}\right)\frac{1}{\bar{z}}$$

or

$$\hat{\gamma} = \frac{3}{\hat{\gamma}}\left(\bar{z} - \frac{\hat{\delta}}{1+\hat{\delta}\hat{\gamma}}\right)^{-1}$$

together with

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

from equation (6.20)

Scheme 2:
Consider

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

and

$$\hat{\gamma} = \left(\frac{\hat{\delta}^2\hat{\gamma}}{1+\hat{\delta}\hat{\gamma}} + \frac{3}{\hat{\gamma}}\right)\frac{1}{\bar{z}}$$

The corresponding iterations are:

$$\hat{\delta}^{(k+1)} = \frac{n*(\hat{\gamma}^{(k)})^2 - \sum_{i=1}^{n} w_i^{(k)}}{\hat{\gamma}^{(k)}\sum_{i=1}^{n} w_i^{(k)}} \quad \textbf{(6.35)}$$

and

$$\hat{\gamma}^{(k+1)} = \left(\frac{(\hat{\delta}^{(k+1)})^2\hat{\gamma}^{(k)}}{1+\hat{\delta}^{(k+1)}\hat{\gamma}^{(k)}} + \frac{3}{\hat{\gamma}^{(k)}}\right)\frac{1}{\bar{s}^{(k)}} \quad \textbf{(6.36)}$$

In this scheme for (k+1) iteration we first obtain the estimate for delta, $\hat{\delta}^{(k+1)}$, as given in equation (6.35) and then use it to obtain $\hat{\gamma}^{(k+1)}$ as given in equation (6.36). This $\hat{\gamma}^{(k+1)}$ is used to obtain $\hat{\delta}^{(k+2)}$ in updating equation (6.35) and so on until convergence is reached using a certain stopping criterion.
This procedure of updating parameter estimates is the same for scheme 3, scheme 4 and scheme 5 given below.
Scheme 3:

Consider

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

and

$$\hat{\gamma} = \frac{3}{\hat{\gamma}} \left(\bar{z} - \frac{\hat{\delta}^2}{1 + \hat{\delta}\hat{\gamma}} \right)^{-1}$$

The corresponding iterations are:

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

and

$$\hat{\gamma}^{(k+1)} = \frac{3}{\hat{\gamma}^{(k)}} \left(\bar{s}^{(k)} - \frac{(\hat{\delta}^{(k+1)})^2}{1 + \hat{\delta}^{(k+1)} \hat{\gamma}^{(k)}} \right)^{-1}$$

Scheme 4: We Consider

$$\hat{\gamma} = \left(\frac{\hat{\delta}^2 \hat{\gamma}}{1 + \hat{\delta}\hat{\gamma}} + \frac{3}{\hat{\gamma}} \right) \frac{1}{\bar{z}}$$

and

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

The corresponding iterations are

$$\hat{\gamma}^{(k+1)} = \left(\frac{(\hat{\delta}^{(k)})^2 \hat{\gamma}^{(k)}}{1 + \hat{\delta}^{(k)} \hat{\gamma}^{(k)}} + \frac{3}{\hat{\gamma}^{(k)}} \right) \frac{1}{\bar{s}^{(k)}} \tag{6.37}$$

and

$$\hat{\delta}^{(k+1)} = \frac{n * (\hat{\gamma}^{(k+1)})^2 - \sum_{i=1}^{n} w_i^{(k)}}{\hat{\gamma}^{(k+1)} \sum_{i=1}^{n} w_i^{(k)}} \tag{6.38}$$

Scheme 5: Consider

$$\hat{\gamma} = \frac{3}{\hat{\gamma}} \left(\bar{z} - \frac{\hat{\delta}^2}{1 + \hat{\delta}\hat{\gamma}} \right)^{-1}$$

and

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

The corresponding iterations are:

$$\hat{\gamma}^{(k+1)} = \frac{3}{\hat{\gamma}^{(k)}}\left(\bar{s}^{(k)} - \frac{(\hat{\delta}^{(k)})^2}{1+\hat{\delta}^{(k)}\hat{\gamma}^{(k)}}\right)^{-1} \tag{6.39}$$

and

$$\hat{\delta}^{(k+1)} = \frac{n*(\hat{\gamma}^{(k+1)})^2 - \sum_{i=1}^n w_i^{(k)}}{\hat{\gamma}^{(k+1)}\sum_{i=1}^n w_i^{(k)}} \tag{6.40}$$

6.3.6 GIG$(-\frac{3}{2}, \delta, \gamma)$

$$g_3(z) = \frac{\delta^3}{\sqrt{2\pi}(1+\delta\gamma)} z^{-\frac{5}{2}} e^{\delta\gamma} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{z} + \gamma^2 z\right)\right\} \tag{6.41}$$

The loglikelihood of the mixing distribution $l_2(\delta,\gamma)$ is also optimized with respect to δ and γ.

$$\begin{aligned} l_2(\delta,\gamma) &= 3n\log\delta - \frac{n}{2}\log(2\pi) - \frac{5}{2}\sum_{i=1}^n \log z_i - n\log(1+\delta\gamma) + n\delta\gamma - \\ &\quad \frac{1}{2}\sum_{i=1}^n\left(\frac{\delta^2}{z_i} + \gamma^2 z_i\right) \end{aligned} \tag{6.42}$$

the derivative with respect to γ is

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

the derivative with respect to δ is

$$\frac{\partial}{\partial\delta} l_2(\delta,\gamma) = \frac{3n}{\delta} - \frac{n\gamma}{(1+\delta\gamma)} + n\gamma - \delta\sum_{i=1}^n \frac{1}{z_i}$$

equating the first equation to zero and simplifying we obtain

$$\hat{\gamma} = \frac{\delta^2 - \bar{z}}{\delta\bar{z}}$$

where $\bar{z} = \frac{1}{n}\sum_{i=1}^n z_i$. Substituting for γ in the second equation and equating and simplifying we obtain

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

note

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

We therefore have

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

Letting $t=\delta^2$, we obtain the quadratic equation

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

where

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

with

$$\begin{aligned} a &= n^2-\sum_{i=1}^{n} z_i\sum_{i=1}^{n}\frac{1}{z_i} \\ b &= n\sum_{i=1}^{n} z_i \\ c &= \left(\sum_{i=1}^{n} z_i\right)^2 \end{aligned}$$

therefore $\delta=\sqrt{t}$ hence, at the $k-th$ iteration of the algorithm, the estimate for δ and γ are

$$\delta^{(k+1)} = \sqrt{t} \tag{6.43}$$

$$\gamma^{(k+1)} = \frac{(\delta^{(k+1)})^2-\bar{z}}{\delta^{(k+1)}\bar{z}} \tag{6.44}$$

These estimates involves computation of unknown values for random variables: $\bar{Z}$, $\{Z_i, i=1,2,...,n\}$ and $\{Z_i^{-1}, i=1,2,...,n\}$. Estimation of the values of these random variables amounts to performing the $E-step$.

E-step

The estimation of the random variables: $\bar{Z}$, $\{Z_i, i=1,2,...,n\}$ and $\{Z_i^{-1}, i=1,2,...,n\}$ is achieved by computing the posterior expectation for $E(Z_i/X_i=x_i)$ and $E(Z_i^{-1}/X_i=x_i)$ using posterior distribution. One attractive and useful feature of the GIG distribution in the mixing mechanism is that its conjugate for the Normal distribution. That is, given a conditional distribution $X/Z \sim N(\mu+\beta z, z)$ and the mixing/prior distribution to be $Z \sim GIG(\lambda,\delta,\gamma)$ the posterior distribution is $Z/X \sim GIG(\lambda-\frac{1}{2},\sqrt{\delta^2+(x-\mu)^2},\alpha)$ where $\alpha=\sqrt{\beta^2+\gamma^2}$. It can easily be shown the moments around the origin of the $GIG(\lambda,\delta,\gamma)$ distribution are given by

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

and this formula hold for negative values of r,i.e for inverse moments too. When mixing with $Z \sim GIG(-\frac{3}{2},\delta,\gamma)$ posterior distribution becomes $Z \sim GIG(-2,\sqrt{\delta^2+(x-\mu)^2},\alpha)$. The posterior expectation required can be computed as:

$$E(Z/X=x)=\frac{\sqrt{\delta^2+(x-\mu)^2}}{\alpha}\frac{K_1(\alpha\sqrt{\delta^2+(x-\mu)^2})}{K_2(\alpha\sqrt{\delta^2+(x-\mu)^2})}$$

$$E(Z^{-1}/X=x)=\frac{\alpha}{\sqrt{\delta^2+(x-\mu)^2}}\frac{K_{-3}(\alpha\sqrt{\delta^2+(x-\mu)^2})}{K_{-2}(\alpha\sqrt{\delta^2+(x-\mu)^2})}$$

These posterior expectation can now be used to compute the parameter estimates via the EM-algorithm for Maximum likelihood estimation of the parameters for the proposed mixed model distribution. Let s_i denote $E(Z_i/X_i=x_i,\theta^{(k)})$ and w_i denote $E(Z_i^{-1}/X_i=x_i,\theta^{(k)})$ where $\theta^{(k)}$ denote the k-th iteration values of the $GIG(-\frac{3}{2},\delta,\gamma)$. Therefore

$$s_i=\frac{\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}}}{\alpha^{(k)}}\frac{K_1(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}})}{K_2(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}})}$$

$$w_i=\frac{\alpha^k}{\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}}}\frac{K_{-3}(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}})}{K_{-2}(\alpha^{(k)}\delta^{(k)}\phi^{(k)}(x_i)^{\frac{1}{2}})}$$

for $i=1,2,...,n$ and $\phi^{(k)}(x)=1+\frac{(x-\mu^{(k)})^2}{(\delta^{(k)})^2}$

Pseudovalues calculated at the E-step can now be used to update the other parameters as follows:

$$\hat{\delta}^{(k+1)} = \sqrt{t} \quad \textbf{(6.45)}$$

$$\hat{\gamma}^{(k+1)} = \frac{(\delta^{(k+1)})^2-\hat{M}}{\delta^{(k+1)}\hat{M}} \quad \textbf{(6.46)}$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^n x_i w_i-\bar{x}\sum_{i=1}^n w_i}{n-\bar{s}\sum_{i=1}^n w_i} \quad \textbf{(6.47)}$$

$$\hat{\mu}^{(k+1)} = \bar{x}-\hat{\beta}^{(k+1)}\bar{s} \quad \textbf{(6.48)}$$

$$\hat{\alpha}^{(k+1)} = \sqrt{(\hat{\gamma}^{(k+1)})^2+(\hat{\beta}^{(k+1)})^2} \quad \textbf{(6.49)}$$

We now consider the Six Models of the finite Inverse Gaussian

6.3.7 M-Step for the Mixing Distribution of Model 1

The mixing distribution given by equation (3.15) is

$$\begin{aligned} g(z) &= \frac{\gamma}{\gamma+\delta}(1+z)g_1(z) \\ &= \frac{\gamma}{\gamma+\delta}(1+z)\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}}z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)} \\ &= \frac{\delta\gamma e^{\delta\gamma}}{\sqrt{2\pi}(\gamma+\delta)}(1+z)z^{-\frac{3}{2}}e^{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)} \end{aligned} \tag{6.50}$$

Therefore

$$l_2 = \sum_{i=1}^{n} \log z_i \tag{6.51}$$

$$\begin{aligned} &= \sum_{i=1}^{n} \log \frac{\delta\gamma e^{\delta\gamma}}{\sqrt{2\pi}(\gamma+\delta)}(1+z)z^{-\frac{3}{2}}e^{-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)} \\ &= \sum_{i=1}^{n}\left\{\log\delta\gamma+\delta\gamma-\frac{1}{2}\log(2\pi)-\log(\gamma+\delta)+\log(1+z_i)-\frac{3}{2}\log(z_i)-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)\right\} \\ &= n\delta\gamma+n\log\delta\gamma-\frac{n}{2}\log(2\pi)-n\log(\gamma+\delta)+\sum_{i=1}^{n}\log(1+z_i)-\frac{3}{2}\sum_{i=1}^{n}\log(z_i)- \\ &\quad \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i}-\frac{\gamma^2}{2}\sum_{i=1}^{n}z_i \end{aligned} \tag{6.52}$$

Therefore

$$\begin{aligned} \frac{\partial}{\partial\delta}l_2 &= n\gamma+\frac{n\gamma}{\delta\gamma}-\frac{n}{\gamma+\delta}-\delta\sum_{i=1}^{n}\frac{1}{z_i} \\ &= n\gamma+\frac{n}{\delta}-\frac{n}{\gamma+\delta}-\delta\sum_{i=1}^{n}\frac{1}{z_i} \\ \frac{\partial}{\partial\delta}l_2 = 0 &\Longrightarrow n\gamma+\frac{n}{\delta}-\frac{n}{\gamma+\delta}-\delta\sum_{i=1}^{n}\frac{1}{z_i}=0 \\ &\therefore n\gamma+\frac{n(\gamma+\delta-\delta)}{\delta(\gamma+\delta)}=\delta\sum_{i=1}^{n}\frac{1}{z_i} \\ &\therefore n\gamma+\frac{n\gamma}{\delta(\gamma+\delta)}=\delta\sum_{i=1}^{n}\frac{1}{z_i} \\ &\therefore n\gamma[1+\frac{1}{\delta(\gamma+\delta)}]=\delta\sum_{i=1}^{n}\frac{1}{z_i} \end{aligned}$$

Therefore

$$\hat{\delta} = \frac{\hat{\gamma}[1+\frac{1}{\hat{\delta}(\hat{\gamma}+\hat{\delta})}]}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{z_i}} \tag{6.53}$$

Similarly

$$\frac{\partial}{\partial\gamma}l_2 = n\delta + \left(\frac{n}{\gamma} - \frac{n}{\gamma+\delta}\right) - \gamma\sum_{i=1}^{n} z_i$$
$$= n\delta + \frac{n\delta}{\gamma(\gamma+\delta)} - \gamma\sum_{i=1}^{n} z_i$$
$$\frac{\partial}{\partial\gamma}l_2 = 0 \Longrightarrow n\delta\left[1 + \frac{1}{\gamma(\gamma+\delta)}\right] = \gamma\sum_{i=1}^{n} z_i$$

Therefore

$$\hat{\gamma} = \frac{\hat{\delta}\left[1 + \frac{1}{\hat{\gamma}(\hat{\gamma}+\hat{\delta})}\right]}{\frac{1}{n}\sum_{i=1}^{n} z_i} \tag{6.54}$$

6.3.8 E-Step

Since Z and $\frac{1}{Z}$ are unobserved; they are therefore the missing values. Hence they are estimated by posterior expectations

$$E(Z/X) \text{ and } E\left(\frac{1}{Z}/X\right)$$

as given in sub-section (6.5)

6.3.9 Iterative Schemes

Let

$$s_i = E(Z_i/X_i)$$

and

$$w_i = E\left(\frac{1}{Z_i}/X_i\right)$$

Then the $k-th$ iterations are as follows:

$$s_i^{(k)} = \frac{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i))}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + \left(\delta^{(k)}\right)^2\phi^{(k)}(x_i)K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + \left(\alpha^{(k)}\right)^2K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.55}$$

$$w_i^{(k)} = \frac{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + (\alpha^{(k)})^2K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})}{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + (\delta^{(k)})^2\phi^{(k)}(x_i)K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})} \tag{6.56}$$

For the log-likelihood function the $k-th$ iteration is given as

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

Iterative Scheme 1

From equations (67) and (68), we obtain the iterative scheme

$$\begin{aligned} \hat{\delta}^{(k+1)} &= \frac{\hat{\gamma}^{(k)}\left[1+\frac{1}{\hat{\delta}^{(k)}(\hat{\gamma}^{(k)}+\hat{\delta}^{(k)})}\right]}{\frac{1}{n}\sum_{i=1}^{n} w_i^{(k)}} \\ &= \frac{\hat{\gamma}^{(k)}\left[1+\frac{1}{\hat{\delta}^{(k)}(\hat{\gamma}^{(k)}+\hat{\delta}^{(k)})}\right]}{\bar{w}^{(k)}} \end{aligned}$$

and

$$\hat{\gamma}^{(k+1)} = \frac{\hat{\delta}^{(k+1)}\left[1+\frac{1}{\hat{\gamma}^{(k)}(\hat{\gamma}^{(k)}+\hat{\delta}^{(k+1)})}\right]}{\bar{s}^{(k)}}$$

where

$$\bar{w} = \sum_{i=1}^{n}\frac{w_i}{n}$$

and

$$\bar{s} = \sum_{i=1}^{n}\frac{s_i}{n}$$

$$\begin{aligned} \hat{\beta}^{(k+1)} &= \frac{\sum_{i=1}^{n} x_i w_i^{(k)} - \bar{x}\sum_{i=1}^{n} w_i^{(k)}}{n - s^{\bar{(k)}}\sum_{i=1}^{n} w_i^{(k)}} \\ &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})w_i^{(k)}}{n - s^{\bar{(k)}}\sum_{i=1}^{n} w_i^{(k)}} \\ &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})w_i^{(k)}}{n(1 - s^{\bar{(k)}}\bar{w}^{(k)})} \\ \hat{\mu}^{(k+1)} &= \bar{x} - \hat{\beta}^{(k+1)}s^{\bar{(k)}} \\ \hat{\alpha}^{(k+1)} &= \left[(\hat{\gamma}^{(k+1)})^2 + (\hat{\beta}^{(k+1)})^2\right]^{\frac{1}{2}} \end{aligned}$$

Iteration Scheme 2

Again, using equations (67) and (68) we can also obtain the following iterative scheme

$$\hat{\gamma}^{(k+1)} = \frac{\hat{\delta}^{(k)}\left[1+\frac{1}{\hat{\gamma}^{(k)}(\hat{\gamma}^{(k)}+\hat{\delta}^{(k)})}\right]}{\bar{s}^{(k)}} \tag{6.57}$$

$$\hat{\delta}^{(k+1)} = \frac{\hat{\gamma}^{(k+1)}\left[1+\frac{1}{\hat{\delta}^{(k)}(\hat{\gamma}^{(k+1)}+\hat{\delta}^{(k)})}\right]}{\bar{w}^{(k)}} \tag{6.58}$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})w_i^{(k)}}{n(1-s^{\bar{(k)}}\bar{w}^{(k)})} \tag{6.59}$$

$$\hat{\mu}^{(k+1)} = \bar{x}-\hat{\beta}^{(k+1)}s^{\bar{(k)}} \tag{6.60}$$

$$\hat{\alpha}^{(k+1)} = \left[(\hat{\gamma}^{(k+1)})^2+(\hat{\beta}^{(k+1)})^2\right]^{\frac{1}{2}} \tag{6.61}$$

6.3.10 M-Step for the Mixing Distribution for Model 2

The mixing distribution is presented by equation (3.21) presented as

$$\begin{aligned} g(z) &= \frac{\delta^2}{1+\delta^2}\left(1+\frac{1}{1+\delta\gamma}\frac{1}{z}\right)g_1(z) \\ &= \frac{\delta^2}{1+\delta^2}\left(1+\frac{1}{1+\delta\gamma}\frac{1}{z}\right)\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}}z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\gamma^2 z+\frac{\delta^2}{z}\right)} \\ &= \frac{\delta^3}{1+\delta^2}\frac{e^{\delta\gamma}}{\sqrt{2\pi}}\left(1+\frac{1}{1+\delta\gamma}\frac{1}{z}\right)z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\gamma^2 z+\frac{\delta^2}{z}\right)} \end{aligned} \tag{6.62}$$

Therefore

$$\begin{aligned} l_2 &= \sum_{i=1}^{n}\log g(z_i) \\ &= \sum_{i=1}^{n}\Big\{3\log\delta+\delta\gamma-\log(1+\delta^2)-\frac{1}{2}\log(2\pi)-\frac{3}{2}\log z_i+ \\ &\quad \log\left(1+\frac{1}{1+\delta\gamma}\frac{1}{z_i}\right)-\frac{\gamma^2}{2}z_i-\frac{\delta^2}{2}\frac{1}{z_i}\Big\} \\ &= 3n\log\delta+n\delta\gamma-n\log(1+\delta^2)-\frac{n}{2}\log(2\pi)-\frac{3}{2}\sum_{i=1}^{n}\log z_i+ \\ &\quad \sum_{i=1}^{n}\log\left(1+\frac{1}{1+\delta\gamma}\frac{1}{z_i}\right)-\frac{\gamma^2}{2}\sum_{i=1}^{n}z_i-\frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} \end{aligned} \tag{6.63}$$

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

$$\begin{aligned}
\therefore \frac{\partial}{\partial\gamma} l_2 &= 0 \Longrightarrow \gamma\sum_{i=1}^{n} z_i = n\delta - \frac{\delta}{(1+\delta\gamma)}\sum_{i=1}^{n} \frac{1}{1+(1+\delta\gamma)z_i} \\
\therefore \hat{\gamma} &= \frac{\delta\left[n - \frac{1}{(1+\delta\gamma)}\sum_{i=1}^{n}\frac{1}{1+(1+\delta\gamma)z_i}\right]}{\sum_{i=1}^{n} z_i} \qquad (6.64)
\end{aligned}$$

Next

$$\begin{aligned}
\frac{\partial}{\partial\delta} l_2 &= \frac{3n}{\delta} + n\gamma - \frac{2n\delta}{1+\delta^2} + \sum_{i=1}^{n} \frac{-\frac{1}{z_i}\gamma(1+\delta\gamma)^{-2}}{1+\frac{1}{1+\delta\gamma}\frac{1}{z_i}} - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= \frac{3n}{\delta} + n\gamma - \frac{2n\delta}{1+\delta^2} - \frac{\gamma}{1+\delta\gamma}\sum_{i=1}^{n}\frac{1}{1+(1+\delta\gamma)z_i} - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= \frac{3n}{\delta} - \frac{2n\delta}{1+\delta^2} + \gamma\left[n - \frac{1}{1+\delta\gamma}\sum_{i=1}^{n}\frac{1}{1+(1+\delta\gamma)z_i}\right] - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= \frac{3n+n\delta^2}{\delta(1+\delta^2)} + \gamma\left[n - \frac{1}{1+\delta\gamma}\sum_{i=1}^{n}\frac{1}{1+(1+\delta\gamma)z_i}\right] - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
\therefore \frac{\partial}{\partial\delta} l_2 = 0 \Longrightarrow \delta &= \frac{\frac{3n+n\delta^2}{\delta(1+\delta^2)} + \gamma\left[n - \frac{1}{1+\delta\gamma}\sum_{i=1}^{n}\frac{1}{1+(1+\delta\gamma)z_i}\right]}{\sum_{i=1}^{n}\frac{1}{z_i}} \qquad (6.65)
\end{aligned}$$

E-Step

Values of random variables Z_i and $\frac{1}{Z_i}$ are not known. So we estimate them by considering posterior expectations

$$E(Z_i/X_i) \text{ and } E\left(\frac{1}{Z_i}/X_i\right)$$

Let

$$s_i = E(Z_i/X_i) \text{ and } w_i = E\left(\frac{1}{Z_i}/X_i\right)$$

Iterations

$$s_i^{(k)} = \frac{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)+A(x_i;\alpha,\delta,\gamma)}{(1+\delta^{(k)}\gamma^{(k)})\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)+B(x_i;\alpha,\delta,\gamma)} \quad (6.66)$$

$$w_i^{(k)} = \frac{(1+\delta^{(k)}\gamma^{(k)})\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)+C(x_i;\alpha,\delta,\gamma)}{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)+D(x_i;\alpha,\delta,\gamma)} \quad (6.67)$$

$$\hat{\gamma}^{(k+1)} = \frac{\delta^{(k)}\left[n-\frac{1}{(1+\delta^{(k)}\gamma^{(k)})}\sum_{i=1}^{n}\frac{1}{1+(1+\delta^{(k)}\gamma^{(k)})s_i^{(k)}}\right]}{\sum_{i=1}^{n}s_i^{(k)}} \quad (6.68)$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})w_i^{(k)}}{n-\frac{1}{n}\sum_{i=1}^{n}s_i^{(k)}\sum_{i=1}^{n}w_i^{(k)}} \quad (6.69)$$

where

$$\begin{aligned} A(x_i;\alpha,\delta,\gamma) &= (1+\delta^{(k)}\gamma^{(k)})\delta^2\phi^{(k)}(x_i)K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right) \\ B(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right) \\ C(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2K_3\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right) \\ D(x_i;\alpha,\delta,\gamma) &= (1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^2\phi^{(k)}(x_i)K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right) \end{aligned}$$

$$\hat{\mu}^{(k+1)} = \bar{x}-\hat{\beta}^{(k+1)}\sum_{i=1}^{n}\frac{s_i^{(k)}}{n} \quad (6.70)$$

$$\hat{\alpha}^{(k+1)} = [(\gamma^{(k+1)})^2+(\beta^{(k+1)})^2]^{\frac{1}{2}} \quad (6.71)$$

6.3.11 M-Step for the mixing distribution for Model 3

The mixing distribution follows equation (3.27) presented as

$$g_{14}(z) = \frac{\delta\gamma^3 e^{\delta\gamma}}{\sqrt{2\pi}(\gamma^3+\delta)}\left(1+\frac{z^2}{1+\delta\gamma}\right)z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\gamma^2 z+\frac{\delta^2}{z}\right)} \quad (6.72)$$

with loglikelihood function

$$\begin{aligned} l_2 &= n\log\delta+3n\log\gamma+n\delta\gamma-\frac{n}{2}\log(2\pi)-n\log(1+\delta\gamma)-n\log(\gamma^3+\delta)+ \\ & \sum_{i=1}^{n}\log(1+\delta\gamma+z_i^2)-\frac{3}{2}\sum_{i=1}^{n}\log z_i-\frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i}-\frac{\gamma^2}{2}\sum_{i=1}^{n}z_i \end{aligned} \quad (6.73)$$

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

$$\left(\frac{n\delta^2}{(1+\delta\gamma)}-\sum_{i=1}^{n} z_i\right)\gamma^2+\sum_{i=1}^{n}\frac{\delta\gamma}{1+\delta\gamma+z_i^2}+\frac{3n\delta}{\gamma^3+\delta}=0 \tag{6.74}$$

$$\left(\frac{n\gamma^2}{1+\delta\gamma}-\sum_{i=1}^{n}\frac{1}{z_i}\right)\delta^2+\sum_{i=1}^{n}\frac{\delta\gamma}{1+\delta\gamma+z_i^2}+\frac{n\gamma^3}{\gamma^3+\delta}=0 \tag{6.75}$$

Both equations are quadratic in γ and δ respectively.

6.3.12 E-Step

Posterior Expectation

$$\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z\big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty \big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{2-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{K_0(\alpha\delta\sqrt{\phi(x)})+\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^2\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)}}{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-1}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^2\phi(x)}{\alpha^2}\frac{K_2(\alpha\delta\sqrt{\phi(x)})}{(1+\delta\gamma)}}{\big[\frac{\alpha}{\delta\sqrt{\phi(x)}}+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha\delta(1+\delta\gamma)\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^3(\phi(x))^{\frac{3}{2}}}{\alpha}K_2(\alpha\delta\sqrt{\phi(x)})}{[\alpha^2(1+\delta\gamma)+\delta^2\phi(x)]K_1(\alpha\delta\sqrt{\phi(x)})}
\end{aligned} \tag{6.76}$$

Similarly,

$$
\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}\big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\int_0^\infty \big(1+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{-2-1}+\frac{z^{0-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{1}{2}\left(\alpha^2 z+\frac{\delta^2\phi(x)}{z}\right)}\,dz} \\
&= \frac{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-2}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{\big[\big(\frac{\delta\sqrt{\phi(x)}}{\alpha}\big)^{-1}+\big(\frac{\delta\sqrt{\phi(x)}}{\alpha}\big)\frac{1}{1+\delta\gamma}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\frac{\alpha^2}{\delta^2\phi(x)}K_2(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{\big[\frac{\alpha}{\delta\sqrt{\phi(x)}}+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^2 K_2(\alpha\delta\sqrt{\phi(x)})+\frac{\delta^2\phi(x)}{1+\delta\gamma}K_0(\alpha\delta\sqrt{\phi(x)})}{\big[\alpha\delta\sqrt{\phi(x)}+\frac{\delta^3\left(\phi(x)\right)^{\frac{3}{2}}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \\
&= \frac{\alpha^3(1+\delta\gamma)K_2(\alpha\delta\sqrt{\phi(x)})+\alpha\delta^2\phi(x)K_0(\alpha\delta\sqrt{\phi(x)})}{\big[\alpha^2\delta(1+\delta\gamma)\sqrt{\phi(x)}+\delta^3\left(\phi(x)\right)^{\frac{3}{2}}\big]K_1(\alpha\delta\sqrt{\phi(x)})} \qquad (6.77)
\end{aligned}
$$

Similarly

$$
E(z^2/X=x)=\frac{(1+\delta\gamma)\alpha^2\delta^2\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})+\delta^4(\phi(x))^2K_3(\alpha\delta\sqrt{\phi(x)})}{[(1+\delta\gamma)\alpha^4+\alpha^2\delta^2\phi(x)]K_1(\alpha\delta\sqrt{\phi(x)})} \qquad (6.78)
$$

The posterior expectations for the $k-th$ iteration are:

$$
s_i^{(k)}=\frac{\alpha^{(k)}\delta^{(k)}(1+\delta^{(k)}\gamma^{(k)})\sqrt{\phi^{(k)}(x)}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+A(x_i;\alpha,\delta,\gamma)}{[(\alpha^{(k)})^2(1+\delta^{(k)}\gamma^{(k)})+(\delta^{(k)})^2\phi^{(k)}(x)]K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})} \qquad (6.79)
$$

$$
w_i^{(k)}=\frac{(\alpha^{(k)})^3(1+\delta^{(k)}\gamma^{(k)})K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+C(x_i;\alpha,\delta,\gamma)}{\big[(\alpha^{(k)})^2\delta^{(k)}(1+\delta^{(k)}\gamma^{(k)})\sqrt{\phi^{(k)}(x)}+(\delta^{(k)})^3\left(\phi^{(k)}(x)\right)^{\frac{3}{2}}\big]K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})} \qquad (6.80)
$$

$$
v_i^{(k)}=\frac{(1+\delta^{(k)}\gamma^{(k)})(\alpha^{(k)})^2(\delta^{(k)})^2\phi^{(k)}(x_i)K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+E(x_i;\alpha,\delta,\gamma)}{[1+\delta^{(k)}\gamma^{(k)}(\alpha^{(k)})^4+(\alpha^{(k)})^2(\delta^{(k)})^2\phi^{(k)}(x_i)]K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})} \qquad (6.81)
$$

where

$$
\begin{aligned}
A(x_i;\alpha,\delta,\gamma) &= \frac{(\delta^{(k)})^3(\phi^{(k)}(x))^{\frac{3}{2}}}{\alpha^{(k)}}K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) \\
C(x_i;\alpha,\delta,\gamma) &= \alpha^{(k)}(\delta^{(k)})^2\phi^{(k)}(x)K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) \\
E(x_i;\alpha,\delta,\gamma) &= (\delta)^4(\phi^{(k)}(x_i))^2K_3(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})
\end{aligned}
$$

Now, define the iterative scheme as follows:
let

$$a_1^{(k+1)} = \left(\frac{n(\delta^{(k)})^2}{(1+\delta^{(k)}\gamma^{(k)})} - \sum_{i=1}^{n} s_i^{(k)} \right)$$
$$b_1^{(k+1)} = \sum_{i=1}^{n} \frac{\delta^{(k)}}{1+\delta^{(k)}\gamma^{(k)}+v_i^{(k)}}$$
$$c_1^{(k+1)} = \frac{3n\delta^{(k)}}{(\gamma^{(k)})^3+\delta^{(k)}}$$

let

$$t^{(k+1)} = \frac{-b_1^{(k+1)} - \sqrt{(b_1^{(k+1)})^2 - 4a_1^{(k+1)}c_1^{(k+1)}}}{2a_1^{(k+1)}} \tag{6.82}$$

using the square root transformation, we have

$$\gamma^{k+1} = \sqrt{t^{(k+1)}} \tag{6.83}$$

Similarly, define

$$a_2^{(k+1)} = \left(\frac{n(\gamma^{(k+1)})^2}{1+\delta^{(k)}\gamma^{(k+1)}} - \sum_{i=1}^{n} w_i^{(k)} \right) \tag{6.84}$$
$$b_2^{(k+1)} = \sum_{i=1}^{n} \frac{\delta^{(k)}\gamma^{(k+1)}}{1+\delta^{(k)}\gamma^{(k+1)}+v_i^{(k)}} \tag{6.85}$$
$$c_2^{(k+1)} = \frac{n(\gamma^{(k+1)})^3}{(\gamma^{(k+1)})^3+\delta^{(k)}} \tag{6.86}$$

let

$$s^{(k+1)} = \frac{-b_2^{(k+1)} - \sqrt{(b_2^{(k+1)})^2 - 4a_2^{(k+1)}c_2^{(k+1)}}}{2a_2^{(k+1)}} \tag{6.87}$$

using the square root transformation, we have

$$\delta^{k+1} = \sqrt{s^{(k+1)}} \tag{6.88}$$

and the $k-th$ iteration for the loglikelihood is given by

$$\begin{aligned} l^{(k)} &= 3n\log\gamma^{(k)} + n\delta^{(k)}\gamma^{(k)} + \beta^{(k)}\sum_{i=1}^{n}(x_i-\mu^{(k)}) - n\log\alpha^{(k)}\pi - n\log(1+\delta^{(k)}\gamma^{(k)}) - \\ &\quad n\log((\gamma^{(k)})^3+\delta^{(k)}) - \frac{1}{2}\sum_{i=1}^{n}\log\phi^{(k)}(x_i) + \sum_{i=1}^{n}\log\left[(\alpha^{(k)})^2(1+\delta^{(k)}\gamma^{(k)}) + \right. \\ &\quad \left.(\delta^{(k)})^2\phi^{(k)}(x_i)\right] + \sum_{i=1}^{n}\log K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)_i}\right) \end{aligned} \tag{6.89}$$

6.3.13 M-Step of the mixing distribution for Model 4

From formula (3.33)

$$\begin{aligned} g(z) &= \frac{\gamma\delta^2}{\delta^3+\gamma}\big(z+\frac{1}{1+\delta\gamma}z^{-1}\big)g_1(z) \\ &= \frac{\gamma\delta^2}{\delta^3+\gamma}\big(z+\frac{1}{1+\delta\gamma}z^{-1}\big)\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}}z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)} \\ &= \frac{\gamma\delta^3 e^{\delta\gamma}}{\sqrt{2\pi(\delta^3+\gamma)}}\big(z+\frac{1}{1+\delta\gamma}z^{-1}\big)z^{-\frac{3}{2}}e^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)} \end{aligned} \tag{6.90}$$

Therefore

$$\begin{aligned} l_2 &= \sum_{i=1}^{n}\log g(z_i) \\ &= \sum_{i=1}^{n}\{\log\gamma+3\log\delta+\delta\gamma-\frac{1}{2}\log(2\pi)-\log(\delta^3+\gamma)-\log(1+\delta\gamma)-\frac{3}{2}\log(z_i)- \\ &\quad \frac{\delta^2}{2}\frac{1}{z_i}-\frac{\gamma^2}{2}z_i+\log((1+\delta\gamma)z_i+\frac{1}{z_i})\} \\ &= n\log\gamma+3n\log\delta+n\delta\gamma-\frac{n}{2}\log(2\pi)-n\log(\delta^3+\gamma)-n\log(1+\delta\gamma)-\frac{3}{2}\sum_{i=1}^{n}\log(z_i)- \\ &\quad \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i}-\frac{\gamma^2}{2}\sum_{i=1}^{n}z_i+\sum_{i=1}^{n}\log((1+\delta\gamma)z_i+\frac{1}{z_i}) \end{aligned} \tag{6.91}$$

Differentiating w.r.t γ we obtain

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

$\frac{\partial}{\partial\gamma}=0$ implies that

$$\frac{n\delta^3}{\gamma(\delta^3+\gamma)}+\frac{n\gamma\delta^2}{1+\delta\gamma}-\gamma\sum_{i=1}^{n}z_i+\sum_{i=1}^{n}\frac{\delta z_i^2}{1+(1+\delta\gamma)z_i^2}=0 \tag{6.92}$$

Similarly

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

$\frac{\partial}{\partial\delta} = 0$ implies that

$$\frac{3n\gamma}{\delta(\delta^3+\gamma)} + \frac{n\delta\gamma^2}{1+\delta\gamma} - \delta\sum_{i=1}^{n}\frac{1}{z_i} + \sum_{i=1}^{n}\frac{\gamma z_i^2}{1+(1+\delta\gamma)z_i^2} = 0 \tag{6.93}$$

6.3.14 E-Step

Values of random variables Z_i, $\frac{1}{Z_i}$ and Z_i^2 are not known. So we estimate them by considering posterior expectations

$$E(Z_i/x_i), E\left(\tfrac{1}{Z_i}/x_i\right) \text{ and } E(Z_i^2/x_i)$$

Let

$$s_i = E(Z_i/x_i), w_i = E\left(\tfrac{1}{Z_i}/x_i\right) \text{ and } v_i = E(Z_i/x_i)$$

The $k-th$ iterations are as follows

$$s_i^{(k)} = \frac{(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^3(\sqrt{\phi^{(k)}(x_i)})^3 K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + A(x_i;\alpha,\delta,\gamma)}{\alpha^{(k)}(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^2\phi^{(k)}(x_i)K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + B(x_i;\alpha,\delta,\gamma)} \tag{6.94}$$

$$w_i^{(k)} = \frac{\alpha^{(k)}(\delta^{(k)})^2(1+\delta^{(k)}\gamma^{(k)})\phi^{(k)}(x_i)K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + C(x_i;\alpha,\delta,\gamma)}{(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})^3 K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + D(x_i;\alpha,\delta,\gamma)} \tag{6.95}$$

$$v_i^{(k)} = \frac{(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^2\phi^{(k)}(x_i)K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + E(x_i;\alpha,\delta,\gamma)}{(\alpha^{(k)})^2(1+\delta^{(k)}\gamma^{(k)})K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) + F(x_i;\alpha,\delta,\gamma)} \tag{6.96}$$

where

$$\begin{aligned}
A(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) \\
B(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^3 K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) \\
C(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^3 K_3(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) \\
D(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) \\
E(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2 K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}) \\
F(x_i;\alpha,\delta,\gamma) &= (\alpha^{(k)})^2 K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})
\end{aligned}$$

For the log-likelihood, the $k-th$ iteration is given as

$$\begin{aligned} l^{(k)} &= n\log(\delta^{(k)}\gamma^{(k)}) + n\delta^{(k)}\gamma^{(k)} + \beta^{(k)}\sum_{i=1}^{n} x_i - n\beta^{(k)}\mu^{(k)} - n\log((1+\delta^{(k)}\gamma^{(k)})\pi((\delta^{(k)})^3+\gamma^{(k)})) - \\ &\sum_{i=1}^{n}\log\phi^{(k)}(x_i) + \sum_{i=1}^{n}\log\left\{(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^2\phi^{(k)}(x)K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right) + \right. \\ &\left.(\alpha^{(k)})^2K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)\right\} \end{aligned} \tag{6.97}$$

6.3.15 Iterative Scheme

From equations (6.92) and (6.93), we obtain the following iterative scheme

$$\gamma^{(k+1)} = \frac{\frac{n(\delta^{(k)})^3}{\gamma^{(k)}((\delta^{(k)})^3+\gamma^{(k)})} + \frac{n\gamma^{(k)}(\delta^{(k)})^2}{1+\delta^{(k)}\gamma^{(k)}} + \sum_{i=1}^{(k)} \frac{\delta^{(k)}z_i^2}{1+(1+\delta^{(k)}\gamma^{(k)})z_i^2}}{\sum_{i=1}^{n} s_i^{(k)}} \tag{6.98}$$

$$\delta^{(k+1)} = \frac{\frac{3n\gamma^{(k+1)}}{\delta^{(k)}((\delta^{(k)})^3+\gamma^{(k+1)})} + \frac{n\delta^{(k)}(\gamma^{(k+1)})^2}{1+\delta^{(k)}\gamma^{(k+1)}} + \sum_{i=1}^{n} \frac{\gamma^{(k+1)}z_i^2}{1+(1+\delta^{(k)}\gamma^{(k+1)})z_i^2}}{\sum_{i=1}^{n} w_i^k} \tag{6.99}$$

with

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n} x_i w_i^{(k)} - \bar{x}\sum_{i=1}^{n} w_i^{(k)}}{n - \bar{s}^{(k)}\sum_{i=1}^{n} w_i^{(k)}} \tag{6.100}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \hat{\beta}^{(k+1)}\bar{s}^{(k)} \tag{6.101}$$

$$\hat{\alpha}^{(k+1)} = \left[(\hat{\gamma}^{(k+1)})^2 + (\hat{\beta}^{(k+1)})^2\right]^{\frac{1}{2}} \tag{6.102}$$

6.3.16 M-Step of the mixing distribution for Model 5

Consider formula (3.39) presented as

$$\begin{aligned} g_{24}(z) &= \frac{\gamma^3}{\delta(\gamma^2+1)}\left[z + \frac{z^2}{1+\delta\gamma}\right]g_1(z) \\ &= \frac{\gamma^3 e^{\delta\gamma}}{\sqrt{2\pi}(\gamma^2+1)}\left[1 + \frac{z}{1+\delta\gamma}\right]z^{-\frac{1}{2}}e^{-\frac{1}{2}\left(\frac{\delta^2}{z}+\gamma^2 z\right)} \end{aligned} \tag{6.103}$$

with the loglikelihood

$$\begin{aligned} l_2 &= 3n\log\gamma + n\delta\gamma - \frac{n}{2}\log(2\pi) - n\log(\gamma^2+1) - n\log(1+\delta\gamma) - \frac{1}{2}\sum_{i=1}^{n}\log z_i + \\ &\sum_{i=1}^{n}[\log(1+\delta\gamma+z_i)] - \frac{\gamma^2}{2}\sum_{i=1}^{n} z_i - \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} \end{aligned} \tag{6.104}$$

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

$$\frac{3n+2n\delta\gamma}{\gamma(1+\delta)} - \gamma\left(\frac{2n}{\gamma^2+1}+\sum_{i=1}^{n} z_i\right) + \delta\left(n+\sum_{i=1}^{n}\frac{1}{1+\delta\gamma+z_i}\right) = 0 \qquad (6.105)$$

$$\frac{n\delta\gamma^2}{1+\delta\gamma} + \gamma\sum_{i=1}^{n}\frac{1}{1+\delta\gamma+z_i} - \delta\sum_{i=1}^{n}\frac{1}{z_i} = 0 \qquad (6.106)$$

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

$$\begin{aligned}
E\big(Z/X=x\big) &= \frac{\int_0^\infty z\big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\int_0^\infty \big(z^{1-1}+\frac{z^{2-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z^{0-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{\delta\sqrt{\phi(x)}}{\alpha}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^2\frac{K_2\left(\alpha\delta\sqrt{\phi(x)}\right)}{1+\delta\gamma}}{K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{\delta\sqrt{\phi(x)}(1+\delta\gamma)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\frac{\delta^2\phi(x)(1+\delta\gamma)}{\alpha}K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{\alpha(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)} \\
&= \frac{\alpha\delta\sqrt{\phi(x)}(1+\delta\gamma)K_1\big(\alpha\delta\sqrt{\phi(x)}\big)+\delta\phi(x)(1+\delta\gamma)K_2\big(\alpha\delta\sqrt{\phi(x)}\big)}{\alpha^2(1+\delta\gamma)K_0\big(\alpha\delta\sqrt{\phi(x)}\big)+\alpha\delta\sqrt{\phi(x)}K_1\big(\alpha\delta\sqrt{\phi(x)}\big)}
\end{aligned} \qquad (6.107)$$

Similarly,

$$\begin{aligned}
E\big(\frac{1}{Z}/X=x\big) &= \frac{\int_0^\infty z^{-1}\big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\int_0^\infty \big(z+\frac{z^2}{1+\delta\gamma}\big)z^{-2}e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\frac{1}{2}\int_0^\infty \big(z^{-1-1}+\frac{z^{0-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz}{\frac{1}{2}\int_0^\infty \big(z^{0-1}+\frac{z^{1-1}}{1+\delta\gamma}\big)e^{-\frac{\alpha^2}{2}\left(z+\frac{\delta^2\phi(x)}{\alpha^2 z}\right)}\,dz} \\
&= \frac{\big[\frac{\delta\sqrt{\phi(x)}}{\alpha}\big]^{-1}K_1(\alpha\delta\sqrt{\phi(x)})+\frac{K_0(\alpha\delta\sqrt{\phi(x)})}{1+\delta\gamma}}{K_0(\alpha\delta\sqrt{\phi(x)})+\frac{\delta\sqrt{\phi(x)}}{\alpha(1+\delta\gamma)}\big]K_1(\alpha\delta\sqrt{\phi(x)})}
\end{aligned}$$

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

$$= \frac{\alpha^2(1+\delta\gamma)K_1(\alpha\delta\sqrt{\phi(x)})+\alpha\delta\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})}{\alpha\delta\sqrt{\phi(x)}K_0(\alpha\delta\sqrt{\phi(x)})+\delta^2\phi(x)K_1(\alpha\delta\sqrt{\phi(x)})} \quad \textbf{(6.108)}$$

Similarly

$$E(Z^2/X=x) = \frac{\alpha(1+\delta\gamma)\delta^2\phi(x)K_2(\alpha\delta\sqrt{\phi(x)})+(\delta\sqrt{\phi(x)})^3K_3(\alpha\delta\sqrt{\phi(x)})}{\alpha^3(1+\delta\gamma)K_0(\alpha\delta\sqrt{\phi(x)})+\alpha^2\delta\sqrt{\phi(x)}K_1(\alpha\delta\sqrt{\phi(x)})} \quad \textbf{(6.109)}$$

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

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

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

Therefore the k-th iterations are:

$$s_i^{(k+1)} = \frac{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}(1+\delta^{(k)}\gamma^{(k)})K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+A(x_i;\alpha,\delta,\gamma)}{(\alpha^{(k)})^2(1+\delta^{(k)}\gamma^{(k)})K_0\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)+B(x_i;\alpha,\delta,\gamma)} \quad \textbf{(6.110)}$$

$$w_i^{(k+1)} = \frac{(\alpha^{(k)})^2(1+\delta^{(k)}\gamma^{(k)})K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+C(x_i;\alpha,\delta,\gamma)}{\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+D(x_i;\alpha,\delta,\gamma)} \quad \textbf{(6.111)}$$

$$v_i^{(k+1)} = \frac{\alpha^{(k)}(1+\delta^{(k)}\gamma^{(k)})(\delta^{(k)})^2\phi^{(k)}(x)K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+E(x_i;\alpha,\delta,\gamma)}{(\alpha^{(k)})^3(1+\delta^{(k)}\gamma^{(k)})K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})+F(x_i;\alpha,\delta,\gamma)} \quad \textbf{(6.112)}$$

(6.113)

where

$$A(x_i;\alpha,\delta,\gamma) = \delta^{(k)}\phi^{(k)}(x)(1+\delta^{(k)}\gamma^{(k)})K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)$$

$$B(x_i;\alpha,\delta,\gamma) = \alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}\right)$$

$$C(x_i;\alpha,\delta,\gamma) = \alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})$$

$$D(x_i;\alpha,\delta,\gamma) = (\delta^{(k)})^2\phi^{(k)}(x)K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})$$

$$E(x_i;\alpha,\delta,\gamma) = (\delta^{(k)}\sqrt{\phi^{(k)}(x)})^3K_3(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})$$

$$F(x_i;\alpha,\delta,\gamma) = (\alpha^{(k)})^2\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})$$

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

$$\hat{\gamma}^{(k+1)} = \frac{\frac{3+2\delta^{(k)}\gamma^{(k)}}{\gamma^{(k)}(1+\delta^{(k)}\gamma^{(k)})} + \delta^{(k)}\left[1+\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+\delta^{(k)}\gamma^{(k)}+s_i^{(k)}}\right]}{\left[\frac{2}{(\gamma^{(k)})^2+1}+\bar{s}^{(k)}\right]} \tag{6.114}$$

$$\hat{\delta}^{(k+1)} = \frac{\frac{n\delta^{(k)}(\gamma^{(k+1)})^2}{1+\delta^{(k)}\gamma^{(k+1)}} + \gamma^{(k+1)}\sum_{i=1}^{n}\frac{1}{1+\delta^{(k)}\gamma^{(k+1)}+s_i}}{n\bar{w}^{(k)}} \tag{6.115}$$

$$\hat{\beta}^{(k+1)} = \frac{\sum_{i=1}^{n} x_i w_i^{(k+1)} - \bar{x}\sum_{i=1}^{n} w_i^{(k+1)}}{n - \bar{s}^{(k+1)}\sum_{i=1}^{n} w_i^{(k+1)}} \tag{6.116}$$

$$\hat{\mu}^{(k+1)} = \bar{x} - \hat{\beta}^{(k+1)}\bar{s}^{(k+1)} \tag{6.117}$$

$$\hat{\alpha}^{(k+1)} = \sqrt{(\hat{\beta}^{(k+1)})^2 + (\hat{\gamma}^{(k+1)})^2} \tag{6.118}$$

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

$$\begin{aligned} l^{(k+1)} = {} & 3n\log\gamma^{(k+1)} + n\delta^{(k+1)}\gamma^{(k+1)} + \beta^{(k+1)}\sum_{i=1}^{n} x_i - n\beta^{(k+1)}\mu^{(k+1)} - \\ & n\log(\alpha^{(k+1)}\pi((\gamma^{(k+1)})^2+1)(1+\delta^{(k+1)}\gamma^{(k+1)})) + \sum_{i=1}^{n}\log\big[\alpha^{(k+1)}\times \\ & (1+\delta^{(k+1)}\gamma^{(k+1)})K_0\big(\alpha^{(k+1)}\delta^{(k+1)}\sqrt{\phi^{(k+1)}(x)}\big) + \delta^{(k+1)}\times \\ & \sqrt{\phi^{(k+1)}(x)}K_1\big(\alpha^{(k+1)}\delta^{(k+1)}\sqrt{\phi^{(k+1)}(x)}\big)\big] \end{aligned} \tag{6.119}$$

6.3.17 M-Step for the mixing distribution of Model 6

We now consider the the mixing distribution given by formula (3.45) expressed as

$$\begin{aligned} g_{34}(z) &= \frac{\delta^2\gamma^3}{(\gamma^3+\delta^3)(1+\delta\gamma)}(z^{-1}+z^2)g_1 z \\ &= \frac{\delta^2\gamma^3}{(\gamma^3+\delta^3)(1+\delta\gamma)}(z^{-1}+z^2)\frac{\delta e^{\delta\gamma}}{\sqrt{2\pi}z^{-\frac{3}{2}}e^{-\frac{1}{2}(\gamma^2 z+\frac{\delta^2}{z})}} \\ &= \frac{\delta^3\gamma^3 e^{\delta\gamma}}{\sqrt{2\pi}(\gamma^3+\delta^3)(1+\delta\gamma)}(z^{-1}+z^2)z^{-\frac{3}{2}}e^{-\frac{1}{2}(\gamma^2 z+\frac{\delta^2}{z})} \end{aligned} \tag{6.120}$$

Therefore

$$\begin{aligned} l_2 &= \sum_{i=1}^{n} \log g(z_i) \\ &= \sum_{i=1}^{n} \Big\{ 3\log\delta + 3\log\gamma + \delta\gamma - \frac{1}{2}\log(2\pi) - \log((\gamma^3+\delta^3)(1+\delta\gamma)) + \log(\frac{1}{z_i} + z_i^2) \\ &\quad -\frac{3}{2}\log(z_i) - \frac{\gamma^2}{2} z_i - \frac{\delta^2}{2}\frac{1}{z_i} \Big\} \\ &= 3n\log\delta + 3n\log\gamma + n\delta\gamma - \frac{n}{2}\log(2\pi) - n\log((\gamma^3+\delta^3)(1+\delta\gamma)) + \sum_{i=1}^{n}\log(\frac{1}{z_i} + z_i^2) \\ &\quad -\frac{3}{2}\sum_{i=1}^{n}\log(z_i) - \frac{\gamma^2}{2}\sum_{i=1}^{n} z_i - \frac{\delta^2}{2}\sum_{i=1}^{n}\frac{1}{z_i} \end{aligned} \tag{6.121}$$

Hence

$$\begin{aligned} \frac{\partial}{\partial\gamma} l_2 &= \frac{3n}{\gamma} + n\delta - \frac{3n\gamma^2}{\delta^3+\gamma^3} - \frac{n\delta}{1+\delta\gamma} - \gamma\sum_{i=1}^{n} z_i \\ &= 3n(\frac{1}{\gamma} - \frac{\gamma^2}{\gamma^3+\delta^3}) + n\delta(1 - \frac{1}{1+\delta\gamma}) - \gamma\sum_{i=1}^{n} z_i \end{aligned} \tag{6.122}$$

$$\begin{aligned} \frac{\partial}{\partial\gamma} l_2 &= 3n\frac{(\gamma^3+\delta^3-\gamma^3)}{\gamma(\gamma^3+\delta^3)} + n\delta(\frac{1+\delta\gamma-1}{1+\delta\gamma}) - \gamma\sum_{i=1}^{n} z_i \\ &= \frac{3n\delta^3}{\gamma(\gamma^3+\delta^3)} + \frac{n\delta^2\gamma}{1+\delta\gamma} - \gamma\sum_{i=1}^{n} z_i \\ &= \frac{3n\delta^3}{\gamma(\gamma^3+\delta^3)} - \gamma(\sum_{i=1}^{n} z_i - \frac{n\delta^2}{1+\delta\gamma}) \\ \frac{\partial}{\partial\gamma} l_2 &= 0 \Longrightarrow \gamma(\sum_{i=1}^{n} z_i - \frac{n\delta^2}{1+\delta\gamma}) = \frac{3n\delta^3}{\gamma(\gamma^3+\delta^3)} \\ \therefore \gamma &= \frac{\frac{3n\delta^3}{\gamma(\gamma^3+\delta^3)}}{(\sum_{i=1}^{n} z_i - \frac{n\delta^2}{1+\delta\gamma})} \\ &= \frac{3n\delta^3}{\gamma(\gamma^3+\delta^3)(\sum_{i=1}^{n} z_i - \frac{n\delta^2}{1+\delta\gamma})} \\ &= \frac{3\delta^3}{\gamma(\gamma^3+\delta^3)(\bar{z} - \frac{\delta^2}{1+\delta\gamma})} \\ \therefore \hat{\gamma} &= \frac{3\delta^3(1+\delta\gamma)}{\gamma(\gamma^3+\delta^3)((1+\delta\gamma)\bar{z} - \delta^2)} \end{aligned} \tag{6.123}$$

Similarly

$$
\begin{aligned}
\frac{\partial}{\partial\delta} l_2 &= \frac{3n}{\delta} + n\gamma - \frac{3n\delta^2}{\gamma^3+\delta^3} - \frac{n\gamma}{(1+\delta\gamma)} - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= 3n\left(\frac{1}{\delta} - \frac{\delta^2}{\gamma^3+\delta^3}\right) + n\gamma\left(1 - \frac{1}{1+\delta\gamma}\right) - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= 3n\frac{(\gamma^3+\delta^3-\delta^3)}{\delta(\gamma^3+\delta^3)} + n\gamma\left(\frac{1+\delta\gamma-1}{1+\delta\gamma}\right) - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= \frac{3n\gamma^3}{\delta(\gamma^3+\delta^3)} + \frac{n\gamma^2\delta}{1+\delta\gamma} - \delta\sum_{i=1}^{n}\frac{1}{z_i} \\
&= \frac{3n\gamma^3}{\delta(\gamma^3+\delta^3)} - \delta\left[\sum_{i=1}^{n}\frac{1}{z_i} - \frac{n\gamma^2}{1+\delta\gamma}\right] \\
\frac{\partial}{\partial\delta} l_2 &= 0 \Longrightarrow \delta\left[\sum_{i=1}^{n}\frac{1}{z_i} - \frac{n\gamma^2}{1+\delta\gamma}\right] = \frac{3n\gamma^3}{\delta(\gamma^3+\delta^3)} \\
\therefore \hat{\delta} &= \frac{3n\gamma^3}{\delta(\gamma^3+\delta^3)}\left(\frac{1}{\sum_{i=1}^{n}\frac{1}{z_i} - \frac{n\gamma^2}{1+\delta\gamma}}\right) \\
&= \frac{3\gamma^3(1+\delta\gamma)}{\delta(\gamma^3+\delta^3)\left[\frac{1+\delta\gamma}{n}\sum_{i=1}^{n}\frac{1}{z_i} - \gamma^2\right]} \\
&= \frac{3\gamma^3(1+\delta\gamma)}{\delta(\gamma^3+\delta^3)\left[(1+\delta\gamma)\bar{w} - \gamma^2\right]} \qquad \textbf{(6.124)}
\end{aligned}
$$

6.3.18 E-Step and Iterations

Values of random variables Z_i and $\frac{1}{Z_i}$ are not known. So we estimate them by considering posterior expectations

$$E(Z_i/X_i) \text{ and } E\left(\tfrac{1}{Z_i}/X_i\right)$$

Let

$$s_i = E(Z_i/X_i) \text{ and } w_i = E\left(\tfrac{1}{Z_i}/X_i\right)$$

and

$$\bar{s} = \sum_{0}^{n}\frac{s_i}{n}$$

$$\bar{w} = \sum_{0}^{n}\frac{w_i}{n}$$

Therefore the k-th iterations are

$$s_i^{(k)} = \frac{(\alpha^{(k)})^3\delta^{(k)}\sqrt{\phi^{(k)}(x)}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) + ((\delta^{(k)})^2\phi^{(k)}(x))^2K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})}{(\alpha^{(k)})^4K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) + \alpha^{(k)}(\delta^{(k)})^3(\phi^{(k)}(x))^{\frac{3}{2}}K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})} \quad (6.125)$$

$$w_i^{(k)} = \frac{(\alpha^{(k)})^4K_3(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) + \alpha^{(k)}(\delta^{(k)}\sqrt{\phi^{(k)}(x)})^3K_0(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})}{\delta^{(k)}(\alpha^{(k)})^3\sqrt{\phi^{(k)}(x)}K_2(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)}) + ((\delta^{(k)})^2\phi^{(k)}(x))^2K_1(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x)})} \quad (6.126)$$

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

$$\hat{\gamma}^{(k+1)} = \frac{3(\delta^{(k)})^3(1+\delta^{(k)}\gamma^{(k)})}{\gamma^{(k)}((\gamma^{(k)})^3+(\delta^{(k)})^3)((1+\delta^{(k)}\gamma^{(k)})\bar{s}^{(k)}-(\delta^{(k)})^2)} \quad (6.127)$$

$$\hat{\delta}^{(k+1)} = \frac{3(\gamma^{(k)})^3(1+\delta^{(k)}\gamma^{(k)})}{\delta^{(k)}((\gamma^{(k)})^3+(\delta^{(k)})^3)\left[(1+\delta^{(k)}\gamma^{(k)})\bar{w}^{(k)}-(\gamma^{(k)})^2\right]} \quad (6.128)$$

The $k-th$ iteration of the log-likelihood function becomes

$$\begin{aligned} l^{(k)} = {} & 3n\log\gamma^{(k)} + n\log\delta^{(k)} + n\delta^{(k)}\gamma^{(k)} + \beta^{(k)}\sum_{i=1}^{n}(x_i-\mu^{(k)}) - n\log(\alpha^{(k)}\pi((\delta^{(k)})^3 + \\ & (\gamma^{(k)})^3)(1+\delta^{(k)}\gamma^{(k)})) - \sum_{i=1}^{n}\log\phi^{(k)}(x_i) + \\ & \sum_{i=1}^{n}\log\left\{(\alpha^{(k)})^3K_2\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right) + (\delta^{(k)}\sqrt{\phi^{(k)}(x_i)})^3K_1\left(\alpha^{(k)}\delta^{(k)}\sqrt{\phi^{(k)}(x_i)}\right)\right\} \end{aligned} \quad (6.129)$$

Remark

The method of moment (MoM) estimates for these models are difficult to obtain directly. We suggest to use the moment estimates for *NIG* as initial values as presented by Karlis (2002) formulation. This is motivated by the fact that all models are related to the *NIG*. That is, they are Normal Weighted Inverse Gaussian distributions. Thus, the *NIG* MoM estimates are within the admissible range.

Moment estimators for *NIG* distribution obtained by Karlis (2002) are as follows:

$$\hat{\mu} = \bar{x} - \hat{\beta}\frac{\hat{\delta}}{\hat{\gamma}}$$

$$\hat{\delta} = \frac{s^2\hat{\gamma}^3}{\hat{\beta}^2+\hat{\gamma}^2}$$

$$\hat{\beta} = \frac{\bar{\gamma}_1 s\hat{\gamma}^2}{3}$$

$$\hat{\gamma} = \frac{3}{s\sqrt{(3\bar{\gamma}_2 - 5\bar{\gamma}_1{}^2)}}$$

where $\bar{x}$ is the sample mean, s^2 is the sample variance, while $\bar{\gamma}_1 = \mu_3/\mu_2^{3/2}$ and $\bar{\gamma}_2 = \mu_4/\mu_2^2 - 3$, with $\mu_k = n^{-1}\sum_{i=1}^{n}(x_i-\bar{x})^k$, i.e. the sample skewness and kurtosis respectively.

therefore it can be easily shown that:

$$\hat{\gamma} = \frac{3}{s\sqrt{(3\bar{\gamma}_2 - 5\bar{\gamma}_1{}^2)}}$$

The other parameters can be obtained sequentially by substituting the value of $\hat{\gamma}$. In this chapter we have constructed the iterative scheme designs for the normal mixtures to be used in Maximum Likelihood parameter estimation via the Expectation maximization (EM) Algorithm applied in the next chapter.

7 EM ALGORITHM ESTIMATION USING NWIG DISTRIBUTIONS TO FINANCIAL DATA

In this chapter we estimate the maximum likelihood (ML) parameters for the NWIG distributions via the EM-algorithm using iterative schemes developed from the previous chapter.

7.1 Introduction

In this section we consider three data sets 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 Figure 7.1, 7.2 and 7.3 for these data sets shows that the data is negatively skewed and exhibiting heavy tails.The Q-Q plot show that the normal distribution is not a good fit for the data especially at the tails. This is typical for the other data sets.

Table 7.1 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 tails than the normal distribution. We observe skewness for the data sets which indicates that the two tails of the returns behave slightly differently.

Table 7.1: Summary Statistics for the data sets 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

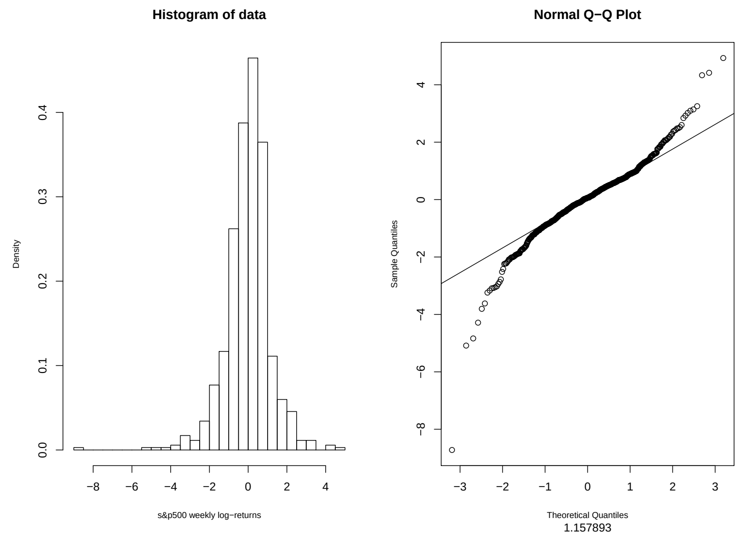

Figure 7.1. Histogram and Q-Q plot for s&p500 weekly log-returns

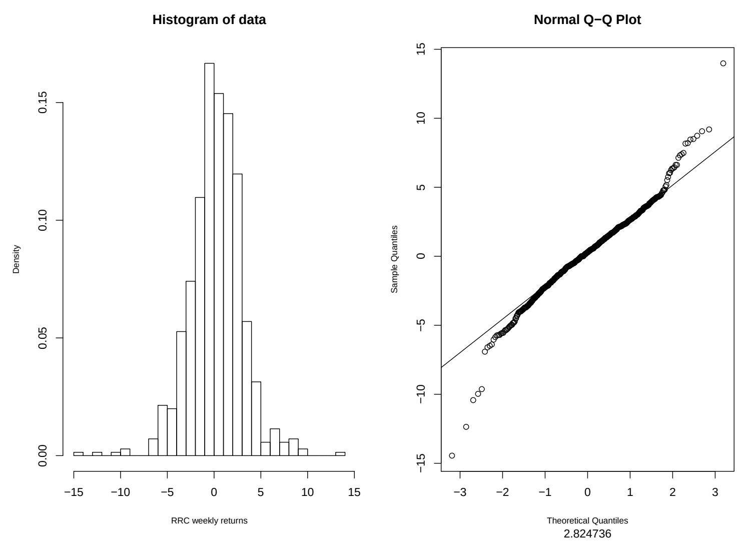

Figure 7.2. Histogram and Q-Q plot for RRC weekly log-returns

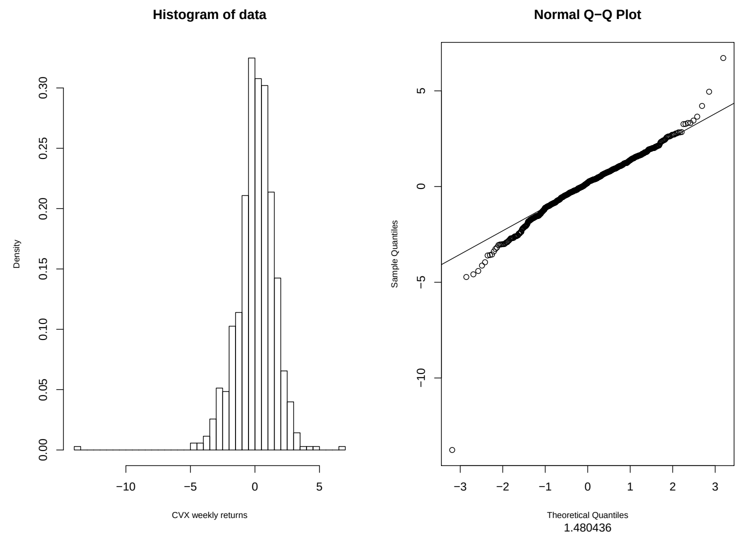

Figure 7.3. Histogram and Q-Q plot for CVX weekly log-returns

We use Karlis (2002) formulation to obtain the initial values for the EM algorithm. For this dataset, the values obtained are

Table 7.2: Method of Moment estimates of NIG 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

The stopping criterion is when

$$\frac{l^{(k)} - l^{(k-1)}}{l^{(k)}} < tol \tag{7.1}$$

where tol is the tolerance level.

7.2 Parameter Estimation for the Normal Inverse Gaussian

MoM and EM parameter estimates at different tolerance levels are presented in table below. The loglikehood value, number of iterations are also given. The monotonicity property of the EM algorithm can be seen from table.

Table 7.3: Maximum Likelihood estimate of NIG for RRC

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.4203574	0.4214410	0.4215579
$\hat{\beta}$	-0.02456226	-0.03578018	-0.03585363	-0.03586155
$\hat{\delta}$	2.950864	3.277072	3.284293	3.285072
$\hat{\mu}$	0.4284473	0.5132705	0.5137393	0.5137899
Loglikelihood	-1696.347	-1695.549	-1695.855	-1695.888
No. iteration		79	129	228
AIC	3400.694	3399.098	3399.71	3399.776

Figure 7.4 and 7.5 below shows the fit for NIG to the RRC data set.

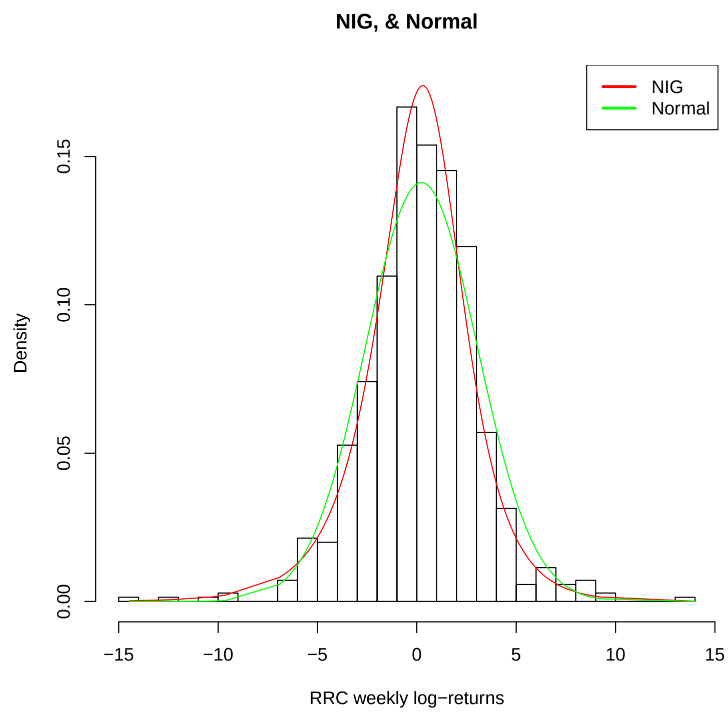

Figure 7.4. Fitting NIG to RRC weekly log-returns

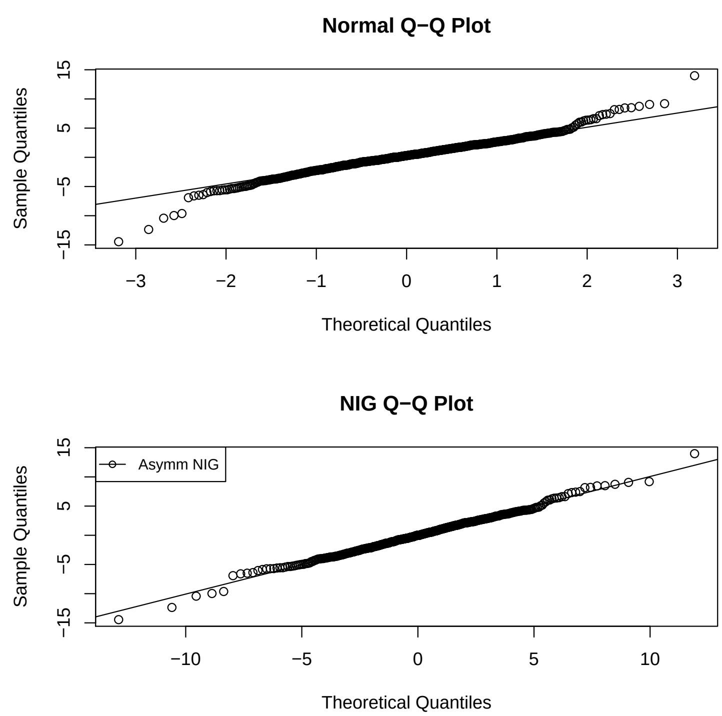

Figure 7.5. Q-Q plot for NIG

7.3 Parameter Estimation for the Normal Reciprocal Inverse Gaussian

The Table below shows the maximum likelihood parameter estimates for NRIG using the RRC data set.

Table 7.4: Maximum likelihood parameter estimates for NRIG using RRC

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.5479919	0.5490824	0.5491998
$\hat{\beta}$	-0.02456226	-0.03894518	-0.03903887	-0.03904892
$\hat{\delta}$	2.950864	2.413854	2.423925	2.425010
$\hat{\mu}$	0.4284473	0.5356479	0.5362332	0.5362960
Loglikelihood	-1777.218	-1696.095	-1696.414	-1696.449
No. iteration		101	154	259
AIC		3400.189	3400.829	3400.898

Figure 7.6 and 7.7 show the fit and Q-Q plot to the RRC data set.

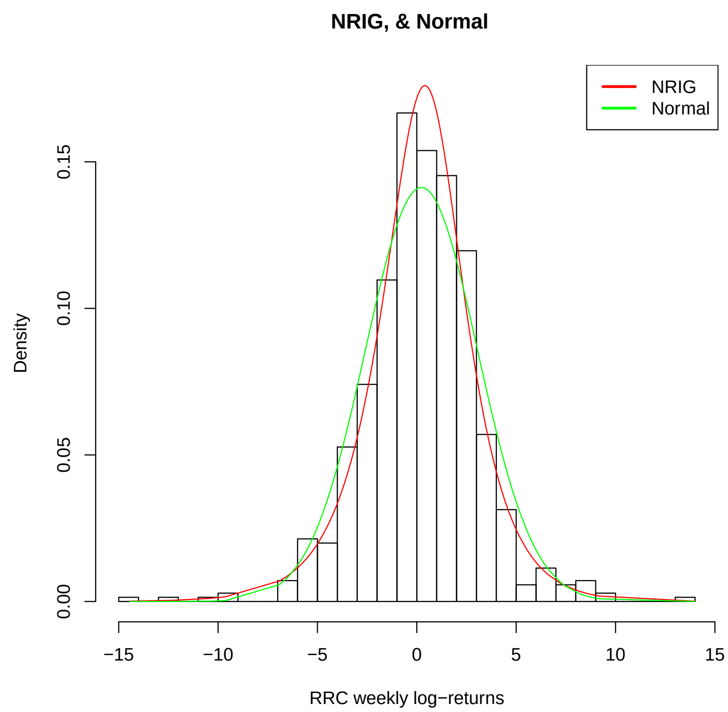

Figure 7.6. Fitting NRIG to RRC weekly log-returns

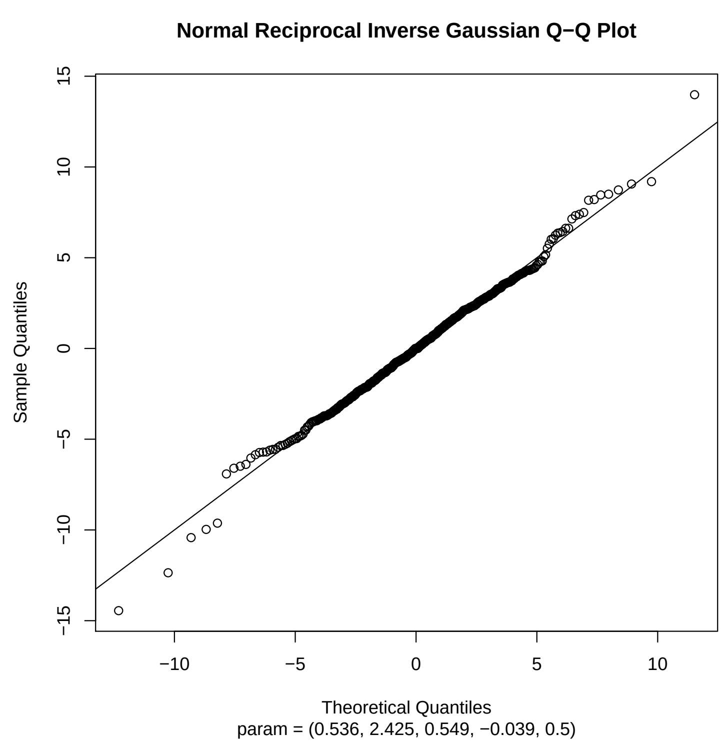

Figure 7.7. Q-Q plot for NRIG

7.4 Parameter Estimation of the GHD when the index parameter is $-\frac{3}{2}$

Table 7.5: Parameter estimates at different tolerance level using RRC

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.277191	0.2777899	0.2778586
$\hat{\beta}$	-0.02456226	-0.0323062	-0.03234023	-0.03234413
$\hat{\delta}$	2.950864	4.095578	4.098373	4.098694
$\hat{\mu}$	0.4284473	0.4880231	0.4882531	0.4882795
Loglikelihood		-1695.205	-1695.459	-1695.488
No. iteration		104	147	235
AIC		3398.41	3398.918	3398.976

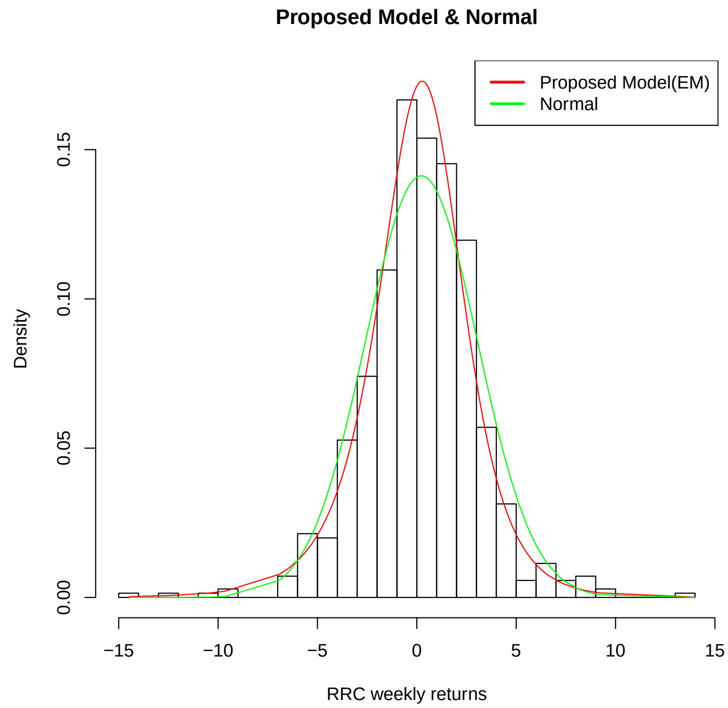

Figure 7.8. Fitting GHD($\lambda = -\frac{3}{2}$) to RRC weekly returns

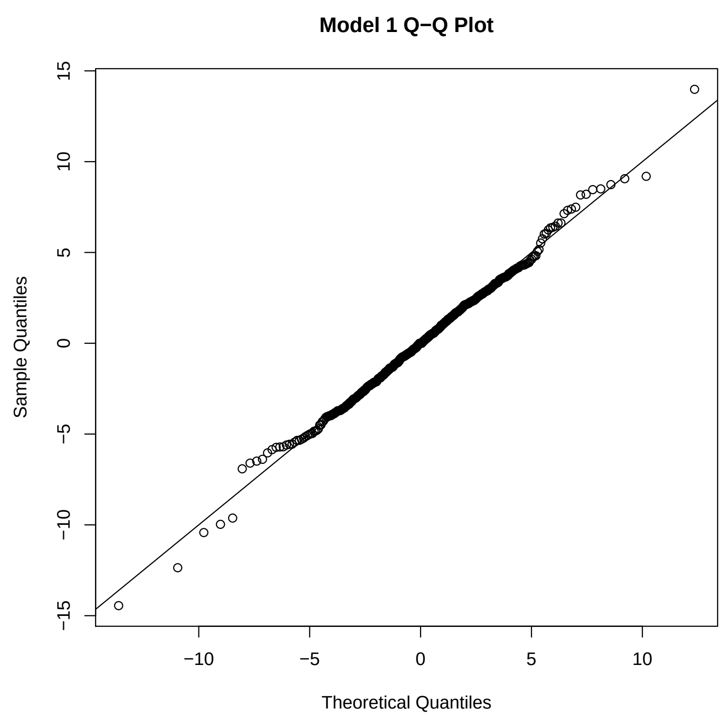

Figure 7.9. Q-Q plot of GHD($\lambda = -\frac{3}{2}$) for RRC weekly log-returns

7.5 Parameter Estimation of the GHD when the index parameter is $\frac{3}{2}$

Table 7.6: Maximum likelihood estimates of GHD($\lambda = \frac{3}{2}$) for RRC weekly log-returns using iteration scheme 1

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.6718806	0.6724	0.6724594
$\hat{\beta}$	-0.02456226	-0.04170169	-0.04177134	-0.04177928
$\hat{\delta}$	2.950864	1.409413	1.417213	1.418104
$\hat{\mu}$	0.4284473	0.5541088	0.5545578	0.554609
Loglikelihood		-1696.843	-1697.156	-1697.192
No. iteration		87	139	245
AIC		3401.686	3402.312	3402.384

Table 7.7: Maximum likelihood estimates of GHD($\lambda = \frac{3}{2}$) for RRC weekly log-returns using iteration scheme 2.

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.6716031	0.672372	0.6724591
$\hat{\beta}$	-0.02456226	-0.04168266	-0.04176951	-0.04177926
$\hat{\delta}$	2.950864	1.403473	1.416612	1.418098
$\hat{\mu}$	0.4284473	0.5539407	0.5545414	0.5546088
Loglikelihood		-1696.738	-1697.146	-1697.192
No. iteration		125	191	325
AIC		3401.476	3402.292	3402.384

Table 7.8: Maximum likelihood estimates of GHD($\lambda = \frac{3}{2}$) for RRC weekly log-returns using iteration scheme 3.

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.6716851	0.6723834	0.6724592
$\hat{\beta}$	-0.02456226	-0.04169365	-0.04177095	-0.04177927
$\hat{\delta}$	2.950864	1.404576	1.416776	1.418099
$\hat{\mu}$	0.4284473	0.5540067	0.5545504	0.5546089
Loglikelihood		-1696.754	-1697.149	-1697.192
No. iteration		122	186	313
AIC		3401.508	3402.298	3402.384

Table 7.9: Maximum likelihood estimates of GHD($\lambda = \frac{3}{2}$) for RRC weekly log-returns using iteration scheme 4.

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.6735225	0.6725625	0.672461
$\hat{\beta}$	-0.02456226	-0.04192363	-0.04179333	-0.04177949
$\hat{\delta}$	2.950864	1.43292	1.419544	1.418127
$\hat{\mu}$	0.4284473	0.5555022	0.554696	0.5546104
Loglikelihood		-1697.662	-1697.238	-1697.193
No. iteration		195	263	397
AIC		3403.324	3402.476	3402.386

Table 7.10: Maximum likelihood estimates of GHD($\lambda = \frac{3}{2}$) for RRC weekly log-returns using iteration scheme 5.

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.964887	0.6725535	0.6724609
$\hat{\beta}$	-0.02456226	-0.06772162	-0.04179207	-0.04177948
$\hat{\delta}$	2.950864	4.799672	1.419435	1.418126
$\hat{\mu}$	0.4284473	0.7313935	0.5546888	0.5546103
Loglikelihood		-1754.313	-1697.236	-1697.193
No. iteration		49	303	431
AIC		35166.626	3402.472	3402.386

Remark:
Parameter estimates can been obtained without updating. For example in scheme 2 we now have the following iteration.

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

and

$$\hat{\gamma}^{(k+1)} = \frac{3}{\hat{\gamma}^{(k)}} (\bar{s}^{(k)} - \frac{(\hat{\delta}^{(k)})^2}{1 + \hat{\delta}^{(k)} \hat{\gamma}^{(k)}})^{-1}$$

We should note that the estimate $\hat{\gamma}^{(k+1)}$ now depends on $\hat{\delta}^{(k)}$ and not $\hat{\delta}^{(k+1)}$.
Applying this to all the 5 schemes we obtain the following

Table 7.11: Iteration schemes at $(tol = 10^{-8})$ without updating of parameters.

Parameter	scheme 1	scheme 2	scheme 3	scheme 4	Scheme 5
$\hat{\alpha}$	0.672459	0.6724575	0.6724577	0.6724575	0.6724577
$\hat{\beta}$	-0.04177924	-0.04177905	-0.04177907	-0.04177905	-0.04177907
$\hat{\delta}$	1.418096	1.418077	1.418079	1.418077	1.418079
$\hat{\mu}$	0.5546087	0.5546076	0.5546077	0.5546076	0.5546077
Loglikelihood	-1697.192	-1697.191	-1697.191	-1697.191	-1697.191
No. iteration	355	432	413	432	413
AIC	3402.384	3402.382	3402.382	3402.382	

Figure 7.10 and 7.11 shows the fit and Q-Q plot to RRC data.

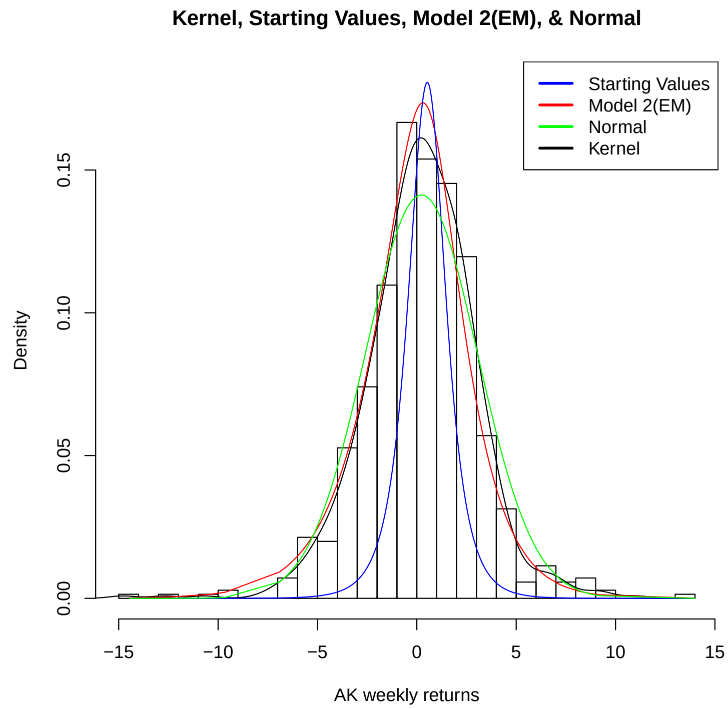

Figure 7.10. Fitting GHD($\lambda = \frac{3}{2}$) to RRC weekly log-returns

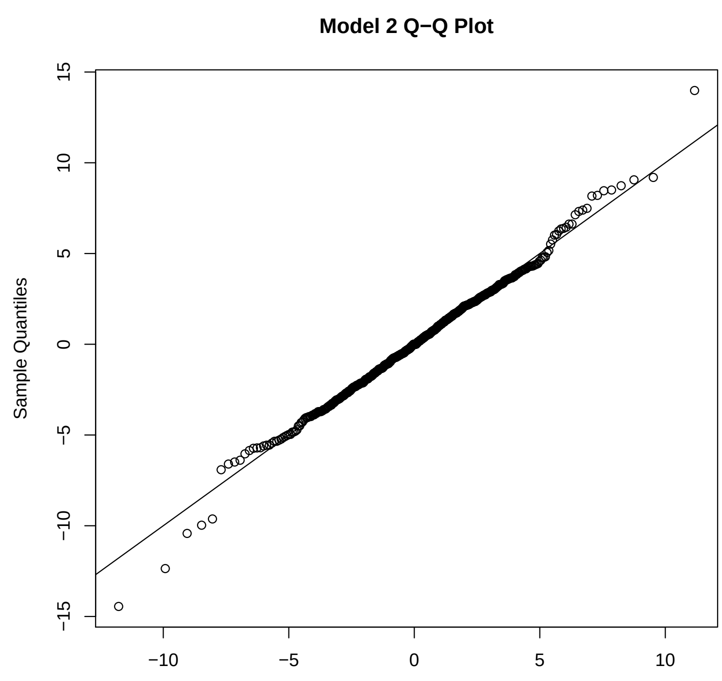

Figure 7.11. Q-Q plot of GHD($\lambda = \frac{3}{2}$) for RRC Weekly log-returns

We now show the fit of the Normal Mixture models with Finite inverse Gaussian mixing distributions.

7.6 Parameter Estimation of Model 1

Table 7.12: Maximum likelihood Parameter Estimates of Model 1 for RRC weekly log-returns based on Iterative scheme 1

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.4990739	0.5443183	0.5448038
$\hat{\beta}$	-0.02456226	-0.03450044	-0.03884595	-0.03887937
$\hat{\delta}$	2.950864	2.287431	2.68337	2.687818
$\hat{\mu}$	0.4284473	0.5067461	0.5349165	0.5351374
Loglikelihood		-1701.816	-1696.447	-1696.404
No. iteration		16	136	264
AIC		3411.632	3400.894	3400.808

Table 7.13: Maximum likelihood Parameter Estimates of Model 1 for RRC weekly log-returns based on iterative Scheme 2.

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.5570388	0.5569838	0.5448147
$\hat{\beta}$	-0.02456226	-0.03913063	-0.03942849	-0.03888028
$\hat{\delta}$	2.950864	2.796674	2.790291	2.687908
$\hat{\mu}$	0.4284473	0.536258	0.5380866	0.535143
Loglikelihood		-1695.589	-1695.597	-1696.403
No. iteration		20	22	239
AIC		3399.178	3399.194	3400.806

The maximum likelihood estimates for the NIG distribution using Karlis (2002) EM-algorithm formulation are

$$\hat{\alpha} = 0.4215579, \hat{\beta} = -0.03586155, \hat{\delta} = 3.285072, \hat{\mu} = 0.5137899.$$

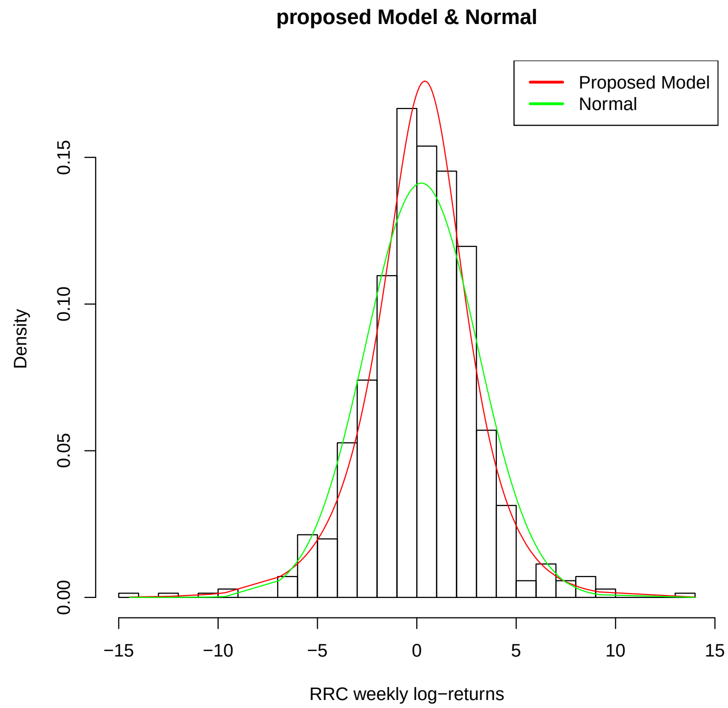

Figure 7.12. Fitting Model 1 to RRC weekly returns

Remark: The Proposed model can be presented as

$$f(x) = \frac{\gamma}{\gamma+\delta} \times NIG + \frac{\delta}{\gamma+\delta} \times NRIG \tag{7.2}$$

Note that the parameter estimates used here are those of the Proposed model. Figure 2 illustrates the Fit of the Proposed model model and its components.

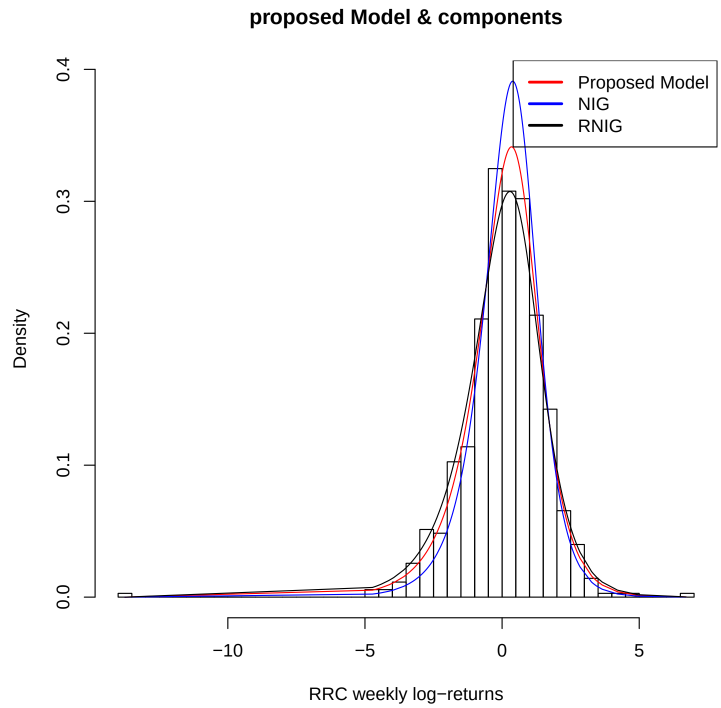

Figure 7.13. Combined plot for model 1 and its components

Using the estimates, $p = 0.1681738$ and $1-p = 8318262$. From Figure 2, NRIG fits the data well compared to NIG. The Proposed model puts sufficient weight on NRIG compared to NIG.

7.7 Parameter Estimation for Model 2

Table 7.14: Maximum likelihood estimates of Model 2 for RRC Weekly log-returns

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.5422119	0.5669625	0.5735228
$\hat{\beta}$	-0.02456226	-0.04351162	-0.04569187	-0.04626743
$\hat{\delta}$	2.950864	4.135342	4.279309	4.3181
$\hat{\mu}$	0.4284473	0.561554	0.575256	0.5789067
Loglikelihood		-1698.1	-1697.796	-1697.745
No. iteration		55	110	271
AIC		3404.2	3403.592	3403.49

Table 7.15:Maximum likelihood estimates of Model 2 for CVX Weekly log-returns

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.4190067	1.342608	1.382169	1.612872
$\hat{\beta}$	-0.1054991	-0.3551768	-0.3718762	-0.4751398
$\hat{\delta}$	0.8324058	2.491154	2.536825	2.805817
$\hat{\mu}$	0.3036691	0.7494867	0.7756322	0.9361947
Loglikelihood		-1226.209	-1226.186	-1226.962
No. iteration		48	55	490
AIC		2460.418	2460.372	2453.912

Table 7.16: Maximum likelihood estimates of Model 2 for s&p500 index Weekly log-returns

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-7})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.6556607	1.305925	1.644426	1.644773
$\hat{\beta}$	-0.1257455	-0.1979512	-0.2684969	-0.2685751
$\hat{\delta}$	0.8	1.633107	1.907611	1.90791
$\hat{\mu}$	0.3036691	0.2371093	0.308193	0.3082736
Loglikelihood		-1047.278	-1049.191	-1049.194
No. iteration		27	270	357
AIC		2102.556	2106.382	2106.388

The Figures below show that the proposed models is a good fit for the data sets.

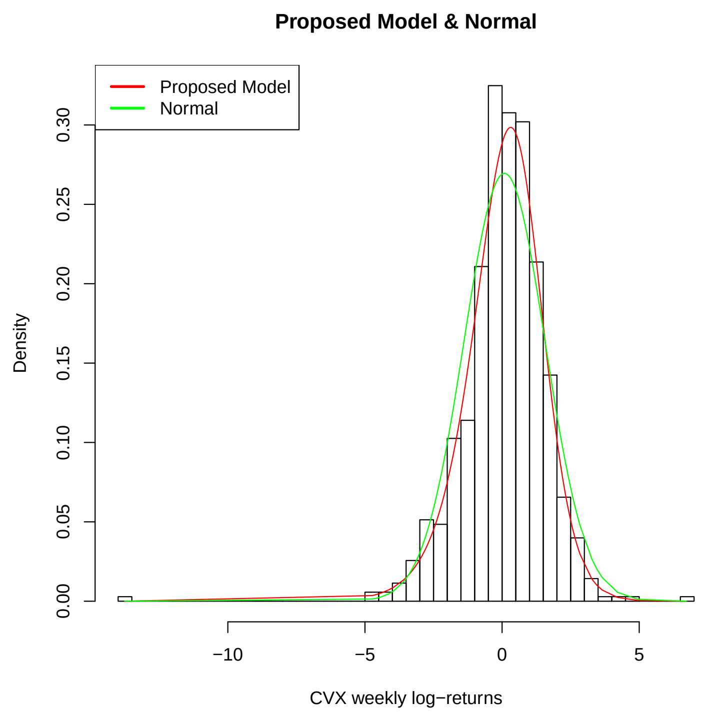

Figure 7.14. Fitting Model 2 to RRC log weekly returns

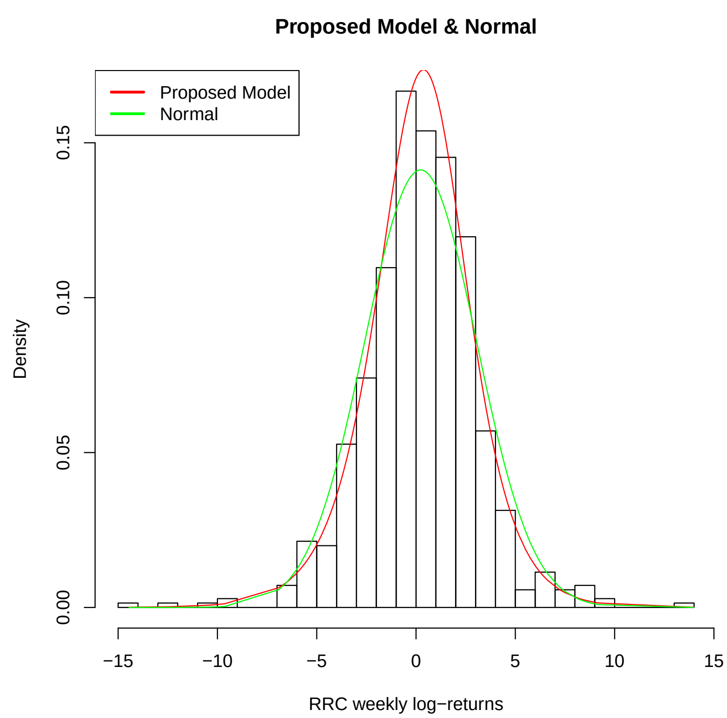

Figure 7.15. Fitting Model 2 to CVX log weekly returns

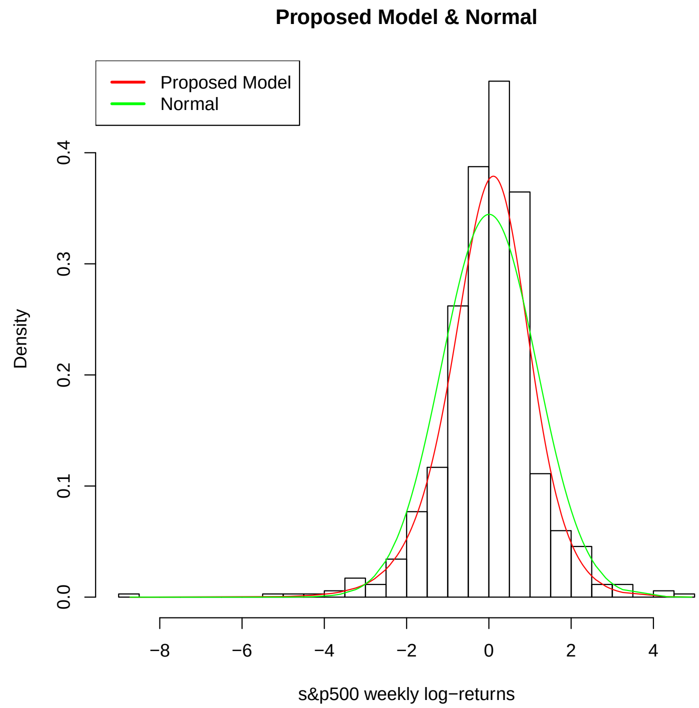

Figure 7.16. Fitting Model 2 to s&p500 index log weekly returns

Remark:

Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\delta^2}{1+\delta^2} \times NIG + \frac{1}{1+\delta^2} \times GHD(-\frac{3}{2}, \alpha, \delta, \beta, \mu) \tag{7.3}$$

Using the estimates we obtain the estimates of p for the data sets to be

Table 7.17: estimates of p for the data sets.

dataset	$\hat{p}$
RRC	0.94910
CVX	0.88729
s&p500	0.78449

7.8 Parameter Estimation for Model 3

Table 7.18: Maximum likelihood Parameter Estimates of Model 3 for $s\&p500$ index

Parameter	Starting Values	$EM(tol = 10^{-3})$	$EM(tol = 10^{-5})$	$EM(tol = 10^{-10})$
$\hat{\alpha}$	0.6556607	1.025172	1.025056	1.025061
$\hat{\beta}$	-0.1257455	-0.082968	-0.082969	-0.082988
$\hat{\delta}$	0.8310044	0.8571405	0.8570473	0.8570354
$\hat{\mu}$	0.1690855	0.1506045	0.150692	0.1507229
Loglikelihood		-1048.305	-1048.288	-1048.301
No. iteration		9	12	32
AIC		2104.61	2104.576	2104.602

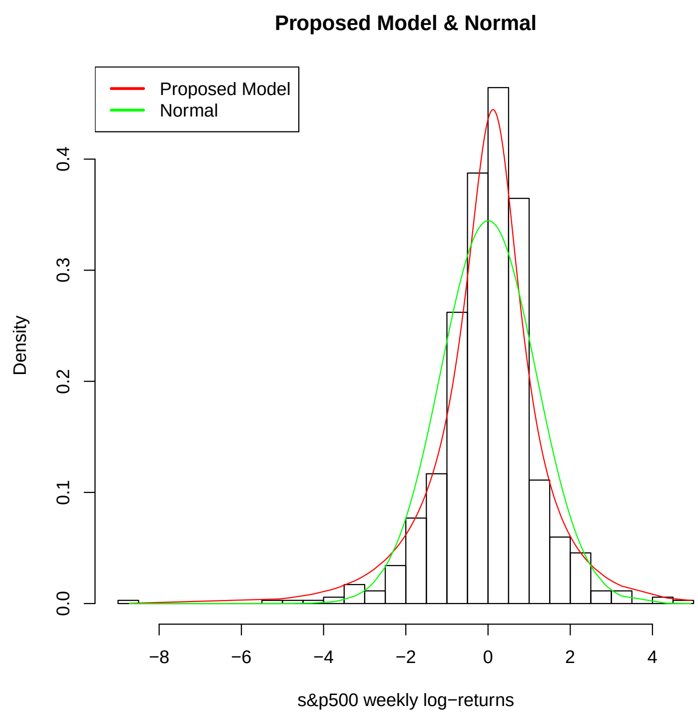

Figure 7.17. Fitting Model 3 to s&p500 index weekly log returns

The proposed model fits the data set well. Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\gamma^3}{\gamma^3+\delta} \times NIG + \frac{\delta}{\gamma^3+\delta} \times GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu) \qquad \textbf{(7.4)}$$

Using the parameter estimates $p = 0.242793$. Therefore, the finite mixture for these data sets is more weighted to the GHD when $\lambda = \frac{3}{2}$ than to the NIG.

7.9 Parameter Estimation for Model 4

Table 7.19: Maximum likelihood parameter estimates of Model 4 for RRC

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.5144623	0.5144017	0.5144511
$\hat{\beta}$	-0.02456226	-0.03578978	-0.0357382	-0.03571578
$\hat{\delta}$	2.950864	2.26434	2.264649	2.265279
$\hat{\mu}$	0.4284473	0.5176135	0.5172807	0.5171165
Loglikelihood		-1696.862	-1696.873	-1696.844
No. iteration		43	47	78
AIC		3401.724	3401.746	3401.688

Table 7.20: Maximum likelihood parameter estimates of Model 4 for CVX

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.4190067	1.167283	1.167218	1.167188
$\hat{\beta}$	-0.1054991	-0.2492112	-0.2491672	-0.2491203
$\hat{\delta}$	0.8324058	1.631138	1.631156	1.631209
$\hat{\mu}$	0.3036691	0.5692185	0.5691717	0.5691122
Loglikelihood		-1222.955	-1222.962	-1222.956
No. iteration		37	41	66
AIC		2453.91	2453.924	2453.912

Table 7.21: Maximum likelihood estimates of Model 5 for s&p500 index

Parameter	Starting Values	$EM(tol = 10^{-5})$	$EM(tol = 10^{-6})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.6556607	1.569046	1.568931	1.568897
$\hat{\beta}$	-0.1257455	-0.2078989	-0.20787	-0.2078466
$\hat{\delta}$	0.8	1.466964	1.466918	1.466935
$\hat{\mu}$	0.3036691	0.2425692	0.2425475	0.2425306
Loglikelihood		-1042.372	-1042.386	-1042.385
No. iteration		36	42	65
AIC		2092.744	2092.772	2092.77

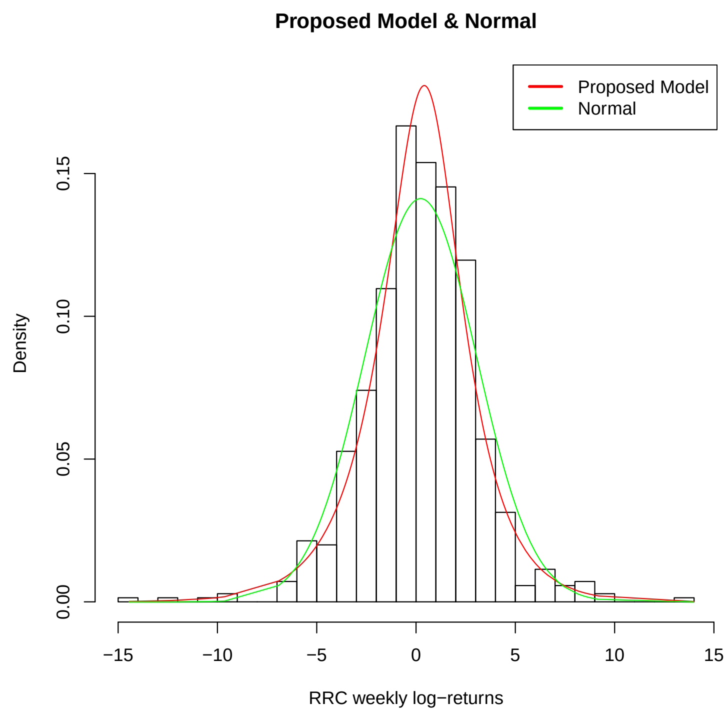

Figure 7.18. Fitting Model 4 to RRC log weekly returns

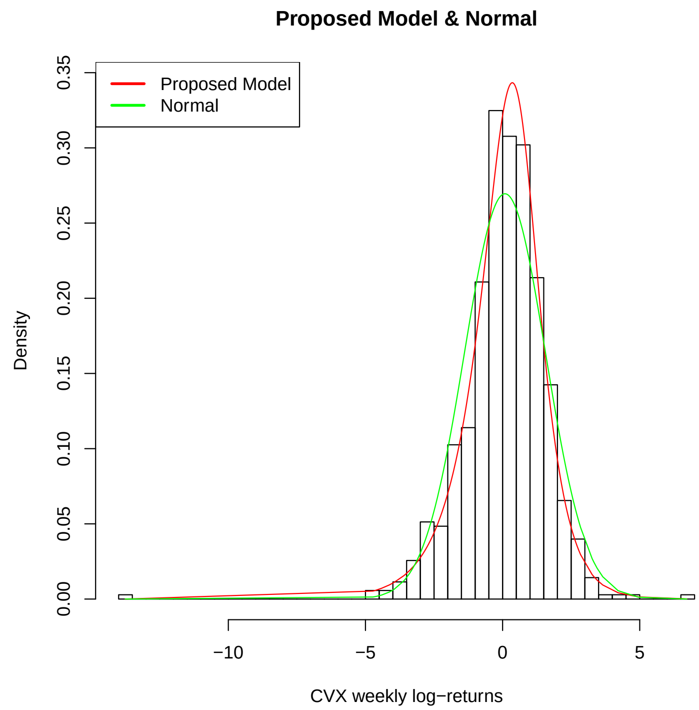

Figure 7.19. Fitting Model 4 to CVX log weekly returns

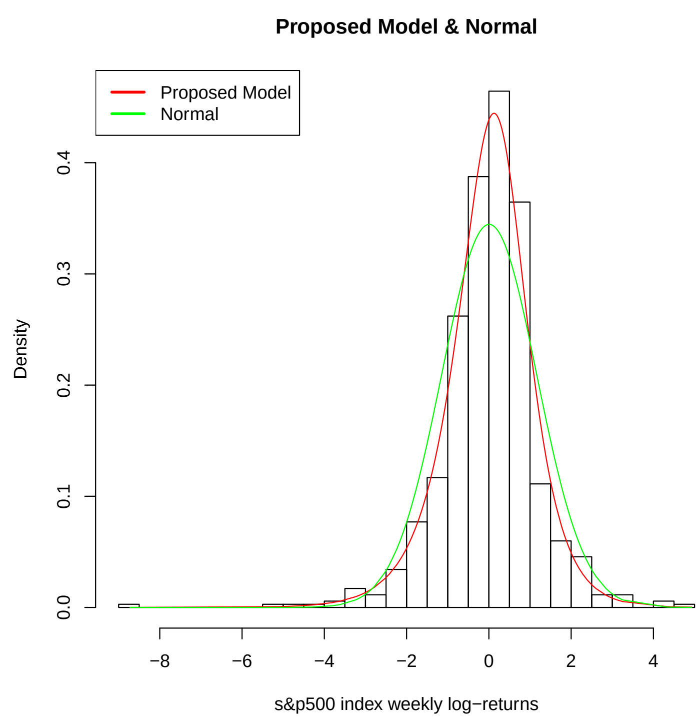

Figure 7.20. Fitting Model 4 to s&p500 index log weekly returns

Remark:

Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\delta^3}{\delta^3+\gamma} \times NRIG + \frac{\gamma}{\delta^3+\gamma} \times GHD(-\frac{3}{2}, \alpha, \delta, \beta, \mu) \tag{7.5}$$

Using the estimates we obtain the estimates of p for the data sets to be

Table 7.22: Estimates of p in Model 6 for the data sets.

dataset	$\hat{p}$
RRC	0.95772
CVX	0.79194
s&p500	0.66996

The finite mixture for these data sets is more weighted to the NRIG than the other special case of the GHD when $\lambda = -\frac{3}{2}$.

7.10 Parameter Estimate for Model 5

Table 7.23: Maximum likelihood parameter estimates of Model 5 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
Loglikelihood		-1697.415	-1697.81	-1697.814
No. iteration		32	50	67
AIC		3402.83	3403.62	3403.628

Table 7.24: Maximum likelihood parameter estimates of Model 5 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
Loglikelihood		-1224.464	-1223.515	-1223.504
No. iteration		30	77	126
AIC		2456.928	2455.030	2455.008

Table 7.25: Maximum likelihood parameter estimates of Model 5 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
Loglikelihood		-1061.985	-1061.673	-1063.643
No. iteration		31	43	510
AIC		2131.970	2131.346	2135.286

Figures below 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.

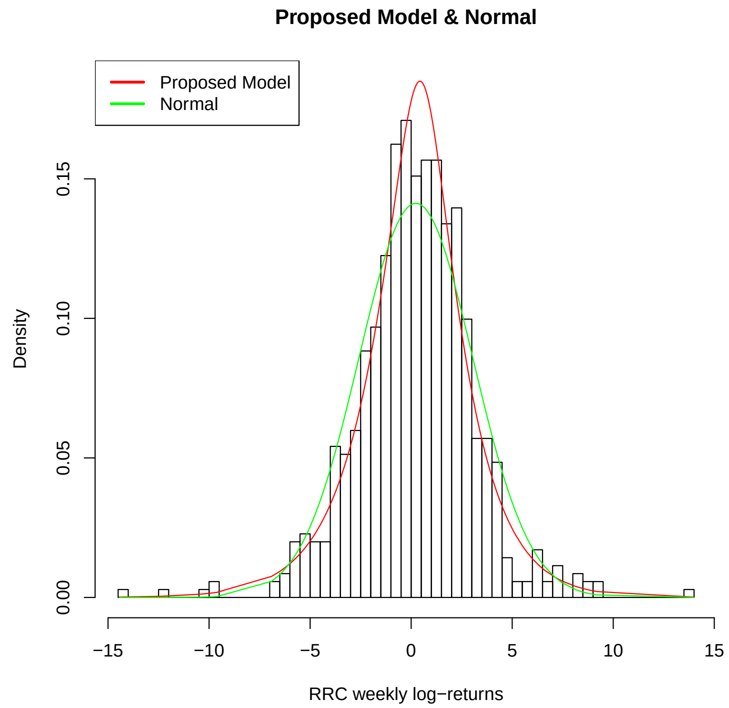

Figure 7.21. Fitting Model 5 to RRC weekly log-returns

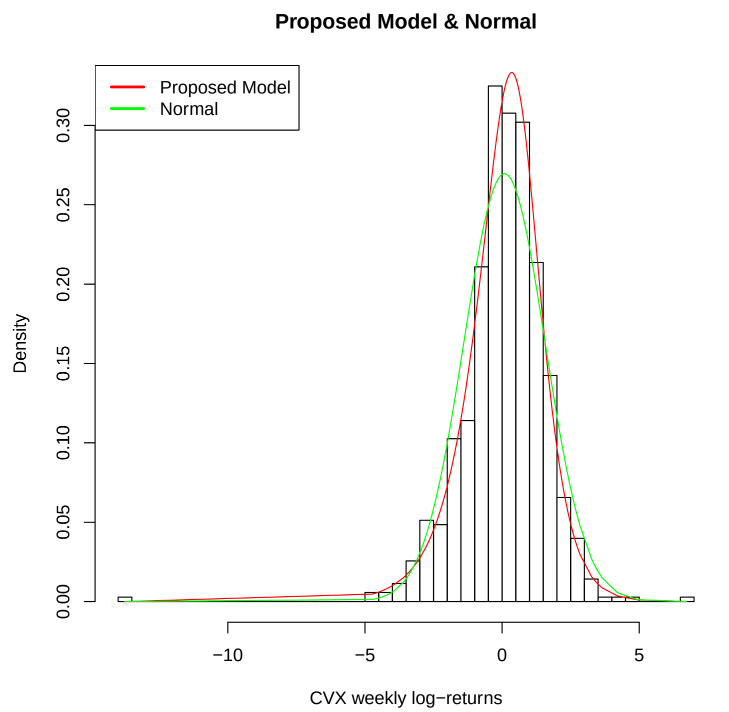

Figure 7.22. Fitting Model 5 to CVX weekly log-returns

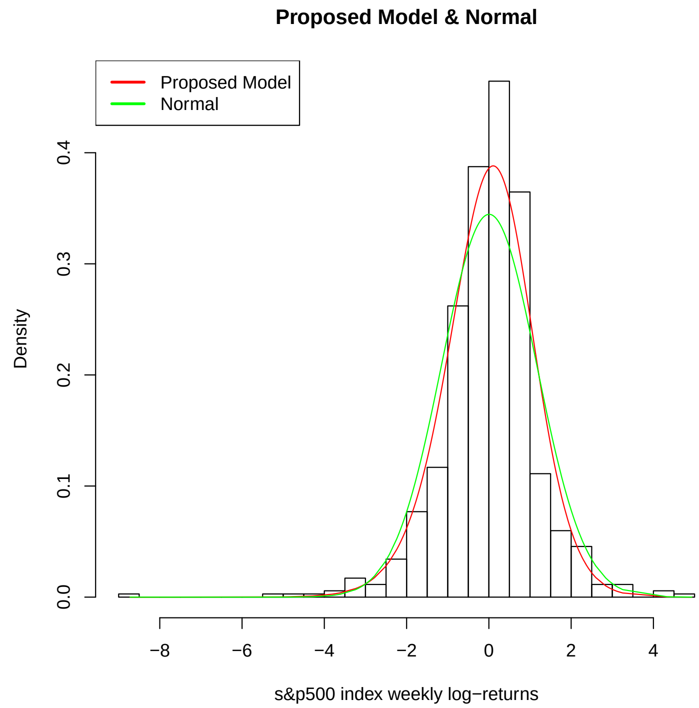

Figure 7.23. Fitting Model 5 to s&p500 index weekly log-returns

Remark:

Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\gamma^2}{\gamma^2+1} \times GHD(\frac{1}{2}, \alpha, \delta, \beta, \mu) + \frac{1}{\gamma^2+1} \times GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu) \tag{7.6}$$

Table 7.26: Estimates of p in Model 5 for the data sets.

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

The finite mixture therefore is flexible in determining between $GHD(\frac{1}{2}, \alpha, \delta, \beta, \mu)$ and $GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu)$ depending on the nature of the data.

7.11 Parameter Estimation for Model 6

For the purpose of comparison, we start by giving the maximum parameter estimates of NIG in the table below.

Table 7.27: Maximum likelihood parameter estimates of NIG for the data sets.

dataset	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$	loglikelihood	No. iterations
RRC	0.4215579	-0.03586155	3.285072	0.5137899	-1695.888	228
CVX	0.9265284	-0.2429664	1.7323	0.5578459	-1221.667	119
s&p500	0.7671588	-0.1299129	0.9661878	0.1727121	-1035.403	74

We now wish to obtain the maximum likelihood parameter estimates for the proposed model via the EM algorithm. Tables 4-6 illustrate monotonic convergence at different levels. The loglikelihood and AIC for each data sets are also provided.

Table 7.28: Maximum likelihood parameter estimates of Model 6 for RRC data set at different levels of tolerance.

Parameter	Starting Values	$EM(tol = 10^{-6})$	$EM(tol = 10^{-7})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.3722511	0.7895924	0.6931385	0.6931218
$\hat{\beta}$	-0.02456226	-0.05281212	-0.04211019	-0.04210727
$\hat{\delta}$	2.950864	2.971323	1.982277	1.982127
$\hat{\mu}$	0.4284473	0.6242585	0.5562482	0.5562297
Loglikelihood		-1677.033	-1683.709	-1683.713
No. iteration		50	280	334
AIC		3362.066	3359.418	3359.426

Table 7.29: Maximum likelihood parameter estimates of Model 6 for CVX data set at different levels of tolerance.

Parameter	Starting Values	$EM(tol = 10^{-6})$	$EM(tol = 10^{-7})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.4190067	1.426509	1.426399	1.426388
$\hat{\beta}$	-0.1054991	-0.2628444	-0.2627827	-0.2627765
$\hat{\delta}$	0.8324058	1.768069	1.767947	1.767935
$\hat{\mu}$	0.3036691	0.5909337	0.5908452	0.5908363
Loglikelihood		-1191.741	-1191.752	-1191.753
No. iteration		131	159	187
AIC		2391.482	2391.504	2391.506

Table 7.30: Maximum likelihood parameter estimates of Model 6 for $s\&p500$ data set at different levels of tolerance.

Parameter	Starting Values	$EM(tol = 10^{-4})$	$EM(tol = 10^{-5})$	$EM(tol = 10^{-8})$
$\hat{\alpha}$	0.6556607	1.349555	1.351381	1.351351
$\hat{\beta}$	-0.1257455	-0.1370256	-0.1372949	-0.1372904
$\hat{\delta}$	0.8310044	1.206995	1.208541	1.208515
$\hat{\mu}$	0.1690855	0.1788236	0.1789885	0.1789858
Loglikelihood		-1035.351	-1035.18	-1035.183
No. iteration		44	66	87
AIC		2078.702	2078.36	2078.366

Comparing the loglikelihood and the AIC, the proposed model fits the data better than the NIG distribution in all the three data sets.
Figures 2, 3 & 4 show how the proposed models fit the data sets. It is clear that the proposed model is a good alternative to the NIG.

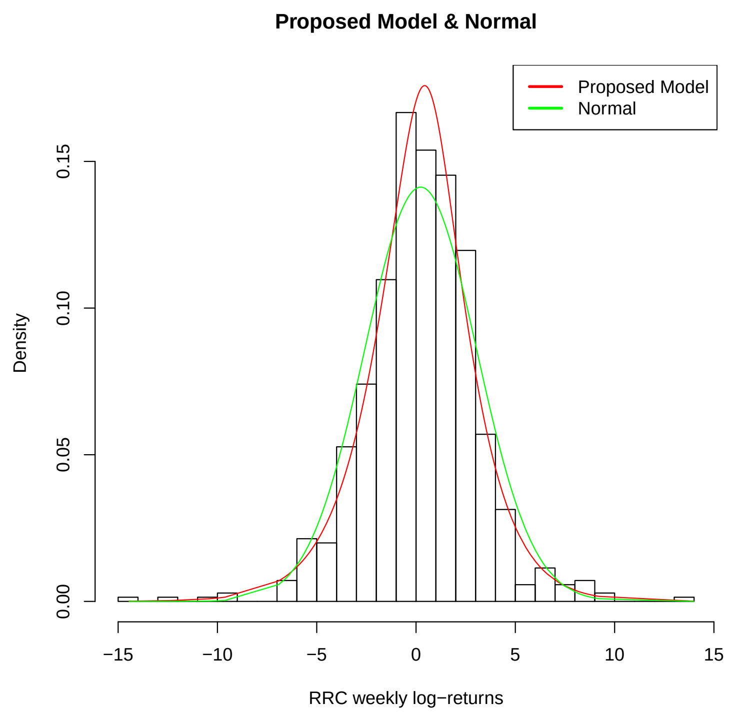

Figure 7.24. Fitting Model 6 to RRC weekly log-returns

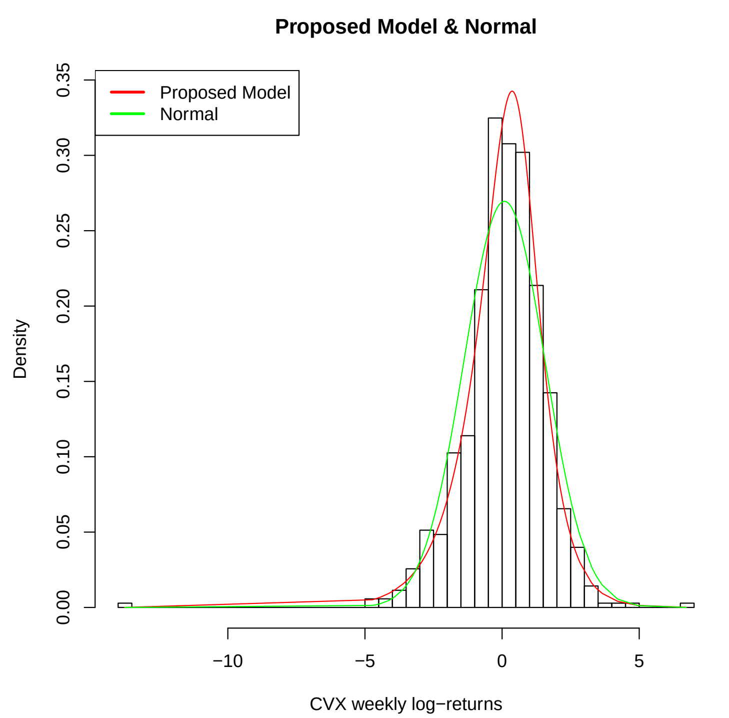

Figure 7.25. Fitting Model 6 to CVX weekly log-returns

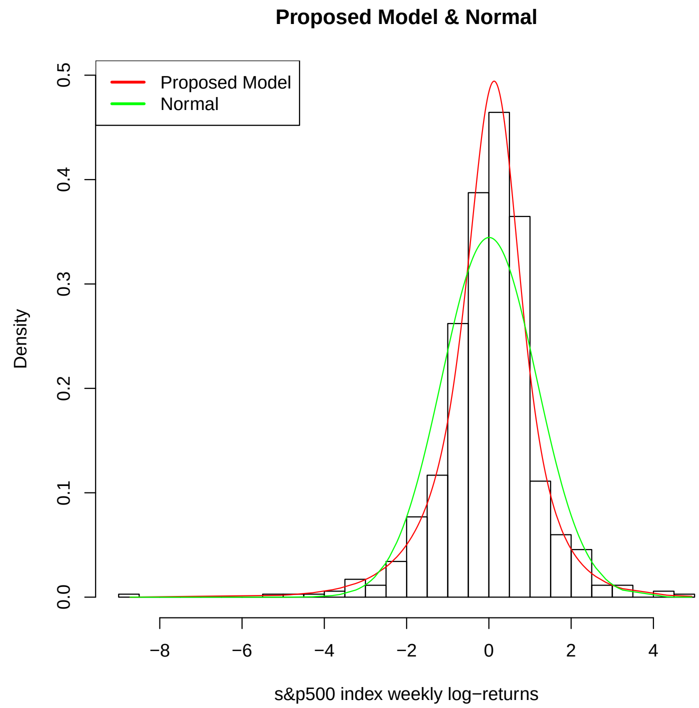

Figure 7.26. Fitting Model 6 to s&p500 index weekly returns

Remark:

Expressing the proposed model in terms of its components we have

$$f(x) = \frac{\delta^3}{\delta^3+\gamma^3} \times GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu) + \frac{\gamma^3}{\delta^3+\gamma^3} \times GHD(-\frac{3}{2}, \alpha, \delta, \beta, \mu) \qquad (7.7)$$

Table 7.31: Estimates of p in Model 6 for the data sets.

dataset	$\hat{p}$
RRC	0.04079
CVX	0.33275
s&p500	0.57922

The finite mixture therefore is flexible in determining between $GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu)$ and $GHD(\frac{3}{2}, \alpha, \delta, \beta, \mu)$ depending on the nature of the data.

8 RISK ESTIMATION USING NWIG DISTRIBUTIONS

In this chapter we consider the two leading measures of risk: Value at Risk (VaR) and Expected Shortfall (ES). Using the financial data sets, we measure the VaR and ES and perform the backtesting procedures.

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

$$P\{X \leq -VaR^X_{1-\alpha}\} = \alpha \tag{8.1}$$

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

$$\int_{-\infty}^{VaR^X_\alpha} f(x)dx = \alpha \tag{8.2}$$

where f(x) is the profit-loss distribution.
Expected Shortfall measures the expected loss in the tail of the distribution. From equation (8.2)

$$\frac{1}{\alpha}\int_{-\infty}^{VaR_\alpha} f(x)dx = 1 \tag{8.3}$$

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

$$E[X|X < VaR_\alpha] = \int_{-\infty}^{VaR_\alpha} x\frac{f(x)}{\alpha}dx \tag{8.4}$$

is the Expected Shortfall denoted as ES_α. This version was used by Yamai and Yoshima (2002) to obtain the ES for a normal distribution.
Equation (8.2) can be expressed in a different version as follows: Defining F(x) as the cdf of the random variable X, let

$$u = F(x) \Longrightarrow x = F^{-1}(u) \tag{8.5}$$

$$\therefore du = f(x)dx$$

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

$$x = VaR_\alpha \Longrightarrow u = \alpha$$

$$\therefore ES_\alpha = \frac{1}{\alpha}\int_0^\alpha F^{-1}(u)du \tag{8.6}$$

$$ES_\alpha = \frac{1}{\alpha}\int_0^\alpha VaR_u du \tag{8.7}$$

as presented by Zhang et al. (2019).
Emmer et al. (2015) proposed quantile approximation for equation (8.7) that takes the following form:

$$ES_\alpha \approx \frac{1}{4}\left[VaR_\alpha + VaR_{0.75\alpha+0.25} + VaR_{0.5\alpha+0.5} + VaR_{0.25\alpha+0.75}\right] \tag{8.8}$$

Kratz et al. (2018) generalize equation (13) to VaR levels of

$$\alpha_j = \alpha + \frac{j-1}{N}(1-\alpha), j = 1,...,N, N \in \mathbb{N} \tag{8.9}$$

When $N = 5$,

$$ES_\alpha \approx \frac{1}{5}\left[VaR_\alpha + VaR_{0.80\alpha+0.20} + VaR_{0.6\alpha+0.4} + VaR_{0.40\alpha+0.60} + VaR_{0.20\alpha+0.80}\right] \tag{8.10}$$

8.1 Risk Estimation and Backtesting

We use the parameter estimates for our proposed model to determine the VaR and ES 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 (Kupiec, 1995) 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_α or x_t<$VaR_{1-\alpha}$, and compare the corresponding failure rate to α.
The likelihood ratio statistic is given by

$$2\log(\frac{\tau(\alpha)}{n})^{\tau(\alpha)}(1-\frac{\tau(\alpha)}{n})^{n-\tau(\alpha)} - 2\log(\alpha^{\tau(\alpha)} - (1-\alpha)^{n-\tau(\alpha)}) \tag{8.11}$$

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.

8.2 Risk Estimation and Backtesting for the Special Cases of Generalised Hyperbolic Distribution

In this section we refer to the *NIG* as GHD($\lambda = -\frac{1}{2}$), *NRIG* as GHD($\lambda = \frac{1}{2}$), the special case of *GHD* when $\lambda = -\frac{3}{2}$ as GHD($\lambda = -\frac{3}{2}$) and the special case of *GHD* when $\lambda = \frac{3}{2}$ as GHD($\lambda = \frac{1}{2}$).
From the previous chapter, the ML estimates for the RRC data set via the EM-algorithm is summarized in the table below

Table 8.1: Estimates for the special cases of GHD.

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

The Kolmogorv-Smirnov and Anderson-Darling test performed on the 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 8.2: Goodness of Fit Test

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

Table below 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.

Table 8.3: AIC, BIC and Log-likelihood Values for the special cases of GHD.

Model	GHD($\lambda = -\frac{1}{2}$)	GHD($\lambda = \frac{1}{2}$)	GHD($\lambda = -\frac{3}{2}$)	GHD($\lambda = \frac{3}{2}$)
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

Model 3 has the lowest AIC and BIC with the highest log-likelihood. It is the best fit for the data. This special case outperforms the Normal Inverse Gaussian distribution.

We use the parameter estimates for RRC for our proposed models to determine the VaR and ES 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 (Kupiec, 1995) 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

$$2\log(\frac{\tau(\alpha)}{n})^{\tau(\alpha)}(1-\frac{\tau(\alpha)}{n})^{n-\tau(\alpha)} - 2\log(\alpha^{\tau(\alpha)} - (1-\alpha)^{n-\tau(\alpha)}) \quad \textbf{(8.12)}$$

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 8.4: VaR Values of RRC weekly log-returns for Normal and Proposed Models GHD special cases.

	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
GHD($\lambda = -\frac{1}{2}$)	-12.175020	-7.483157	-4.387882	4.621687	7.300979	11.305172
GHD($\lambda = \frac{1}{2}$)	-11.676119	-7.396380	-4.414590	4.635737	7.248426	10.976183
GHD($\lambda = -\frac{3}{2}$)	-12.770428	-7.524902	-4.344605	4.605084	7.328694	11.666360
GHD($\lambda = \frac{3}{2}$)	-11.206503	-7.271316	-4.422422	4.646686	7.176342	10.659890

Table 8.5: ES Values of RRC log-returns Based on Normal and GHD special cases.

	0.001	0.01	0.05
GHD($\lambda = -\frac{1}{2}$)	-14.31580521	-9.51044987	-6.32267305
GHD($\lambda = \frac{1}{2}$)	-13.54898243	-9.25410370	-6.26915453
GHD($\lambda = -\frac{3}{2}$)	-15.35943879	-9.77595177	-6.35304744
GHD($\lambda = \frac{3}{2}$)	-12.88318596	-8.98494206	-6.18936754

Table 8.6: 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
GHD($\lambda = -\frac{1}{2}$)	2	5	33	28	11	1
GHD($\lambda = \frac{1}{2}$)	2	5	33	28	10	1
GHD($\lambda = -\frac{3}{2}$)	2	5	33	27	11	1
GHD($\lambda = \frac{3}{2}$)	2	5	33	27	11	1

GHD($\lambda = -\frac{3}{2}$) has the highest VaR and ES value indicating that it perform well than the other models at the tails.

Table 8.7: 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
GHD($\lambda = -\frac{1}{2}$)	0.2067157	0.4191802	0.7134756	0.20316	0.1632629	0.7381375
GHD($\lambda = \frac{1}{2}$)	0.2067157	0.4181802	0.7134756	0.144112	0.1632629	0.7381375
GHD($\lambda = -\frac{3}{2}$)	0.2067157	0.4191802	0.7134756	0.20316	0.287939	0.7381375
GHD($\lambda = \frac{3}{2}$)	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 GHD($\lambda = -\frac{3}{2}$) (Normal-$GIG(-\frac{3}{2}, \delta, gamma)$) outperforms the other models at level 0.99.

In this section we obtained VaR using NWIG distributions. We considered: $GHD(-\frac{1}{2}, \delta, \gamma)$, $GHD(\frac{1}{2}, \delta, \gamma)$, $GHD(-\frac{3}{2}, \delta, \gamma)$ and $GHD(\frac{3}{2}, \delta, \gamma)$. We have shown that these mixing distribution are NWIG distributions.

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

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

NIG for VaR computation.

8.3 Risk Estimation and Backtesting for Normal Finite Weighted Inverse Gaussian Distribution

In this section we consider the **Six** models constructed based on finite weighted inverse Gaussian mixing distribution. The data used in this research is the Shares of Chevron (CVX) weekly returns for the period 3/01/2000 to 1/07/2013 with 702 observations. The histogram for the weekly log-returns in 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 8.8 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.

Table 8.8: Summary Statistics for CVX weekly log-returns.

Minimum	Standard.dev	skewness	exc.kurtosis	Maximum	Mean	N
-13.76000	1.480436	-1.297339	8.10113	6.71400	0.08711	702

The proposed models are now fitted to CVX 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 Karlis (2002).

$$\hat{\alpha} = 0.4190067, \hat{\beta} = -0.1054991, \hat{\delta} = 0.8324058, \hat{\mu} = 0.3036691.$$

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

Table 8.9: Maximum likelihood parameter estimates for Model 1-6

Parameter	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{\mu}$
Model 1	1.124238	1.574226	-0.2517274	0.572402
Model 2	1.612872	2.805817	-0.4751398	0.9361947
Model 3	0.976286	0.9163648	-0.1451396	0.4193146
Model 4	1.167188	1.631209	-0.2491203	0.5691122
Model 5	1.395091	1.480994	-0.2974406	0.6455598
Model 6	1.426388	1.767935	-0.2627765	0.5908363

The parameter estimates from table 8.9 are now fitted to RRC weekly log-returns. Figures 8.2, 8.3, 8.4 and 8.5 show the histogram and Q-Q plots of the RRC returns fitted with the proposed models. Figure 2-7 show that the proposed model fit the data well.

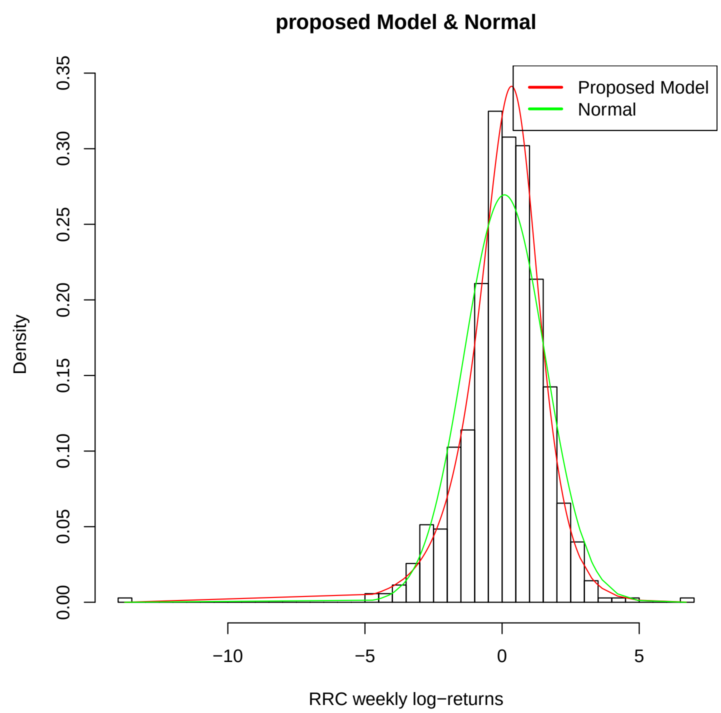

Figure 8.1. Fitting Model 1 to CVX weekly log returns

The table below 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.

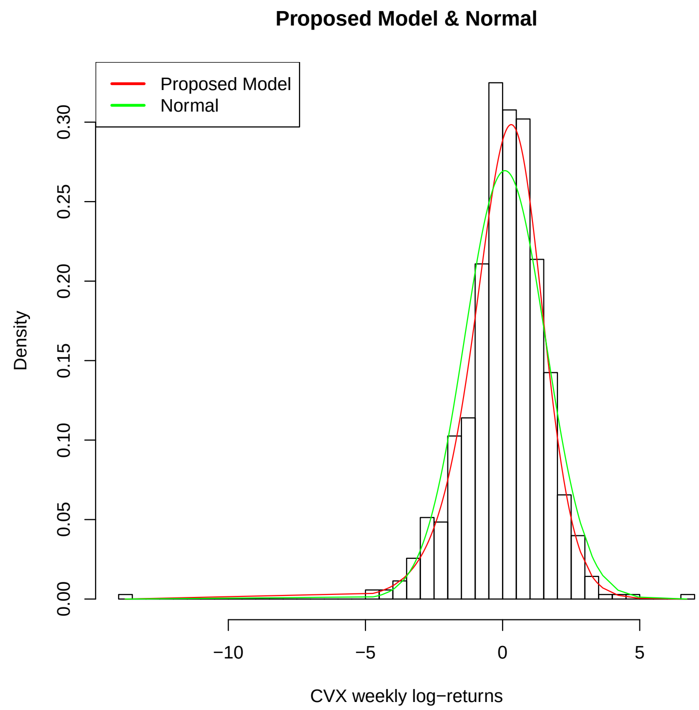

Figure 8.2. Fitting Model 2 to CVX log-weekly returns

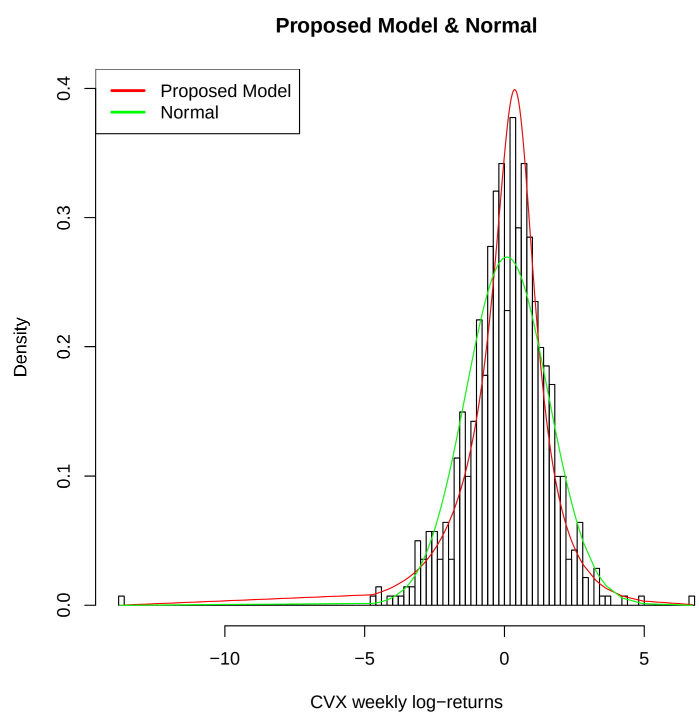

Figure 8.3. Fitting Model 3 to CVX log-weekly returns

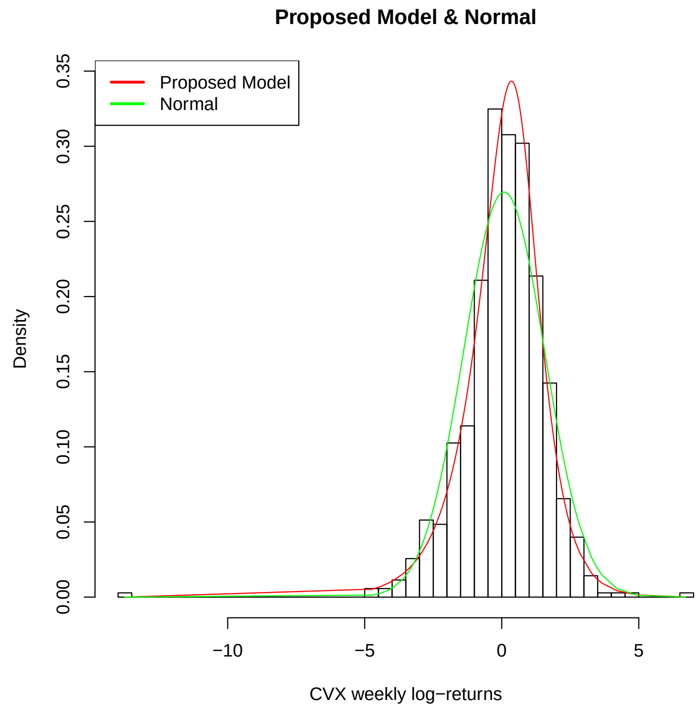

Figure 8.4. Fitting Model 4 to CVX log-weekly returns

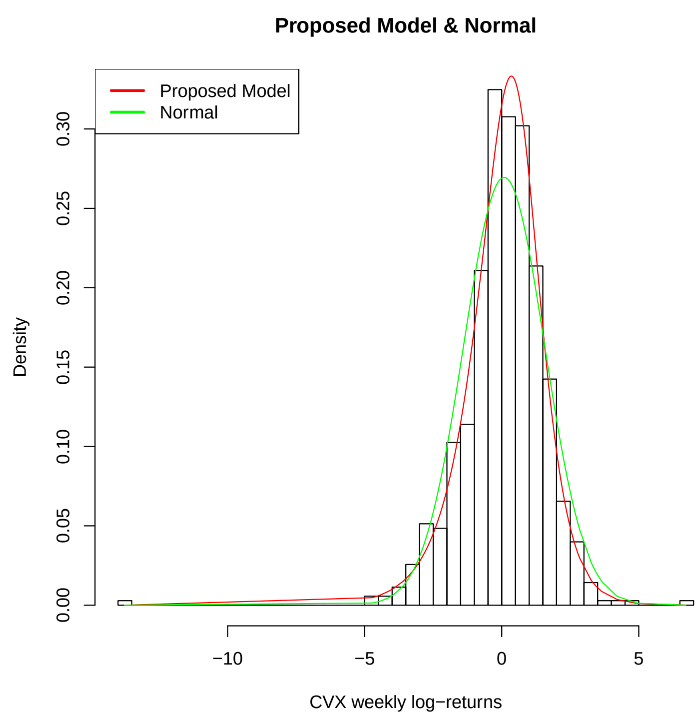

Figure 8.5. Fitting Model 5 to CVX log-weekly returns

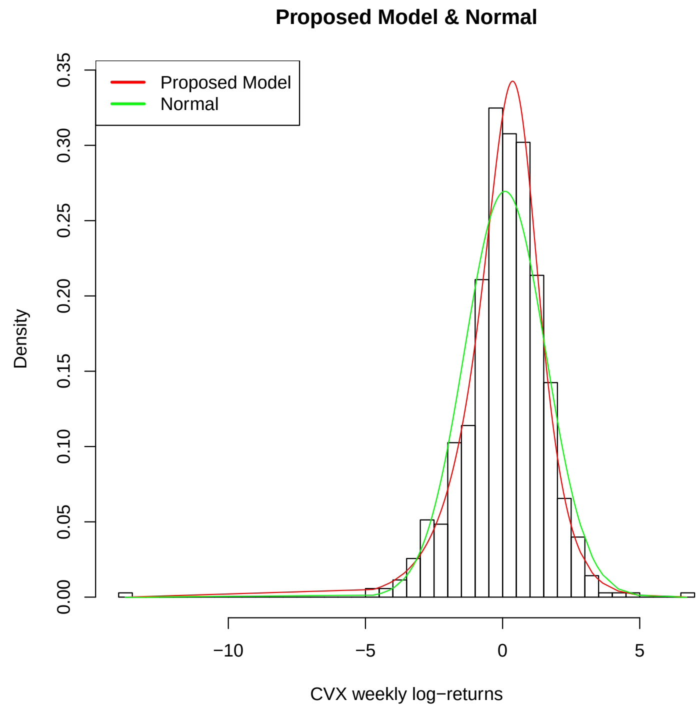

Figure 8.6. Fitting Model 6 to CVX log-weekly returns

Table 8.10: AIC, BIC and Log-likelihood Values for Model 1-6.

Model	Model 1	Model 2	Model 3	Model 4
AIC	2453.276	2461.924	2473.254	2453.912
BIC	2471.492	2480.140	2491.47	2472.128
Log-likelihood	-1222.638	-1226.962	-1232.627	-1222.956

Model	Model 5	Model 6
AIC	2455.008	2391.506
BIC	2473.224	2409.722
Log-likelihood	-1223.504	-1191.753

Remark:
Model 6 has the lowest AIC and BIC with the highest log-likelihood. It is the best fit for the data. Using Karlis (2002) formulation, the Normal Inverse Gaussian (NIG) parameter estimates for the $EM-algorithm$ are: $\hat{\alpha} = 0.9265284, \hat{\delta} = 0.9265284, \hat{\beta} = -0.2429664, \hat{\mu} = 0.5578459$. The $loglikelihood = -1221.667$ at $tol = 10^{-8}$ with 119

iterations. Therefore a finite mixture of $GIG(-\frac{3}{2}, \delta, \gamma)$ and $GIG(\frac{3}{2}, \delta, \gamma)$ is versatile compared to Inverse Gaussian (IG) distribution.

Table 8.11: VaR Values of CVX log-returns Based on the Normal and Model 1-6.

	0.001	0.01	0.05
Normal	-4.487787	-3.356904	-2.347995
Model 1	-6.373158	-4.008336	-2.376487
Model 2	-5.603337	-3.706563	-2.321636
Model 3	-6.6614465	-4.1473175	-2.4303994
Model 4	-6.1845506	-3.9279027	-2.3550354
Model 5	-6.005398	-3.881535	-2.358399
Model 6	-5.618886	-3.685542	-2.286172

Table 8.12: ES Values of RRC log-returns Based on the Normal and Model 1-6.

	0.001	0.01	0.05
Model 1	-7.417996	-5.033958	-3.392502
Model 2	-6.417102	-4.532881	-3.181197
Model 3	-7.776542	-5.236823	-3.499916
Model 4	-7.177348	-4.907374	-3.333809
Model 5	-6.918954	-4.805581	-3.304015
Model 6	-6.450259	-4.527236	-3.154907

Table 8.13: Number of violations of VaR for Model 1-6 at Different levels.

	0.001	0.01	0.05
Normal	3	10	42
Model 1	1	5	41
Model 2	1	6	42
Model 3	1	4	39
Model 4	1	6	42
Model 5	1	6	42
Model 6	1	6	42

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

Table 8.14: P-value for the Kupiec Test for Each Distribution at Different levels.

	0.001	0.01	0.05
Normal	0.04222549	0.2879388	0.245805
Model 1	0.7381375	0.4191802	0.3190668
Model 2	0.7381375	0.691514	0.245805
Model 3	0.7381375	0.2126411	0.5066538
Model 4	0.7381375	0.691514	0.245805
Model 5	0.7381375	0.691514	0.245805
Model 6	0.2067157	0.691514	0.245805

Remark:
At 5 percent level of significant, the Normal distribution is rejected at levels at the level 0.001. The Normal weighted Inverse Gaussian distributions were all effective and well specified on all levels of VaR.
In this work we constructed a class of weighted inverse Gaussian Distribution by considering a finite mixture of two special cases of Generalized Inverse Gaussian distribution. We considered the cases when the indexes are $-\frac{1}{2}$, $\frac{1}{2}$, $-\frac{3}{2}$ and $\frac{3}{2}$. These special cases are also weighted inverse Gaussian distributions.

We further used the class as mixing distributions to construct the Normal Variance-Mean Mixtures. The parameter estimates were obtained using the Expectation Maximization (EM) algorithm. We obtained a monotonic convergence for the iterative schemes of the models using the method of moments estimates of NIG as initial values.
We used AIC, BIC and loglikelihood for model selection. The model with the mixing distribution based on a finite mixture for $GIG(-\frac{3}{2},\delta,\gamma)$ and $GIG(\frac{3}{2},\delta,\gamma)$ was found to be the best model. The results show that the six models are sufficient for VaR computation.

8.3.1 Hypothetical Example

Suppose a company invest 100 million dollars in RRC, the absolute terms of economic capital to be set aside at $\alpha = 0.001$ will be

Table 8.15: Economical Capital.

	VaR(%)	ES (%)	VaR(Amount)	ES (Amount)
Normal	8.50	10.2	8,500,000	10,200,000
Model 1	12.1	14.3	12,100,000	14,300,000
Model 2	11.7	13.5	11,700,000	13,500,000
Model 3	12.8	15.4	12,800,000	15,400,000
Model 4	11.2	12.9	11,200,000	12,900,000

Remark: The Normal distribution underestimate risk and hence the required economic capital. Comparing with Model 3, the NIG distribution underestimate risk by more than 1.1 million dollars for ES.

9 DEPENDENCE MODELLING

9.1 Introduction

The joint distribution combines the information from the marginal distribution and the way in which the variables depend on each other.However it expresses this dependence implicitly. We cannot immediately see the nature of the Independence simply by looking at the formula for the joint distribution function. Copulas provide an alternative approach that expresses the interdependence between the variables explicitly. They allow us to deconstruct the joint distribution of a set of variables into compliments that can be adjusted individually.

9.1.1 Definition of a Copula

A copula is a function that expresses a multi variable cumulative distribution in terms of the individual marginal cumulative distribution. For a bi-variate distribution the cumulative distribution function

$$F_{X,Y}(x,y) = \text{Prob}(X = x, Y = y) = C_{X,Y}(F_X(x), F_Y(y))$$

Thus the copula function $C_{X,Y}$ is usually written in the more compact form

$$C(u,v) = F_{X,Y}(x,y) \tag{9.1}$$

Where;

$$u = F_X(x) \text{ and } v = F_Y(y)$$

The definition cab be extended to the multivariate case where we have;

$$C(u_1,u_2,...,u_d) = F_{x_1,x_2,...,x_d}(x_1,x_2,...,x_d)$$

Where

$$u_i = F_{X_i}(x_i) \tag{9.2}$$

9.1.2 Three Properties of Copula

Copulas must also satisfy three properties to ensure that they correctly capture the properties we would expect of a joint distribution in all circumstances.
Propety 1: A copula is an increasing function of its inputs.

$$C(u_1, u_2, ..., u_i^*, ..., u_d) > C(u_1, u_2, ..., u_i, ..., u_d)$$

For;

$$u_i^* > u_i; i = 1, 2, ..., d \quad (9.3)$$

Property 2

$$C(1, 1, ..., 1, u_1, 1, .., 1) = u_i \quad (9.4)$$

Property 3

$$C(u_1, u_2, .., u_d) \in [0, 1] \quad (9.5)$$

9.2 Sklar's Theorem

Sklar(1959) annotated that the dependent structure of a set of random variables can be captured independently using copulas.
Sklar's Theorem
Let F be joint cumulative distribution function with marginal cumulative distribution functions: $F_1, F_2, ..., F_d$.
Then there exists a copula function C such that for all $x_1, x_2, ..., x_d \in (-\infty, \infty)$

$$F(x_1, x_2, .., x_d) = C(F_1(x_1), F_2(x_2), ..., F_d(x_d)) \quad (9.6)$$

in the case of variables that have a continuous distribution, the copula is unique.
The Converse of Skalar's Theorem
If C is a copula and $F_1, F_2, ..., F_d$ are uni-variate cumulative distribution functions, then the function F defined above is a joint cumulative distribution with marginal cumulative distribution functions $F_1, F_2, ..., F_3$.
i.e

$$C(F_1, F_2, ..., F_d) = F_{X_1, X_2, ..., X_d}(x_1, x_2, ..., x_d)$$

9.2.1 Associations in Variables

Variables are said to be associated is there is some form of statistical relationship between them; whether causal or not. to facilitate comparison, measures of association can be constructed.
Coefficients of association are generally designed so that their values vary between -1 and +1. their absolute values increase with the strength of the relationships.
They have +1 (or -1) when there is perfect positive (or negative) association.
Any one particular type of coefficient of association measures a particular form of association.
For example, Pearson's Correlation Coefficient measures the degree to which there is linear relationship between the variables.
Concordance is another particular form of association. Broadly, two random variables are considered concordant if small values of one are likely to be associated with small values of the other and vice versa. Spearman's Rho and Kendell's Tau are two examples of measures of concordance.
Remarks:
1) A positive association between two variables does not necessary imply that one is dependent on the other. For example both might be dependent on the third(perhaps unobserved) variable with neither being directly dependent on the other.
2) A common pitfall is to forget that correlation does not imply Association.
Pearson's rho measures how strongly the variables are related.
Pearson's rho is defined as

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{VarXVarY}}$$

The estimate is given by

$$\rho_{XY} = \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x})(x_{i2}-\bar{x})}{\sqrt{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)^2}}$$

Where;
$\bar{x_1} = \sum_{i=1}^{n}\frac{x_i1}{n}, \bar{x_2} = \frac{1}{n} = \sum_{i=1}^{n} x_i2$ For a random sample $[x_i1, x_i2, i = 1, 2, ..., n]$
Remark.
Desirable properties of a measure of concordance or association between two variables should have a number of properties. these include in-variance which requires that the measure of concordance does not change if we apply the same monotone function to the value of each variable. Pearson's rho does not have this property. Remarks; Correlation Coefficients.
The commonly used measures of concordance that are more robust than Pearson's rho and are more invariant are Spearman's rho often called the rank correlation and Kendells tau.

Measures of rank correlation look at the association between the position (or rank) of observations in a data series when they are arranged in order.

Spearmans rho S_ρ can be calculated as;

$$S_\rho = \frac{\sigma}{N(N^2-1)} \sum_{i=n}^{N} d_i^2$$

N is the number of (pairs of) observations and d_1 the difference in the rank for the ith observation
Where ;
Both Kendell's tau and Spearmans rho are dependence measures which are rank based and therefore invariant with respect to monotone transformations of the marginals. Their range of values is the interval [-1, 1]. Additionally they can be expressed solely in terms of their associated copula and therefore their values does not depend on the marginal distribution. often there are closed-form expressions in term of the copula parameters available.
Kendels Tau denoted by T, is defined as the probability of discordance of two variables X_1 and X_2

$$T(x_1, x_2) = prob((x_{11} - x_{22})(x_{12} - x_{22}) > 0) - prob((x_{11} - x_{22})(x_{12} - x_{22}) < 0) \quad \textbf{(9.7)}$$

where (x_{11}, x_{12}) and $(x_{21}, x_{22})$are independent and identically distributed co*** of (x_1, x_2) non parametric estimation of Kenddel's T is treated in details in chapter 8 of Mollander et al., (2014)
In particular to estimate Kendells's T from a random sample $[x_1 1, x_i 2, i = 1, 2, ..., n]$ of size n from the joint distribution of (x_1, x_2) we consinder all

$$\frac{n}{2} = \frac{n(n-1)}{2} \text{unordered pairs} x_i = (x_i 1, x_i 2) \text{ and } x_j = (x_j 1, x_j 2) fori, j = 1, 2, m, n. = 0$$

9.3 Tail Dependence

The correlation measure described above each try to summarize the nature and extend of the associations between variables into a single statistic. Two key shortcomings of such statistics are;

1. Information lost in the summarization process

2. They capture the interdependence through the whole distribution. This may be of less interest than the inter-dependency in the tails of the distributions.
   it is often the case in insurance and investment application that large losses tend to occur together.

So the relationships between the variables at the extremes of the distributions are of particular importance. These can be measured using the coefficients of upper and lower tail dependence.

9.3.1 The coefficient of upper tail dependence

The coefficient of upper tail dependence is given as

$$\begin{aligned}\lambda_u &= \lim_{u\to 1} Prob(X > F_X{}^{-1}(U)/Y > F_Y{}^{-1}(U)) \\ &= \lim_{u\to 1} \frac{1-2u+C(u,u)}{1-u}\end{aligned} \tag{9.8}$$

The coefficient is a probability so it is a value between 0 and 1. The coefficient of upper tail dependence indicates whether high values of one random variable X tend to be linked with high values of another random variable Y. Specifically, the coefficient of upper tail dependence is the limiting value of this probability as $\psi \to 1-1$;that is as we move from below into the upper tail.

9.3.2 The Coefficient of Lower Tail Dependence

$$\begin{aligned}\lambda_L &= \lim_{u\to 0^+} Prob(X <= F_X{}^{-1}(U)/Y <= F_Y{}^{-1}(U)) \\ &= \lim_{u\to 0^+} C(u,u)\end{aligned}$$

Again, this coefficient is a probability so it takes a value between 0 and 1. The coefficient of lower tail dependence indicates whether low values of one random variable X are linked with low values of another random variable Y. Specifically, the coefficient of lower tail dependence is the limiting value of the probability as;
$U \to 0^+$ i.e are more further into the lower tail from above.

9.4 Types of Copulas

1. Fundamental Copulas

   (a) Independence (product) copula

   (b) Co-monotonic (minimum) copula
   (c) Counter-monotonic (maximum) copula

2. Implicit Copulas
   (a) Gaussian Copula
   (b) Student's t Copula

3. Explicit Copulas (Archimedean copulas)
   (a) Gumbel Copula
   (b) Clayton Copula
   (c) Frank Copula

Archimedean copulas are inadequate for more than 2 variables since they do not allow different dependence patterns between pairs of variables, Vine copulas or hierarchical copulas can be considered.
Lo et al. (2013) used canonical vine (C-vine) copulas. Kraus and Czando (2017) use D -vine copula and Olechrin and Teteneva (2017) use hierachial copula to estimate VAR. Bynn and Sony (2021) sought the best copula calculating VAR of a portfolio with many assets. They used the vine copulas and hierachial copulas. As for the marginal distribution we can use different distributions to each asset in a portfolio. It has been shown that the normal distribution is inappropriate to model the return distribution of financial assets. The return distributions of financial assets are slightly skewed and fat tailed. There has been expensive research for good alternatives to the normal distribution in literature.
For example, Venkiataraman (1997) used a quasi-bayesian maximum likelihood estimation procedure. Hull and White (1998) used a transform to multivariate normal distribution which is updating schemes such as GARIH. Ebercein and Keeler (1995) used hyperboli8c distribution. Modem et al. (1998) use Variance Gamma (VG) distribution. Bandoff Nelsen (1997) and Mabitreda et al. (2015) used Normal Inverse Gaussian distribution. Byun and Sony (2021) Used NIG distribution as marginal distribution. The NIG is known to have better return portfolio than VG distribution. (Ericksen et al. 2009, Gencil and Yeng 2010) and the calculation of VAR using NIG is also better than other models such as GARCH or VG. (Welhelmson, 2009, Kim and Song, 2011, Doric and Doric, 2011.) Bolvinken and Benth (2000), Godin et al. (2012) have also used the NIG distribution for modeling return distributions of financial assets.
Copulas are popular in modelling a joint distribution of several asset returns in finance. With copulas we can construct a multivariate distribution with different marginal distributions by separating the dependence from marginal distributions. Also, there are many copulas that can incorporate the proper dependence structure of the data. Embrechls et al. (2002) show that copulas are useful in identifying the dependence structure annually returns of assets. Bynn and Song (2021) used copulas to identify the dependence structures and to

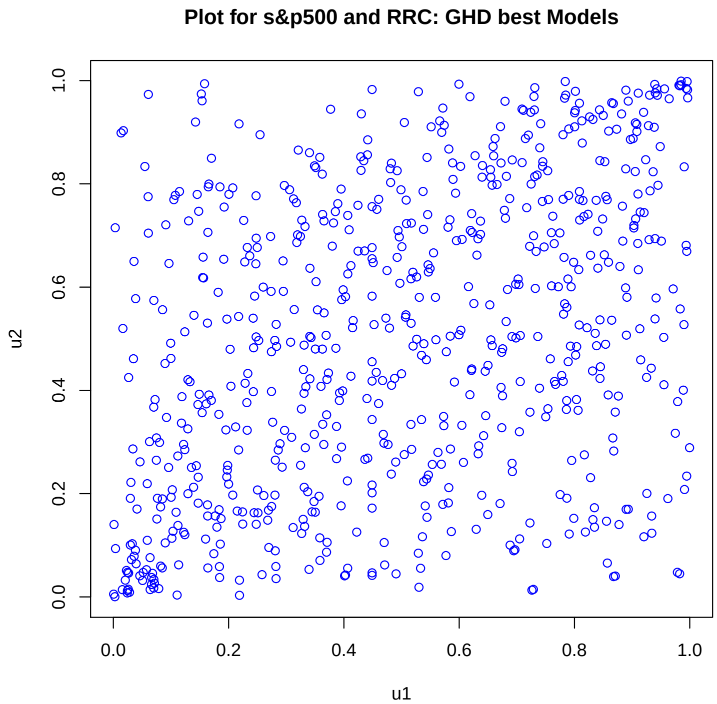

Figure 9.1. Cumulative returns plot for RRC and S&p500 index

generate a multivariate distribution for returns of assets in a portfolio. The VAR of the portfolio was computed using the resulting multivariate distribution. Wu et al. (2007) use exchangeable Archimedean copulas to calculate VAR. When there are more than two assets in a portfolio, elliptical and exchangeable.

9.5 Fitting Bivariate Returns Using Copulas with GHD (Model 3)

In this section we consider bivariate returns for the RRC and s&p500 index. The scatter plot for the returns is present in the figure below

We Have used the special case of the GHD when index parameter is $-\frac{3}{2}$ to model the univariate distribution for the individual returns.

9.5.1 Parameter Estimation for selected Copulas

The copula parameter estimation of RRC and s&p500 index, for "best model" is obtained using maximum likelihood method as follows: **Gaussian Copula**

$$C(u_1,u_2;\theta) \quad = \quad \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{\frac{1}{2}}} \left[\frac{-(s^2-2\theta st+t^2)}{2(1-\theta^2)}\right] dsdt \tag{9.9}$$

$$\hat{\theta} = 0.4211174$$

t-copula

$$C(u_1, u_2; \theta_1, \theta_2) \quad = \quad \int_{-\infty}^{t^{-1}(u_1)} \int_{-\infty}^{t^{-1}(u_2)} \frac{1}{2\pi(1-\theta_2^2)^{\frac{1}{2}}} \left[1 + \frac{s^2 - 2\theta_2 st + t^2}{\nu(1-\theta_2^2)}\right]^{-\frac{\theta_1+2}{2}} dsdt \tag{9.10}$$

$$\begin{aligned} \hat{\theta}_1 &= 0.4217174 \\ \hat{\theta}_2 &= 4.210349 \end{aligned}$$

Clayton Copula

$$C(u_1, u_2; \theta) = (\mu_1^{-\theta} + \mu_1^{-\theta} - 1)^{\frac{-1}{\theta}} \tag{9.11}$$

$$\hat{\theta}_1 = 0.639253$$

Gumbel Copula

$$C(u_1, u_2; \theta) = \exp\{-[(-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}]^{\frac{1}{\theta}}\} \tag{9.12}$$

$$\hat{\theta}_1 = 1.35326$$

Frank Copula

$$C(u_1, u_2; \theta) = -\frac{1}{\theta} ln(1 + \frac{e^{\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}) \tag{9.13}$$

$$\hat{\theta}_1 = 2.69309$$

Joe Copula The maximum likelihood parameter estimate

$$\hat{\theta}_1 = 1.418371$$

9.5.2 Goodness of Fit Test

We now perform goodness of fit test based on Kendall's process for bivariate copula data as investigated by Genest and Rivest (1993) and Wang and Wells (2000). We have also computed the Cramer-von Mises (CvM) and Kolmogorov-Smirnov test statistics, respectively, as well as their corresponding p-values using bootstrapping.

Gaussian Copula

	statistics	p-value
CvM	0.1143797	0.17
KS	0.7894206	0.29

Clayton Copula

	statistics	p-value
CvM	0.1527572	0.08
KS	1.004526	0.07

Gumbel Copula

	statistics	p-value
CvM	0.472664	0
KS	1.434981	0

Frank Copula

	statistics	p-value
CvM	0.3074956	0
KS	1.334435	0

Joe Copula

	statistics	p-value
CvM	5.29647	0
KS	1.334435	0

The best copula using the AIC technique for this data was found to be the **Student's t Copula**.

9.6 Fitting Bivariate Returns Using Copulas with Model 6

In this section we consider bivariate returns for the RRC and s&p500 index. The scatter plot for the returns is present in the figure below

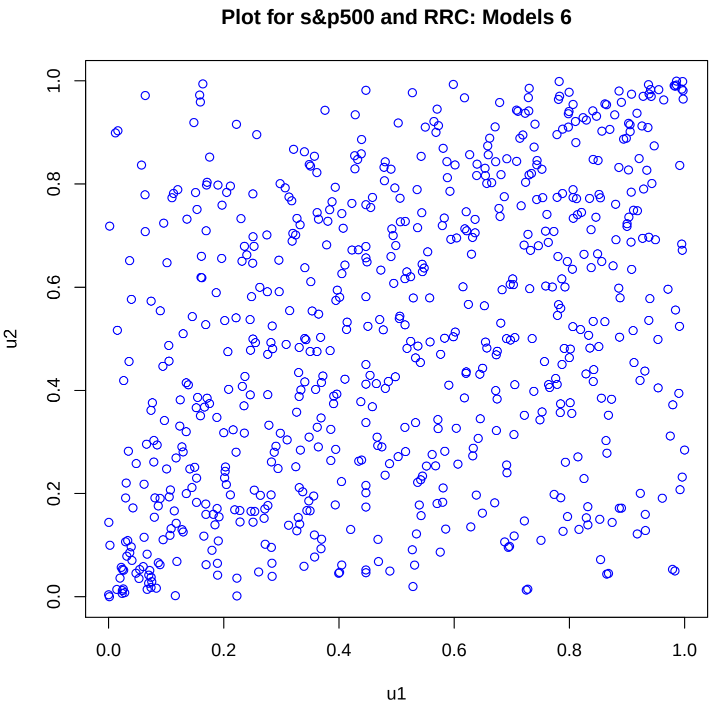

Figure 9.2. Cumulative returns plot for RRC and S&p500 index

We Have used Model 6 to model the univariate distribution for the individual returns.

9.6.1 Parameter Estimation for selected Copulas

The copula parameter estimation of RRC and s&p500 index, for "best model" is obtained using maximum likelihood method as follows: **Gaussian Copula**

$$C(u_1, u_2; \theta) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{\frac{1}{2}}} \left[\frac{-(s^2 - 2\theta st + t^2)}{2(1-\theta^2)} \right] dsdt \quad \textbf{(9.14)}$$

$$\hat{\theta} = 0.4213349$$

t-copula

$$C(u_1, u_2; \theta_1, \theta_2) = \int_{-\infty}^{t^{-1}(u_1)} \int_{-\infty}^{t^{-1}(u_2)} \frac{1}{2\pi(1-\theta_2^2)^{\frac{1}{2}}} \left[1 + \frac{s^2 - 2\theta_2 st + t^2}{\nu(1-\theta_2^2)} \right]^{-\frac{\theta_1+2}{2}} dsdt$$

(9.15)

$$\hat{\theta}_1 = 0.4243634$$
$$\hat{\theta}_2 = 4.765168$$

Clayton Copula

$$C(u_1, u_2; \theta) = (\mu_1^{-\theta} + \mu_1^{-\theta} - 1)^{\frac{-1}{\theta}} \quad \textbf{(9.16)}$$

Gaussian Copula

	statistics	p-value
CvM	0.1425418	0.15
KS	0.9674326	0.14

$$\hat{\theta}_1 = 0.6188705$$

Gumbel Copula

$$C(u_1, u_2; \theta) = \exp\{-[(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{\frac{1}{\theta}}\} \tag{9.17}$$

$$\hat{\theta}_1 = 1.350768$$

Frank Copula

$$C(u_1, u_2; \theta) = -\frac{1}{\theta} ln(1 + \frac{e^{\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}) \tag{9.18}$$

$$\hat{\theta}_1 = 2.700615$$

Joe Copula The maximum likelihood parameter estimate

$$\hat{\theta}_1 = 1.408573$$

9.6.2 Goodness of Fit Test

Clayton Copula

	statistics	p-value
CvM	0.116485	0.2
KS	0.8321657	0.24

Gumbel Copula

	statistics	p-value
CvM	0.4794335	0
KS	1.440484	0

Frank Copula

	statistics	p-value
CvM	0.3082723	0
KS	1.332006	0

Joe Copula

	statistics	p-value
CvM	5.033468	1
KS	3.230643	1

The best copula using the AIC technique for this data was found to be the **Student's t Copula**.

10 CONCLUSION AND RECOMMENDATIONS

The objective of this research was to construct a new class of distribution known as Normal Weighted Inverse Gaussian (NWIG) distributions for Value at Risk (VaR) and Expected Shortfall (ES) computation.

10.1 Normal Weighted Inverse Gaussian Distribution based on Special Cases of Generalised Inverse Gaussian

Generalized Inverse Gaussian distribution with parameters: λ,δ,γ denoted by $GIG(\lambda, \delta, \gamma)$ nests a number of special cases by varying the parameter λ. When used as a mixing distribution in the Normal Variance Mean Mixture (NVMM) we obtain the Generalised Hyperbolic Distribution (GHD). The Normal Inverse Gaussian (NIG) distribution is obtained using Inverse Gaussian (IG) as the mixing distribution. It can also be obtained as a special case of GHD when the index parameter $\lambda = -\frac{1}{2}$. In literature, the other special cases have been assumed if not neglected. The Inverse Gaussian is obtained as a special case of GIG when $\lambda = -\frac{1}{2}$. When $\lambda = \frac{1}{2}$, we have $GIG(\frac{1}{2}, \delta, \gamma)$ which is the Reciprocal Inverse Gaussian (RIG) distribution. Other cases we have considered in our work are when $-\frac{3}{2}$, $\frac{3}{2}$, $-\frac{5}{2}$ and $\frac{5}{2}$. We have further shown that these special cases are weighted Inverse Gaussian distribution.
These special cases have been used as mixing distribution for the NVMM to obtain the other special cases for the GHD. We obtained the properties of the mixed models. The cases when $\lambda = \frac{1}{2}$ and $\frac{3}{2}$ have been found to be alternatives to the NIG. The case when $-\lambda = \frac{3}{2}$ has outperformed all the other models in application to the three financial data set considered in our research.

10.2 Normal Weighted Inverse Gaussian Distribution based on finite Cases of Generalised Inverse Gaussian

We extended the class of NVMM by using finite mixtures of Weighted Inverse Gaussian (WIG) as mixing distribution. This idea is motivated by the fact that finite mixtures are more flexible than single distributions. The four special cases of the GIG distribution when $-\frac{1}{2}$, $\frac{1}{2}$, $-\frac{3}{2}$ and $\frac{3}{2}$ give **Six** can be combined in six different ways. Therefore we obtain **Six** finite mixtures of the special cases. We also showed that these cases can be expressed as Weighted Inverse Gaussian distributions.
We obtained Normal Variance Mean Mixtures using the six cases as mixing distributions and obtained their mean and variance. One case in particular, when the mixing distribution is a finite mixture GIG with indexes $-\frac{3}{2}$ and $\frac{3}{2}$, had the least AIC compared to the other

cases. The Model performed better than the "best" special case of the GHD identified in the section above.

10.3 Parameter Estimation

The most common Parametric method of estimation are Method of Moment (MoM) and Maximum Likelihood (ML) method. Of all the **10** proposed methods, its only the *NIG* that MoM estimates can be obtained directly. Numerical techniques are required to obtain the estimates for the other models.
We obtained ML estimates via the EM-algorithm. We identified iterative schemes based on explicit solution to the normal equation for the special cases of the GHD. In cases where the normal equation were quantities difficult to solve, a subtle approach was adopted. We designed iterative schemes based on a representation of the normal equations. This new approach, when applied to the special cases of the GHD gave same results. The approach preserves the properties of the EM-algorithm and we obtained a monotonic convergence for all models. The MoM estimates for the *NIG* were used as initial values for all the proposed models.

10.4 Application

We have used three data sets for this work: Range Resource Corporation (RRC), Shares of Chevron Corporation (CVX) and s&p500 index weekly log returns with 702 observations. The data sets exhibit non-normal characteristics: Skewed, fat tailed and leptokurtic. The Proposed Models fit the data sets well. The Kolmogorv-Smirnov and Anderson-Darling test performed on the models produce high p-values, a strong evidence that returns data follow the proposed models. Using the loglikelihood and AIC, the proposed models are found good alternative to NIG. outperforms the NIG. In particular, the special case of the GHD with $\lambda = -\frac{3}{2}$ outperforms the *NIG*. Intrestingly, the NVMM with a finite mixture of GIG of indexes $-\frac{3}{2}$ and $\frac{3}{2}$ outperforms all the proposed models.

10.5 Risk Measures

The most popular measures for financial risk are Value at Risk (VaR) and Expected Shortfall (ES). 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. We used the class of Normal Weighted Inverse Gaussian distribution to perform VaR and ES for the RRC, CVX and S&p500 index. Backtesting for VaR and ES was also performed. The (Normal-$GIG(-\frac{3}{2}, \delta, \gamma)$) turned to be the best model for economical capital allocation.

10.6 Dependence Modeling

Copulas provide an alternative approach that expresses the interdependence between the variables explicitly. We have used the Normal-$GIG(-\frac{3}{2}, \delta, \gamma)$ and $Model6$ as marginals for the asset returns. A goodness of fit procedure has been performed for the elliptical copulas and Archmedian copulas used.

References

[1] Aas, K. and Haff, H. (2005) NIG and Skew student's t: Two special cases of the Generalised Hyperbolic. Norwegian Computing Center.

[2] Aas, K. and Haff, H. (2006) The Generalised Hyperbolic Skew Student's t-distribution. Journal of Financial Econometrics, 4(2): 275–309.

[3] Aas, K and Dimkos, K. (2005) Risk Estimation Using the Multivariate Normal Inverse Gaussian Distribution. The Journal of Risk. 8, 39-60.

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

[5] Akman, H. O. and Gupta, R. C. (1992) A comparison of various estimators of the mean of an inverse Gaussian distribution. J. Stat. Comp. and Simulation, 40, 71-81.

[6] Artzner P., F. Delbaen, J-M. Eber and D. Heath (1997), 'Thinking Coherently', Risk Journal, 10(11): 68–71.

[7] Artzner P., F. Delbaen, J-M. Eber and D. Heath (1997) 'Coherent Measures of Risk', Mathematical Finance, 9(3): 203–28.

[8] Azzalini, A. (1985) A class of distribution which includes the normal ones. Scandinavian Journal of Statistics, 12:171-178.

[9] Azzalini, A. (1986) On a class of distributions which includes the normal ones. Statistica, 46:199-208. Bakouch, H. S; Al-Zahrain, B. M.; Al-Shamrani, A. A; Marchi, A. A and Louzada, F. (2012) An Extended Lindley Distribution. Journal of the Korean Statistical Society, 41: 75-85.

[10] Bams, D. and Wilhouwer, J. L. (2001) Empirical Issues In Value at Risk. Astin Bulletin, 31(2): 299-315

[11] Barndorff-Nielsen, O. E. (1977) Exponentially decreasing distributions for the logarithm of particle size. Proc. R. Soc. Lond. A, 353:409-419.

[12] Barndorff-Nielsen, O. E. (1997) Normal Inverse Gaussian Distribution and Stochastic Volatility Modelling. Scandinavian Journal of Statistics, 24:1-13.

[13] Barndorff-Nielsen, O.E., Prause, K. (2001) Apparent scaling. Finance and Stochastics 5, 103-113.

[14] Dempster, A. P., Laird, N. M. and Rubin, D. (1977) Maximum likelihood from incomplete data using the EM algorithm. Journal Roy. Statist. Soc. B, 39:1-38.

[15] Basel Committee.(2013) Fundamental review of the trading book: a revised market risk framework.

[16] Bolvinken, E. and Benth, F. E. (2000) Quantification of Risk in Neruregian stock via Normal Inverse Gaussian Distribution. In proceedings of the AFRRI (2000) Colloquim, Tromso, Norway, 87-98

[17] Bhati, D., Kumawat, P. and E. Gomez-Deniz (2015). A New Count Model Generated from MixedPoisson Transmuted Exponential Family with an application to Health Care Data.Communication in Statistics: Theory and Methods

[18] Corlu, C.G. and Corlu, A. (2014) Modeling Exchange Rate Returns: Which Flexible distributions to use? Quantitative Finance, 15(11):1851-1864.

[19] Dempster, A. P., Laird, N. M. and Rubin, D. (1977) Maximum likelihood from incomplete data using the EM algorithm. Journal Roy. Statist. Soc. B, 39:1-38.

[20] Deric, D. and Deric, E. N. (2011) Return Distribution and Value at Risk Estimation for BELEX 15. Yungslow Journal of Operations Research. 21, 103-118

[21] E. Eberlein and U. Keller. Hyberbolic distributions in finance. Bernoulli, 1(3): 281-299, 1995.

[22] Fisher, R. A. (1934) The Effect of Methods of Ascertainment upon the Estimation of Frequencies. The Annals of Human Genetics.

[23] Ghitany, M. E; Atich, B. and Nadarajah, S. (2018) Lindley distribution and Its Application. Mathematics and Computers in Simulations, 78(4): 493-506.

[24] Gneiting, T. (2011) Making and evaluating point forecasts. Journal of the American Statistical Association, 106(494):746–762.

[25] Goncu, A. and Yang, U. (2016) Variance Gamma and Normal Inverse Gaussian models: goodness of fit to Chinese high frequency index returns. The North American Journal of Economics and finance. 36,279-292.

[26] Gupta, R. C. and Akman, O. (1995) On Reliability Studies of a Weighted Inverse Gaussian Distribution Model. Statistical Papers, 38: 445-452.

[27] Gupta, R. C. and Kundu, D. (2011) Weighted inverse Gaussian–a versatile lifetime model. Journal of Applied Statistics, 38:12,2695-2708.

[28] A Probabilistic Representation of the skew normal distribution. Scandinavian Journal of Statistics, 13:271-275

[29] Hull, J. and White, A. (1998) Value at Risk when daily changes in the market variables are not normally distributed. Journal of Derivatives. 5, 9-19.

[30] Jorgensen, B, Seshadri, V. and Whitmore, G. A. (1991) On the Mixture of Inverse Gaussian Distribution with its Complementary Reciprocal. Scandinavian Journal of Statistics, 18: 77-89.

[31] Karlis, D. (2002) An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution. Statistics & Probability Letters, 57:43-52.

[32] Kim, T. and Song, S. (2011) "Value at Risk estimation using NIG and VG distributions" Journal of Korean Data Analysis Society, 13, 1775-1778.

[33] Kostas, F. (2007) Tests of fit for symmetric variance gamma distributions. UADPhilEcon, National and Kapodistrian University of Athens, 14 Euripidou Street, 105 59 Athens, Greece.

[34] Kupiec, P. (1995) Techniques for verifying the accuracy of risk measurement models.

[35] Lin, T; Lee, J; Yen, S (2007) Finite Mixture Modeling Using Skew Normal Distribution. Statistica Sinica, 7: 241-266.

[36] Lindley, D. V. (1958) Fiducial Distributions and Bayes Theorem. JRSS 20(1):102-107

[37] Mahmoudi, E. and Zakerzadeh, H. (2010). Generalized Poisson-Lindley distribution.Communications in Statistics - Theory and Methods,39:10, 1785-1798.

[38] Marc, Paolella, (2007). Intermediate Probability: A Computational Approach. Swiss Banking Institute, University of Zurich, Switzerland

[39] 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.

[40] K. Prause. Modelling financial data using generalized hyperbolic distributions. PhD Thesis. FDM preprint 48, university of Freiburg, 1999

[41] Nadarajah, S., Zhang, B. and Chan, S. (2014) Estimation methods for expected shortfall. Quantitative Finance, 14(2): 271-291.

[42] , H.; Jamalizadel, A.; Shafiel, S. and Balakrishnan, N. (2018) Mean Mixture of Normal Distribution: Properties, Inference and Application. Metrika

[43] Rappoport, P. (1993) A new approach: Average shortfall. Technical report, J.P. Morgan, 1993.

[44] Sankaran, M. (1970a). The Discrete Poisson-Lindley Distribution.Biometrics,26, 145-149.

[45] Shanker, R. and Hogos, F. (2015): On Poisson Lindley Distribution and Its Applications to Biological Sciences. Biometrics and Biostatistics International Journal, 2(4): 1-5.

[46] Venkataraman, S (1997). Value at Risk for a mixture of normal distributions: The use of quasi-bayesian elimination techniques. Economic Perspectives. Federal Reserve Bank of Chicago, 21 2-13.

[47] Wimmerstedt, L. (2015) Backtesting Expected Shortfall: the design and implementation of different backtest. Master's Thesis in Mathematical Statistics. Royal Institute of Technology.

[48] Yamai, Y. and Yoshiba, T. (2002) On the validity of Value at Risk comparative analyses with expected shortfall. Monetary and Economic Studies.

[49] Zakerzadeh, H. and Dolati, A. (2010). Generalized Lindley Distribution.Journal of MathematicalExtension3, 13-25

[50] 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

[51] Zhang, Y. and Nadarajah, S. (2017). A review of backtesting for values at risk. Communications in Statistics- Theory and Methods, 1-24

11 List of Publications

1. Maina, C. , Weke, P. , Ogutu, C. and Ottieno, J. EM-algorithm for a normal-weighted inverse Gaussian distribution: Normal-reciprocal inverse Gaussian distribution, Advances and Applications in Statistics 72 (2022), 1-24.
   DOI: 10.17654/0972361722001

2. Maina, C. , Weke, P. , Ogutu, C. and Ottieno, J. Properties, estimation and application to financial data for generalized hyperbolic distribution when the index parameter is -3/2 , Advances and Applications in Statistics 71(1) (2021), 55-84. DOI: 10.17654/AS071010055

3. Maina, C. , Weke, P. , Ogutu, C. and Ottieno, J. Normal- mixture with application to financial data, Far East Journal of Theoretical Statistics 63(2) (2021), 93-125. DOI: 10.17654/0972086321003

4. Maina, C. , Weke, P. , Ogutu, C. and Ottieno, J. Modelling Skewed And Heavy-tailed Data Using A Normal Weighted Inverse Gaussian Distribution. Afr. Stat. 17 (1) (2022) 3165 - 3187. DOI:10.16929/as/2022.3165.300