Modelling and Analysis of Financial Time Series using Some non-Gaussian Models

Thesis submitted to the
Cochin University of Science and Technology
for the Award of Degree of

Doctor of Philosophy

under the Faculty of Science

by

Sri Ranganath C. G.

DEPARTMENT OF STATISTICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
COCHIN-682 022

January 2018

CERTIFICATE

Certified that the thesis entitled **"Modelling and Analysis of Financial Time Series using Some non-Gaussian Models "** is a bonafide record of work done by Mr. Sri Ranganath C. G. under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included anywhere previously for the award of any degree or title.

Kochi- 22

January 2018

Dr. N. Balakrishna

Professor,

Department of Statistics,

Cochin University of

Science and Technology.

CERTIFICATE

Certified that all the relevant corrections and modifications suggested by the audience during pre-synopsis seminar and recommended by the Doctoral committee of the candidate has been incorporated in the thesis.

Kochi- 22

January 2018

Dr. N. Balakrishna

Professor,

Department of Statistics,

Cochin University of

Science and Technology.

DECLARATION

This thesis contains no material which has been accepted for the award of any other Degree or Diploma in any University and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due references are made in the text of the thesis.

Kochi- 22 **Sri Ranganath C. G.**

January 2018

ACKNOWLEDGEMENTS

I am deeply indebted to many who have generously helped me in completing this thesis and I take this opportunity to express my sincere thanks to each and every one of them.

I wish to express my deep sense of respect and gratitude to my thesis supervisor, Prof. N. Balakrishna, Department of Statistics, Cochin University of Science and Technology (CUSAT), who has been a constant source of inspiration during the course of my Ph.D work. He has always been patient towards my shortcomings and kept encouraging me to work in a better way. Without his help and support, perhaps, I would have not been able to write this thesis.

I am obliged to Prof. K. R. Muraleedharan Nair, Prof. V. K. Ramachandran Nair, Prof. K. C. James, former Heads, Department of Statistics, CUSAT and Prof. Asha Gopalakrishnan, Head, Department of Statistics, CUSAT for their extensive support. I take this opportunity to record my sincere respect and heartiest gratitude to Prof. P. G. Sankaran, Pro-Vice Chancellor and former Head, Department of Statistics, CUSAT and Prof. S. M. Sunoj, Dr. G. Rajesh, Associate Professor, and the other faculty at Department of Statistics, CUSAT for their valuable suggestions and wholehearted support.

I express deep gratefulness to all my teachers, who provided insight and expertise.

I convey my sincere thanks to the administrative and non-teaching staff, Department of Statistics, CUSAT for the help they rendered.

Discussions with my friends and other research scholars of the department often helped me during the work. I express my sincere thanks to all of them for their valuable suggestions and help.

I thank all my friends and well wishers who have been constantly encouraging me throughout my study.

I also would like to thank CUSAT, Department of Science and Technology (DST), Govt. of India and Kerala State Council for Science, Technolgy and Environment (KSCSTE) for the financial support provided during my research.

I am deeply indebted to my beloved father, mother and brother for their encouragement, prayers and blessings given to me.

Above all, I bow before the grace of the Almighty.

Sri Ranganath C. G.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The analysis of time series in classical setup utilizes Gaussian linear models to explain the phenomenon. With this assumption, the inference procedures result in explicit forms and tractable solutions, while the real life data are more explained by the non-Gaussian models. In this direction, many non-Gaussian models are available in the literature.

Nowadays, modelling of high frequency data in financial markets is of great interest. The information from trade durations, volume of the trade of a commodity and the price of each commodity can be utilized to study the movement of stock in the market. Engle and Russell (1998) proposed conditional duration models to analyse this type of irregularly spaced financial transaction data. Due to the availability of these intraday day price data, more sophisticated models are introduced to study such data. The complex structure and the intractable forms of the likelihood function motivated many researchers in proposing new estimation methods to explain the properties of the estimators in detail. In view of this, many researchers have contributed estimation methods to the existing literature.

Our main objective in this thesis is to study and develop non-Gaussian time series models and to device or utilise some estimation methods in analysing financial data.

1.2 Time series examples

Time series is a series of observations observed over a period of time 't'. Typically, the observations can be collected over an entire interval, randomly sampled on an interval or at fixed time points. Different types of sampling require different approaches to data analysis. The time series analysis is concerned with analysing and modelling the observations in order to extract the inherent information of the data. Let us discuss some examples where the data is recorded against time.

Example 1.1. *In the first example, we consider the GDP growth (annual %) of India obtained from World Bank website. The data consists of 56 observations from 1961 to 2016. The time series plot is shown in Figure 1.1.*

Example 1.2. *Here, we consider the monthly price of Coconut Oil (Philippines/Indonesia in Indian Rupee per Metric Ton). The data is obtained from World Bank website. The data consists of 240 observations from July 1997 to June 2017. The time series plot of the data is shown in Figure 1.2.*

Example 1.3. *In the third example, we consider daily maximum of BSE Index from May* 18^{th} 2006 *to June* 27^{st} 2007. *There are 277 observations. The plots of the actual time series is shown in Figure 1.3. (https://in.finance.yahoo.com)*

Some other examples are (1) monthly price of commodities, (2) currency exchange rate of two countries, (3) daily price of share indices, (4) inflation rate, (5) transaction data in security markets etc.

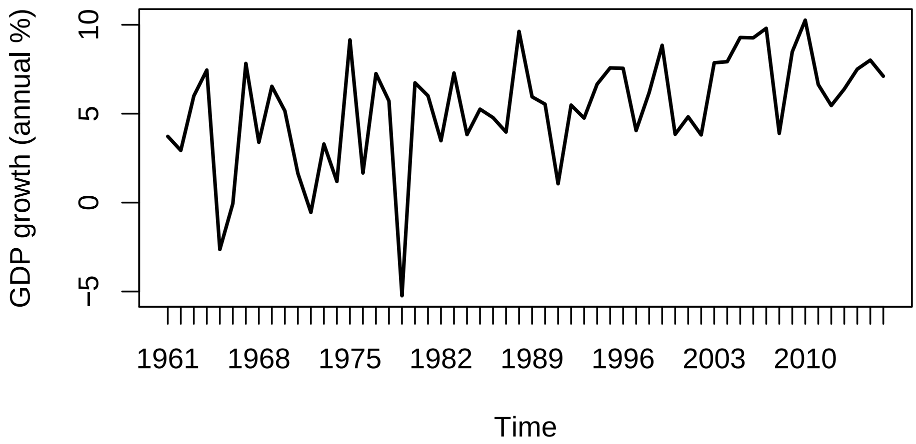

FIGURE 1.1: GDP growth (annual %) of India from 1961 to 2016.

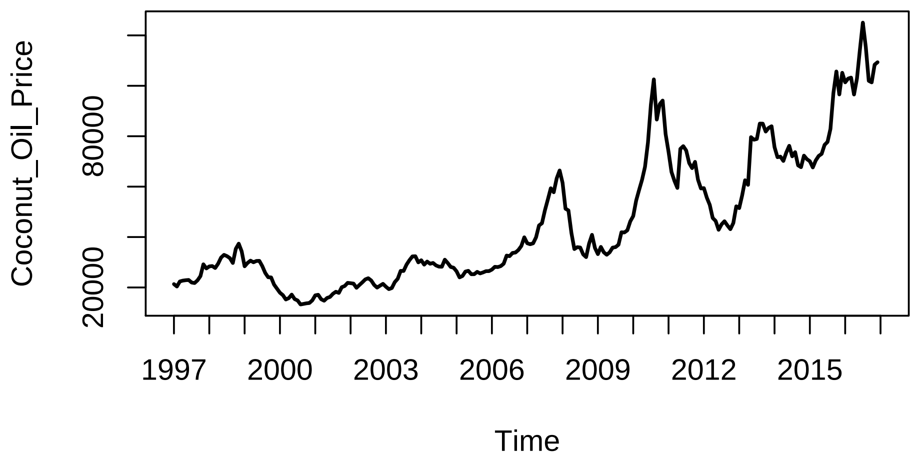

FIGURE 1.2: Monthly coconut oil price from July 1997 to June 2017

1.3 Basic Concepts

1.3.1 Stochastic Process

Let $(\Omega, \mathscr{F}, \mathscr{P})$ be a given probability space. A collection of random variables $\{Z_t, t \geq 0\}$ defined on the probability space $(\Omega, \mathscr{F}, \mathscr{P})$ is called a stochastic process.

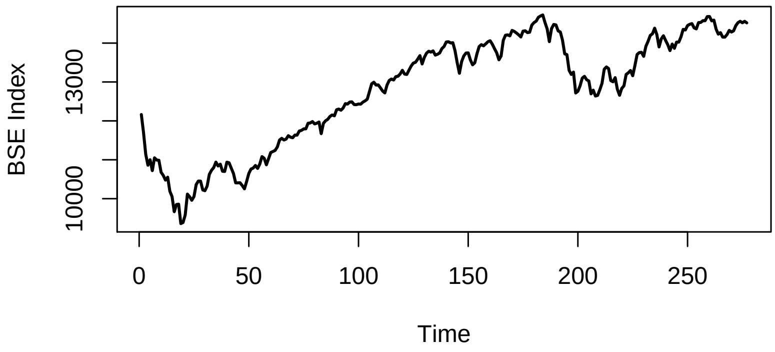

FIGURE 1.3: Time series plot of daily maximum of BSE Index values

In other words, a stochastic process is a collection $\{Z_t, t \in T\}$ of random variables Z_t, T being some indexed set, usually an interval of real numbers. An observed time series $\{z_1, z_2, \ldots z_n\}$ can be thought of as a particular realization function from a certain stochastic process. Let $\{Z_t, t \in T\}$ be a stochastic process. For a fixed $\omega : Z_t(\omega)$ is a function on T, called a sample function or realization of the process. It is usually written as $Z(t)$ or Z_t. The mean function and variance function of the process are defined as $\mu_t = E(Z_t)$ and ${\sigma_t}^2 = V(Z_t) = E(Z_t - \mu_t)^2$, where $V(.)$ is the variance function of the process.

To analyse time series data, we need to find suitable mathematical model for the data. The model is utilised to study the nature of the data. As the observations are possibly unpredictable, the series $\{z_t, t = 1, 2, \ldots, n\}$ can be thought of as a particular realization from a certain stochastic process.

1.3.2 White noise process

The process $\{a_t\}$ is said to be white noise with mean 0 and variance σ^2, written $\{a_t\} \sim WN(0, \sigma^2)$, if and only if $\{a_t\}$ has zero mean and covariance function

$$\nu(k) = \begin{cases} \sigma^2 \quad if \quad k = 0 \\ 0 \quad if \quad k = 0 \end{cases} .$$

If the random variable a_t are independent and identically distributed (iid) with mean zero and variance σ^2, then we shall write $\{a_t\} \sim iid(0, \sigma^2)$.

1.3.3 Gaussian Time Series

The process $\{a_t\}$ is a Gaussian time series if and only if the finite dimensional distribution function of $\{a_t\}$ are all multivariate normal.

1.3.4 Stationary Process

Stationarity is a kind of invariant property. It is a way to model the dependence structure. There are two important forms of stationarity, weak stationarity and strict stationarity.

1.3.4.1 Strict Stationarity

The time series $\{Z_t\}$ is said to be strictly stationary if we assume that the common distribution function of the stochastic process does not change by a shift in time, i.e., if the joint densities of $(Z_{t_1}, Z_{t_2}, \ldots, Z_{t_k})'$ and $(Z_{t_{1+h}}, Z_{t_{2+h}}, \ldots, Z_{t_{k+h}})'$ are the same for all positive integers k and for all $t_1, t_2, \ldots, t_k,\ h \in \mathbb{Z}$. It supposes the distributions are invariant over time. This is a very strong condition that is difficult to verify in practice. A weaker version of stationarity is frequently referred to in the literature as weak stationarity, covariance stationarity, stationarity in the wide sense or second order stationarity.

1.3.4.2 Second order Stationarity/Weak stationarity

The time series $\{Z_t\}$ with index set $\mathbb{Z} = 0, \pm 1, \pm 2, \ldots$ is said to be second order stationary if

1. $E|{Z_t}^2| < \infty$ for all $t \in \mathbb{Z}$,
2. $EZ_t = m$ for all $t \in \mathbb{Z}$ and
3. $\nu_Z(r, s) = \nu_Z(r + t, s + t)$ for all $r, s, t \in \mathbb{Z}$,

where $\nu_Z(r, s) = Cov(Z_r, Z_s)$. If $\{Z_t\}$ is strictly stationary and its first two moments are finite, then it is also weakly stationary, but the converse is not true. If $\{Z_t\}, t \in \mathbb{Z}$ is a weakly stationary Gaussian process then $\{Z_t\}$ is strictly stationary.

For a strictly stationary process, since the distribution function is same for all t, the mean function $E(Z_t) = E(Z_{t-k}) = \mu$ is a constant, provided $E(Z_t) < \infty$. Likewise, if $E(Z_t^2) < \infty$, then $Var(Z_t) = Var(Z_{t-k}) = \sigma^2$ for all t and hence is also a constant.

1.3.5 Autocovariance Function

If $\{Z_t\}$ is a process such that $Var(Z_t) < \infty$ for each $t \in T$, then the autocovariance function $\nu_Z(.,.)$ of $\{Z_t\}$ is defined by $\nu_Z(r,s) = Cov(Z_r, Z_s) = E[(Z_r - EZ_r)(Z_s - EZ_s)],\ r, s \in T$.

If $\{Z_t\}, t \in \mathbb{Z}$ is stationary, then $\nu_Z(r,s) = \nu_Z(r-s,0)$ for all $r, s \in \mathbb{Z}$. So the autocovariance function of a stationary process can be redefined as $\nu_Z(k) = \nu_Z(k,0) = Cov(Z_t, Z_{t-k})$ for all $t, k \in \mathbb{Z}$. The function $\nu_Z(.)$ will be referred to as the autocovariance of $\{Z_t\}$ and $\nu_Z(k)$ as its value at lag 'k'.

1.3.6 Autocorrelation Function

The correlation coefficient between Z_t and Z_{t-k} is called autocorrelation function (ACF) at lag 'k' and is given by

$$\rho_Z(k) = Corr(Z_t, Z_{t-k}) = \frac{Cov(Z_t, Z_{t-k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t-k})}}. \tag{1.1}$$

The ACF is a way to measure the linear relationship between an observation at time t and $t \pm k$.

1.3.7 Partial Autocorrelation Function

The Partial Autocorrelation Function (PACF) of a stationary process, $\{Z_t\}$, denoted by $\phi_{k,k}$ for $k = 1, 2, \ldots$ is defined by

$$\phi_{1,1} = Corr(Z_1, Z_0) = \rho_1$$

and

$$\phi_{k,k} = Corr(Z_k - \hat{Z}_k, Z_0 - \hat{Z}_0),\ \ k \geq 2,$$

where $\hat{Z}_k = l_1 Z_{k-1} + l_2 Z_{k-2} + \ldots + l_{k-1} Z_1$ is the linear predictor. Both $(Z_k, \hat{Z}_k)$ and $(Z_0, \hat{Z}_0)$ are correlated with $\{Z_1, Z_2, \ldots, Z_{k-1}\}$. By stationarity, the PACF, $\phi_{k,k}$ is the correlation between Z_t and Z_{t-k} obtained by fixing the effect of $Z_{t-1}, \ldots, Z_{t-(k-1)}$.

1.3.8 Decomposition of Time Series

An observed time series may exhibit time dependent factors such as trend and seasonality, as well as stochastic time dependence. The decomposition of a time series into various components can be achieved using suitable additive or multiplicative models. The additive structural decomposition of a given time series $\{W_t\}$ is given by

$$W_t = m_t + S_t + E_t, \qquad t = 1, 2, \ldots, n,$$

where m_t is a slowly changing function known as the trend component, S_t is a function with known period s referred to as seasonal component and $E_t = W_t - m_t - S_t$ is a random error component. Our aim is to estimate and extract the deterministic components m_t and S_t. After eliminating the time-dependent factors, the residuals or the noise component $\{E_t\}$ may be independent and identically distributed or stationary. These residuals (random components) are treated as a realization of some stationary stochastic process.

1.4 Linear Time Series

A time series Z_t is said to be linear if it can be written as

$$Z_t = \mu + \sum_{i=0}^{\infty} \varphi_i a_{t-i},$$

where μ is the mean of Z_t, $\varphi_0 = 1$, and $\{a_t\}$ is a sequence of iid random variables with mean 0 and variance σ^2 i.e., $\{a_t\}$ is a white noise series. It is referred to as the innovation or shock at time t. The dynamic structure of Z_t is governed by the coefficients φ_i, which are called the $\varphi-$weights of Z_t in the time series literature. These models are econometric and statistical models employed to describe the pattern of the $\varphi-$weights of Z_t. These models include Autoregressive, Moving Average, Autoregressive Moving Average, Autoregressive Integrated Moving Average etc.

1.4.1 Autoregressive Models

The time series $\{Z_t\}$ is said to be an autoregressive process of order p, $(AR(p))$ if it satisfies the equation

$$Z_t - \phi_1 Z_{t-1} - \ldots - \phi_p Z_{t-p} = a_t, \ \ t \in \mathbb{Z}, \tag{1.2}$$

where a_t is a sequence of uncorrelated random variable with zero mean and finite variance termed as innovations. In terms of backshift operators

$$Z_t - \phi_1 B Z_t - \ldots - \phi_p B^p Z_t = a_t,$$

$\Rightarrow \Phi(B)Z_t = a_t$ where $\Phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$, B is the backshift operator and is defined such that $B^k Z_t = Z_{t-k}$ and $\Phi(B)$ is referred to as the characteristic polynomial associated with an $AR(p)$ process. The process $\{Z_t\}$ is a linear function of its own past values. The resulting $AR(p)$ process is weakly stationary if the roots of the characteristic equation $\Phi(B) = 0$ lie outside the unit circle. For a stationary $AR(p)$ process, the autocorrelation function, $\rho_Z(k)$, is obtained by solving a set of difference equations called the Yule-Walker equations given by

$$(1 - \phi_1 B - \phi_2 B^2 \ldots \phi_p B^p)\rho_Z(k) = 0, \quad k > 0.$$

In particular, the autoregressive process of first order (p=1) is of practical importance. The first order autoregressive process, $AR(1)$ is defined by

$$Z_t = \phi Z_{t-1} + a_t, \tag{1.3}$$

where $\{a_t\} \sim WN(0, \sigma^2)$ and ϕ satisfies the condition $-1 < \phi < 1$ for the process to be weakly stationary. The model (1.3) is like the regression model in which the deviation from the mean at time t is regressed on itself, but with a lag of one time period. Under stationarity, we have $E(Z_t) = 0$, $V(Z_t) = \sigma^2/(1 - \phi^2)$ and the autocorrelation function is given by $\rho_Z(k) = \phi^k$, $k = 0, 1, \ldots$.The ACF of a weakly stationary $AR(1)$ series decays exponentially to zero with ϕ is positive, but decays exponentially to zero and oscillates in sign when ϕ is negative. If we assume that the innovation sequence $\{a_t\}$ is iid, then the $AR(1)$ sequence is Markovian. The PACF of Z_t is

$$\phi_{11} = \rho_1 = \phi \qquad \textit{and} \qquad \phi_{kk} = 0 \quad \text{for } k > 1.$$

1.4.2 Moving Average Model

In this model the observed time series is represented as a finite moving average process. A moving average model of order q ($MA(q)$) is defined by

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}, \tag{1.4}$$

or $Z_t = \Theta(B)a_t$, where $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots \theta_q B^q$, is the characteristic polynomial associated with the $MA(q)$ model, where $\theta_i' s$ are constant, $\{a_t\} \sim WN(0, \sigma^2)$. In this model, the observation at time t, Z_t is expressed as a linear function of the present and past shocks. MA models are always weak stationary because they are finite linear combinations of a white noise sequence for which the first two moments are time invariant. From (1.4) $E(Z_t) = 0$, $V(Z_t) = \sigma^2 \sum_{j=1}^{q} \theta_j^2$ and the ACF is,

$$\rho_Z(k) = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \ldots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2}, & k = 1, 2, \ldots q \\ 0, & k > q \end{cases}. \tag{1.5}$$

The ACF for a $MA(q)$ model vanishes after lag 'q'. When $q = 1$, (1.4) reduces to an $MA(1)$ model. The form of an $MA(1)$ model is

$$Z_t = a_t - \theta a_{t-1}, \quad \{a_t\} \sim WN(0, \sigma^2) \tag{1.6}$$

with the condition $-1 < \theta < 1$ for the process to be invertible. The unconditional variance of an $MA(1)$ process is given by $V(Z_t) = (1 + \theta^2)\sigma^2$. The ACF of the $MA(1)$ process is

$$\rho_Z(k) = \begin{cases} \frac{-\theta}{1+\theta^2}, & k = 1 \\ 0, & k > 1 \end{cases}. \tag{1.7}$$

and the PACF is given by

$$\phi_{kk} = \frac{-\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}}, \qquad k \geq 1. \tag{1.8}$$

The PACF of an $MA(1)$ model decays to zero exponentially.

1.4.3 Autoregressive Moving Average Models

A useful generalization of the pure autoregressive and pure moving average is the mixed autoregressive moving average (ARMA) process. It contains the ideas of AR and MA models into a compact form so that the number of parameters used is kept small, achieving parsimony in parametrization. An $ARMA(p, q)$ model is represented as

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \ldots - \theta_q a_{t-q}, \tag{1.9}$$

that is

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) Z_t = (1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q) a_t$$

or

$$\Phi(B) Z_t = \Theta(B) a_t, \tag{1.10}$$

where $\Phi(B)$ and $\Theta(B)$ are polynomials of degree p and q in B. The model has p AR terms and q MA terms. The model is stationary if AR(p) component is stationary and invertible if MA(q) component is so. Letting $p = q = 1$, we have the simplest and important example of an autoregressive moving average process, the $ARMA(1,1)$ process is given by,

$$Z_t - \phi Z_{t-1} = a_t - \theta a_{t-1}. \tag{1.11}$$

The process is stationary if $-1 < \phi < 1$, and invertible if $-1 < \theta < 1$. The mean and variance of $ARMA(1,1)$ model is $E(Z_t) = 0$ and $Var(Z_t) = \nu_0 = E(Z_t^2)$. The autocorrelation function is given by

$$\rho_Z(k) = \frac{(1 - \theta\phi)(\phi - \theta)}{1 - 2\theta\phi + \theta^2} \phi^{k-1} \text{ for} k \geq 1. \tag{1.12}$$

The autocorrelation function decays exponentially as the lag k increases . The damping factor starts from initial value ρ_1, which also depends on θ. One may refer Box et al. (1994) for detailed analysis of linear time series.

1.4.4 Autoregressive Integrated Moving Average Models

Many empirical time series do not have homogeneous stationary behaviour. A time series Z_t is said to follow an Autoregressive Integrated Moving Average Model (ARIMA) model if the d^{th} difference $W_t = \bigtriangledown^d Z_t = (1 - B)^d Z_t$ is a stationary ARMA process. Since W_t is then a stationary process, ARMA models can be used to describe W_t. The corresponding model can be written as

$$\Phi(B)(1 - B)^d Z_t = \Theta(B)a_t, \tag{1.13}$$

where $\Phi(B)$, $\Theta(B)$ and $\{a_t\}$ are defined as in (1.10). The model (1.13) is called the autoregressive integrated moving average model of order (p, d, q) and is denoted by $ARIMA(p, d, q)$.

1.5 Box Jenkins Modelling Techniques

The Box-Jenkins methodology uses ARMA or ARIMA models to model the data. It is achieved by a three stage iterative procedure based on identification, estimation and diagnostic checking. The first step is to identify a suitable model that may fit the time series data. In the second stage, the unknown parameters in the model are estimated. The third stage is to check for the model adequacy. If the model fails the diagnostic checks, the appropriate model modification can be made by repeating the first two stages. We shall explain each stages in more detail.

1.5.1 Identification

In this stage, by using the data, or the information regarding the data generation process, is utilized to identify and build a model. One of the principle tool for identifying the model is the ACF and PACF plots. The sample ACF plot and PACF plot are compared to theoretical behaviour of these plots when the order is known. These are estimated from the data. The autocorrelation function of an autoregressive process of order p tails off, its partial autocorrelation function has a cut off after lag p. Conversely, the autocorrelation function of moving average process of order q has a cut off after lag q, while its partial autocorrelation tails off. A mixed process is suggested when both the autocorrelation and partial autocorrelation tail off. The autocorrelation function for a mixed process, containing a $p-th$ order autoregressive component and a $q-th$ order moving average component, is a mixture of exponentials and damped sine waves after the first $q-p$ lags. The PACF function is dominated by a mixture of exponentials and damped sine waves after the first $p-q$ lags. Canonical correlation methods, AIC and BIC are the other approach to model selection. According to Akaike (1973) Information Criterion (AIC), one should select the model that minimizes

$$AIC = -2\ \log(\text{maximum likelihood}) + 2\ \text{k}, \tag{1.14}$$

where $k = p + q + 1$, if the model contains an intercept term and $k = p + q$ otherwise. Another approach to determine the ARMA orders is to select a model that minimizes the Schwarz (1978) Bayesian Information Criterion (BIC) defined

as

$$BIC = -2\ \log(\text{maximum likelihood}) + \text{k}\ \log(\text{n}), \tag{1.15}$$

where $k = p + q + 1$ and n is the (effective) sample size.

1.5.2 Parameter Estimation

One of the important aspect of time series analysis is to estimate the model parameters and hence various methods for estimating the parameters in a stationary time series are available in literature (see Box et al. (1994)). Among them, the main approaches to fitting Box-Jenkins models are non-linear least squares and maximum likelihood estimation. The least square (LS) estimator of a parameter is obtained by minimizing the residual sum of squares function. For pure AR models, the LS estimator leads to the linear ordinary least squares (OLS) estimators. If moving average components are present, the LS estimator becomes non-linear and we must resort to numerical optimization techniques. The Maximum Likelihood (ML) estimators are those values of the parameters that maximize the likelihood function. The advantage of the method of ML estimation is that, all of the information in the data is used rather than just the first and second moments, as is the case with least squares. Also, many large sample properties are known under general conditions for ML estimators. Some of the other methods for estimating the model parameters are Method of Moments (MM) and Generalized Method of Moments (GMM). For pure AR models Yule-Walker estimate is also another choice of estimation procedure when the ML estimation becomes difficult.

1.5.3 Daignostic Method

The third stage in Box-Jenkins approach is called model diagnostic checking. After estimating the parameters in a model, it is necessary to check whether the model assumptions are satisfied. It involves techniques like over-fitting, residual plots and checking that the residuals are approximately uncorrelated. If the assumptions are not met, the model should be respecified. A good model should be able to produce residuals that are approximately uncorrelated, that is, the autocorrelations of the residuals should be close to being uncorrelated after taking into account the effect of estimation. If the model is correctly specified, and the parameters are reasonably close to the true values, then the residuals should have nearly the properties of a white noise which gives the interpretation that the chosen model extracts almost all of the information from the data. To check on the correlation of the noise terms in the model, we consider the sample autocorrelation function of the residuals. The asymptotic distribution of the residual autocorrelation play a central role. The key reference on the distribution of residual autocorrelation in ARIMA model is Box and Pierce (1970), the results of which were generalized in McLeod (1978). From the asymptotic distribution of the residual autocorrelation we can derive tests for individual residual autocorrelation and overall tests for an entire group of residual autocorrelation assuming that the model is adequate. These overall tests are often called portmanteau tests. The main idea underlying these portmanteau tests is to identify if there is any dependence structure which is yet unexplained by the fitted model.

Ljung-Box test

Ljung and Box (1978) propose the portmanteau statistic

$$Q(m) = n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}(k)^2}{n-k},$$

as a test statistic for the null hypothesis $H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$ against the alternative hypothesis $H_a : \rho_i \neq 0$ for some $i \in \{1, \ldots, m\}$. The decision rule is to reject H_0 if $Q(m) > \chi^2_\alpha$, where χ^2_α denotes the $100(1-\alpha)th$ percentile of a chi-squared distribution with m degrees of freedom.

1.5.4 Forecasting

One of the primary objective of building a model for a time series is to be able to forecast the values for that series at future time. To assess the precision of those forecast is also of equal importance. For a linear time series model, Minimum Mean Square Error (MMSE) forecasting method is widely used. To derive MMSE, we first consider the stationary ARMA model,

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \ldots \theta_q a_{t-q}$$

or

$$\Phi(B)(1-B)^d Z_t = \Theta(B)a_t.$$

In terms of moving average

$$Z_t = \frac{\Theta(B)}{\Phi(B)}a_t = \Psi(B)a_t = \sum_{j=0}^{\infty}\varphi_j B^j a_t = a_t + \varphi_1 a_{t-1} + \varphi_2 a_{t-2} + \ldots \tag{1.16}$$

with $\varphi_0 = 1$. For $t = n + l$, we have

$$Z_{n+l} = \sum_{j=0}^{\infty} \varphi_j a_{t+l-j}. \tag{1.17}$$

Suppose at time $t = n$, we have the observations $Z_n, Z_{n-1}, Z_{n-2}, \dots$ and wish to forecast $l - step$ ahead value, Z_{n+l}, as a linear combination of the observations $Z_n, Z_{n-1}, Z_{n-2}, \dots$. Since Z_t for $t = n, n-1, n-2, \dots$ can all be written in the form of (1.17), we can let the MMSE forecast $Z_n(l)$ of Z_{n+l} be

$$\hat{Z}_n(l) = \varphi_1^* a_n + \varphi_{l+1}^* a_{n-1} + \varphi_{l+2}^* a_{n-2} + \dots,$$

where the φ_j^* are to be determined. The mean square error of the forecast is

$$E(Z_{n+l} - \hat{Z}_n(l))^2 = \sigma^2 \sum_{j=0}^{l-1} \varphi_j^2 + \sigma^2 \sum_{j=0}^{\infty} (\varphi_{l+1} - \varphi_{l+j}^*)^2,$$

which is seen to be minimized when $\varphi_{l+j}^* = \varphi_{l+j}$. Hence,

$$\hat{Z}_n(l) = \varphi_l a_n + \varphi_{l+1} a_{n-1} + \varphi_{l+2} a_{n-2} + \dots \ .$$

But using (1.17) and the fact that

$$E(a_{n+j}|Z_n, Z_{n-1}, \dots) = \begin{cases} 0, & j > 0, \\ a_{n+j}, & j \le 0, \end{cases}$$

we have

$$E(Z_{n+l}|Z_n, Z_{n-1}, \dots) = \varphi_l a_n + \varphi_{l+1} a_{n-1} + \varphi_{l+2} a_{n-2} + \dots \ .$$

Thus, the minimum mean square error forecast of Z_{n+l} is given by its conditional expectation. That is, $\hat{Z}_n(l) = E(Z_{n+l}|Z_n, Z_{n-1}, \ldots)$. $\hat{Z}_n(l)$ is usually read as the $l-step$ ahead of the forecast of Z_{n+l} at the forecast origin n. The forecast error is

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l).$$

1.5.5 Illustration of Box-Jenkins Methodology

Here, we shall analyse and model the examples discussed in Section 1.2.

Example 1: The GDP data is taken from world bank website https://data.worldbank.org. As the data is not stationarity, we consider the differenced series and the plot is shown in Figure 1.4. The stationarity of the data is confirmed from Box.test() function in R which gives the p-value$<$ 0.01. The ACF and PACF plots are shown in Figure 1.5.

From ACF and PACF plots we suggest a MA(1) model for the series. The least square fit of the model is

$$\hat{Z}_t = \hat{a}_t - 0.8627\, a_{t-1}.$$

To check the adequacy of the model, we plot the ACF of residuals. Also, we perform the Box.test() function in R program and obtained the p-value as 0.25 with Ljung-Box statistic value 1.336, which is less than the Chi-square critical value 10.117 at degrees of freedom 20, which does not reject the null hypothesis of independence in the series. To check the normality of the innovations, the Anderson-Darling (AD) normality test is performed with 'nortest' package in R and obtained the p-value as

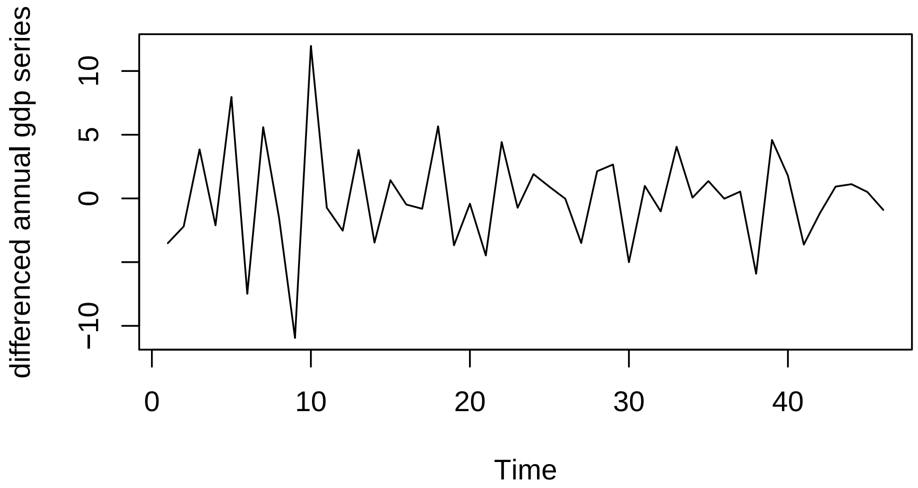

FIGURE 1.4: Stationary series of annual gdp

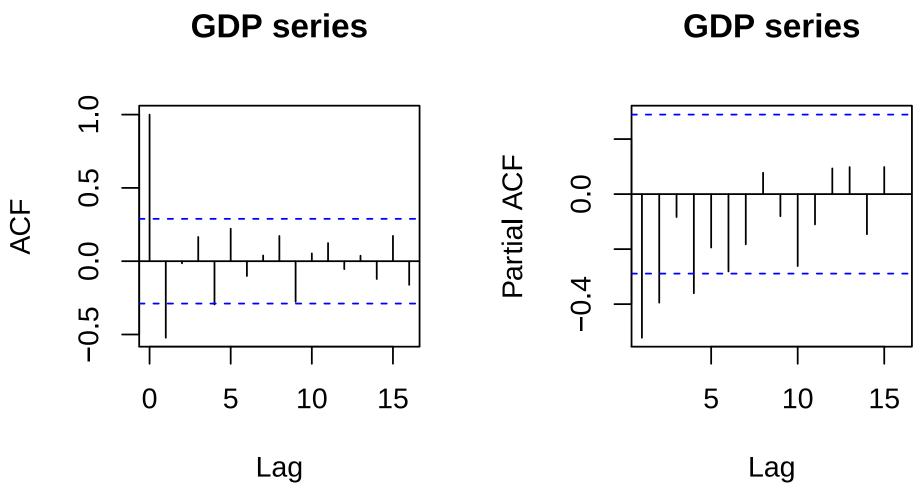

FIGURE 1.5: ACF and PACF plot for differenced annual gdp growth series

0.4789. This AD test performs the test for the composite hypothesis of normality. The Q-Q plot is given in Figure 1.6. From the p-value of AD test and Figure 1.6 the normality assumptions are validated. So the MA(1) model fits the data.

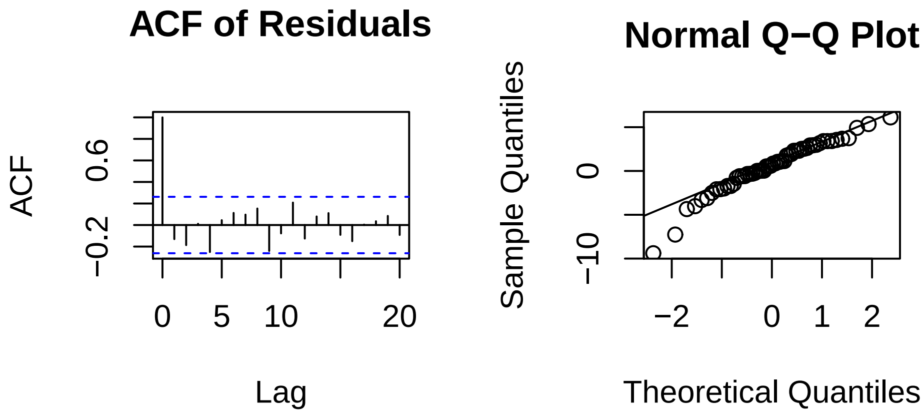

FIGURE 1.6: ACF and Q-Q plot of residuals for annual gdp growth series

Example 2: This example is the monthly price (Indian Rupee per Metric Ton) of Coconut Oil (Philippines/Indonesia) which consists of 240 observations. As a first step, we check the stationarity of the data. As the data is not stationary we consider the first differenced series. The time series plot of the first differenced series is given in Figure 1.7. The stationarity is tested with Augmented Dickey Fuller (ADF) test and the differenced series is stationary with p-value < 0.01. Now, to identify a model we plot the ACF and PACF of the data which is shown in Figure 1.8.

From the ACF and PACF we can see a significant spike in the PACF and a small correlation in the ACF. So, an AR(1), MA(1) or ARMA(1,1) may fit the differenced data. We have tested ARMA(p,q) models of orders $0 \leq p \leq 4, 0 \leq q \leq 4$. To compare these models we consider the AIC values. The resulting values are given in Table 1.1. From the Table 1.1 and Figure 1.8 we conclude that, an ARMA(1,1) model is of good fit, as ARMA(1,1) is having the lowest AIC value. The maximum likelihood estimates of the parameters are obtained as $\hat{\phi}_1 = -0.7077$ and $\hat{\theta}_1 =$

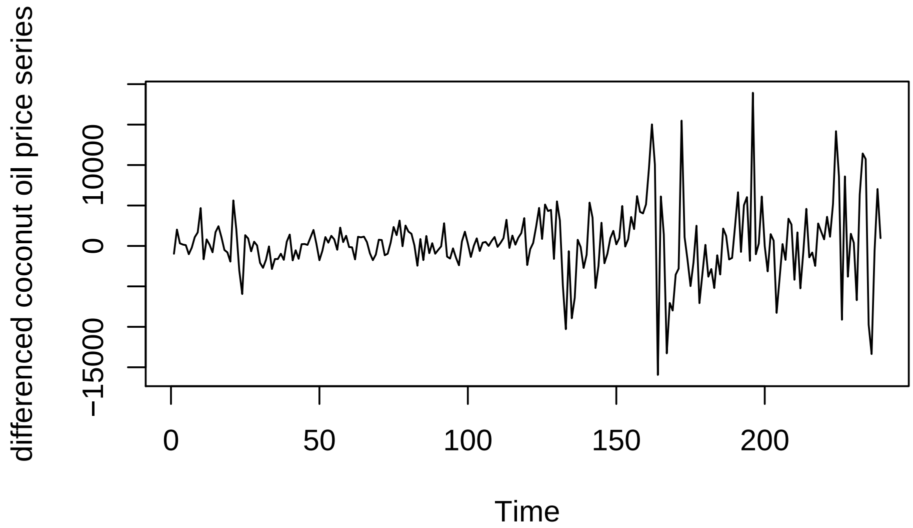

FIGURE 1.7: Stationary series of differenced monthly coconut oil price from July 1997 to June 2017

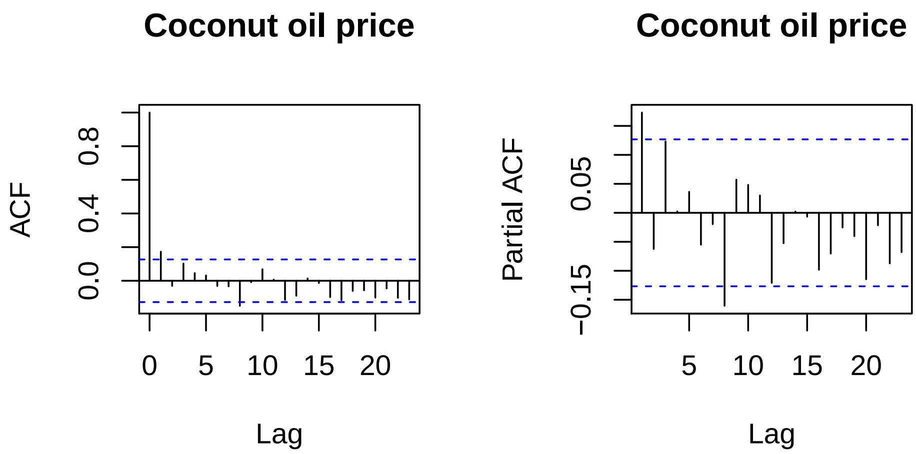

FIGURE 1.8: ACF and PACF plot for differenced monthly coconut oil price

0.8733. Thus the fitted model is

$$\hat{Z}_t = 365.07 + \hat{Z}_{t-1} - 0.7077\,\hat{Z}_{t-1} + \hat{a}_t - 0.8733\,a_{t-1}.$$

To check the adequacy of the model, the model diagnosis are done. We have plotted the residuals of ACF in Figure 1.9, to check the autocorrelation in residuals . In Figure 1.9 the ACF shows no significant correlation, which implies that the residuals are uncorrelated. Also, the p-value obtained from the Ljung-Box test is 0.5097, which does not reject the null hypothesis of independence in the series. Now, to check the normality assumption we consider the Q-Q plot given in Figure 1.9. We also perform the AD test and obtain the p-value$<$ 0.0001, which shows that the assumption of normality is not justified.

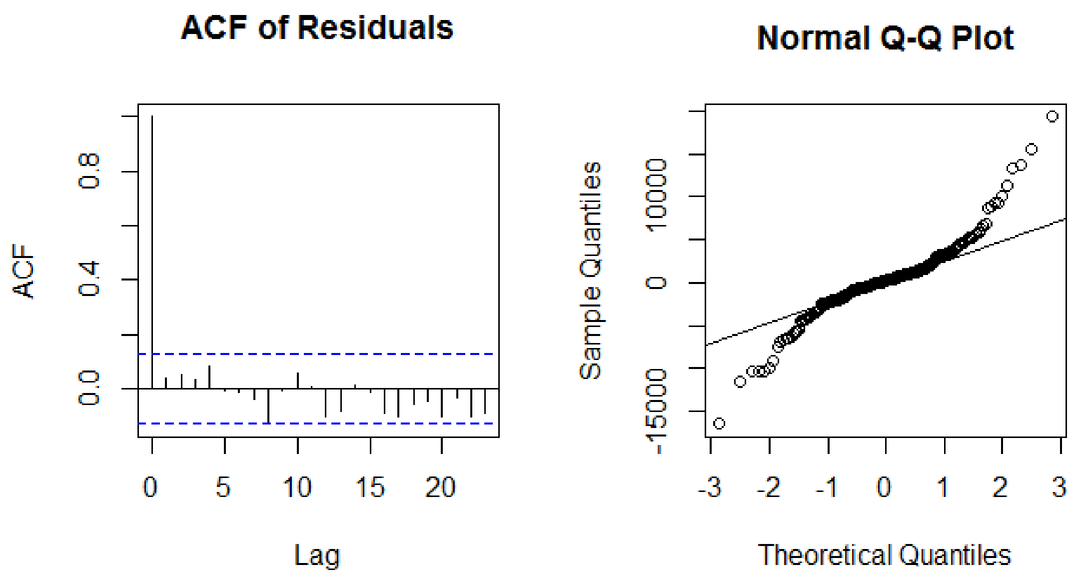

FIGURE 1.9: ACF and Q-Q plot of residuals for coconut Oil price data

TABLE 1.1: AIC of fitted models for differenced coconut oil price data

Order(p,q) $q \longrightarrow, p \downarrow$	0	1	2	3	4
0	AIC	4663.814	4664.643	4663.755	4665.650
1	4664.950	**4661.486**	4661.859	4665.313	4667.311
2	4666.017	4661.869	4663.547	4665.480	4664.233
3	4664.352	4666.344	4662.127	4656.952	4656.592
4	4666.348	4668.352	4666.263	4656.891	4658.540

Here, one of the assumptions of the innovation series is not satisfied. So to model this data or similar data sets, either some transformation to the data should be performed or a non-Gaussian time series model should be introduced. Also, one can utilise the available non-Gaussian time series models mentioned in the literature.

1.6 Non- Gaussian time series

The Box-Jenkins methodology of time series analysis focusses on stationary ARMA models with Gaussian innovations. However, many real life situations are better explained by non-Gaussian distributions. From Example 2, we can see the existence of such type of data. While modelling these type of data, the usual practice is to make some transformations to data, so that the changes results in normal distribution. But, in many cases, the transformation method results in poor results (cf. Lawrance (1991)). So the characteristics of the data is studied by non-Gaussian time series models. In view of this, many non-Gaussian models are developed during the last four decades. This is evident from the studies of H. L. Nelson and Granger (1979), Weiss (1977) and Yakowitz (1973). In the case of Gaussian models both the series $\{Z_t\}$ and the innovations $\{a_t\}$ have normal distributions whereas it is not the case in non-Gaussian models. As the literature of the non-Gaussian models are distribution specific, two approaches are studied in detail. Either to construct a model which have predesignated marginal distribution or to consider a time series model with pre-specified distribution for innovation. Some of the examples of former class of models are Exponential autoregressive and Gamma autoregressive models of Gaver and Lewis (1980) and Lawrance and Lewis (1985), Laplace AR models of Dewald and Lewis (1985), Linnik AR models of Anderson and Arnold (1993), Mittag-Leffler AR model of Jayakumar and Pillai (1993), Inverse Gaussian model of Abraham and Balakrishna (1999), Cauchy model of Balakrishna and Nampoothiri (2003), Gumbel extreme value model of Balakrishna and Shiji (2014) and the examples for latter class are non-Gaussian ARMA model of Li and McLeod (1988), first order autoregressive model with exponential innovations of Andel (1988), and the models proposed by Bell and Smith (1986), Tiku et al. (2000), Hughes et al. (2007), Hurlimann (2012), Bondon (2009) and so on.

1.7 Outline of the thesis

The studies on financial time series reveals the purpose to introduce new class of linear models. Chapter 2 discusses the specific literature of financial time series. The models for financial time series can be broadly classified as observation driven and parametric driven models. In observation driven models, the volatility (conditional variance) is assumed to be a function of the past observations, which introduces the heteroscedasticity in the models. The autoregressive conditional heteroscedastic (ARCH) model of Engle (1982) and Generalized ARCH (GARCH) models of Bollerslev (1986) are examples of these. The parameter driven models assume that the volatilities are generated by some latent models, in terms of unobservable variables. These models are referred to as Stochastic Volatility (SV) models. The Log-normal Stochastic Volatility model of Taylor (1986) is one of the examples of parameter driven model. Then we discuss the concept of financial duration and different types of financial duration processes. The statistical properties of financial durations and the type of models used in modelling the duration are also discussed. One of the difficulties in these duration models is to estimate the parameters. Also, we discuss some of the available estimation procedures in analysing the duration models.

Generalized Error distribution (GED) is a natural generalization of the normal distribution. As particular cases it includes the normal and Laplace distributions. As the distribution has leptokurtic and mesokurtic forms for its different values of its shape parameter it is used in the financial studies. This distribution was introduced by Subbotin (1923) and was later used in robustness studies by Box and Tiao (1962). In Chapter 3, we propose an ARMA model with GED innovations. Firstly, we discuss the model and its properties. We estimate the parameters of the model using conditional maximum likelihood and Generalised method of moments. Thus the resulting estimators are shown to be consistent and asymptotically normal. In the next Section, we carry out the simulation studies. As the last Section of the Chapter, the applicability of the proposed model is illustrated using BSE 500 data.

The method involved in the parameter estimation for a specified model is very important as it affects the efficiency of the model. In AR(1) model, the least square estimators is chosen as an estimate of the coefficient of autocorrelation, which shows less efficiency when the distribution under study is heavy tailed or asymmetric. In Chapter 4, we propose Hurwicz estimator for the coefficient of a first order autoregressive process with GED innovations. We study the performance of the Hurwicz estimator and least square estimator through a Monte Carlo simulation study. Finally, the prediction interval of the one-step ahead observation of the autoregressive process is considered. Also, some simulation studies are employed to investigate the performance of the two estimators in predicting the future observation.

In the study of financial time series, modelling the non-negative random variables is of great importance to understand the evolution of conditional variances. The changes in volatility (conditional variance) over time can be analysed by utilising the statistical models for time-dependent variances. In these models, the time-dependent variances are random variables generated by an underlying stochastic process. One among them is the lognormal stochastic volatility model introduced by Taylor (1986). In Chapter 5, we propose a Lindley SV model. Here, the volatilities are generated by Lindley first order autoregressive process. The Lindley-SV model and the second order properties of the model are studied. The parameter estimation is carried out using the method of moments. The applicability of the model is illustrated by analysing financial data set.

The empirical analysis of durations between the occurrences of certain financial events is important in understanding the market behaviour. In Chapter 6, we propose a Lindley Autoregressive Conditional Duration (ACD) model. The increasing nature of the hazard function of Lindley ACD model makes it as an alternative distribution in modelling duration data. The parameter estimation of the proposed model is carried out by conditional maximum likelihood method. The performance of the estimation method are studied through simulation methods. Also, we study the applicability of the Lindley ACD model in modelling real data.

In Chapter 7, we give a review on the existing literature of the Stochastic Conditional Duration (SCD) models. One of difficulties in studying these models is in estimating the parameters. So the parameter estimation procedures are analysed and further, we propose Bayesian Monte Carlo Markov Chain (MCMC) methods to estimate the parameters of an Inverse Gaussian SCD model. Simulation studies are conducted to analyse the behaviour of the proposed method. In the last Section, the applicability of model is illustrated through IBM trades data and Brunt Crude Oil data.

In Chapter 8, the major conclusions of the works in the thesis are presented. Also, we briefly discuss some related problems as future study.

Chapter 2

Models for Financial Time Series

2.1 Introduction

One of the basic assumptions in the classical time series is the assumption of constant variance with respect to time. But many of the financial time series usually exhibit the characteristic feature that the variances change with time. Here, the main objective is to model the volatility and forecast its future value. The key feature that distinguishes financial time series is that both the financial theory and its empirical time series contain an element of uncertainty. In view of this, a large number of non-linear and non-Gaussian time series models are introduced in the literature, see Tong et al. (1995), Tsay (2005) and the references therein. Another objective of analysing financial time series is to study and model the behaviour of the market using statistical techniques.

Due to advances in information technology, the observations recorded daily or at time scale have gained importance. The study and modelling of these high frequency data are important in empirical study of market infrastructure and high frequency trading. These models which are used to analyse the time interval between each trade are called the duration models. The study of these models helps to understand the dynamic behaviour of durations.

In the next Section of this Chapter we briefly discuss the characteristics of financial time series and the important tools useful in their analysis. In Section 2.3 we discuss about the volatility models in finance. The conditional duration models are explained in Section 2.4. An outline of the Markov Chain Monte Carlo method and Metropolis-Hastings (MH) algorithm are discussed in Section 2.5 .

2.2 Characteristics of Financial Time Series

2.2.1 Asset returns

Asset is a generic term used for the products being traded. According to Campbell et al. (1997) the financial studies involve returns instead of prices. It is because the returns of an asset is complete and scale free summary of the investment opportunists and the return series are more easy to handle. For a given series of prices $\{P_t\}$, the corresponding series of returns is defined by

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1, \quad t = 1, 2, \ldots,$$

and the log-return series is defined by

$$y_t = log\left(\frac{P_t}{P_{t-1}}\right), \; t = 1, 2, \ldots\cdot$$

Many empirical studies (see Bollerslev et al. (1992), Mandelbrot (1963), Fama (1965), Shephard (1996)) show that a set of common features among financial data that are known as 'stylized facts'. Cont (2001) provides a comprehensive survey of these stylized facts, which include :

1. Fat tails : the unconditional distribution of returns have fatter tails than that expected from a normal distribution.

2. Asymmetry : the unconditional distribution is negatively skewed.

3. Aggregated normality.

4. Absence of serial correlation in $\{y_t\}$.

5. Volatility clustering: volatility of returns is serially correlated.

6. Time varying cross-correlation.

Example 2.1. *To illustrate the stylized facts we consider the data of log-returns on Nikkei 225 stock market index for the Tokyo Stock Exchange (TSE). The sample period is from 18 November 2010 to 17 November 2017. Figure 2.1 displays the log-returns on Nikkei stock market index and 2.2 displays the histogram of the log-returns.*

The summary of the data is given in Table 2.1.

TABLE 2.1: Summary statistics of the Nikkei 225 return series

Mean	0.008705
Median	0.013247
Std Deviation	1.122376
Skewness	-0.09373
Kurtosis	13.97556
Minimum	-16.1354
Maximum	8.891297

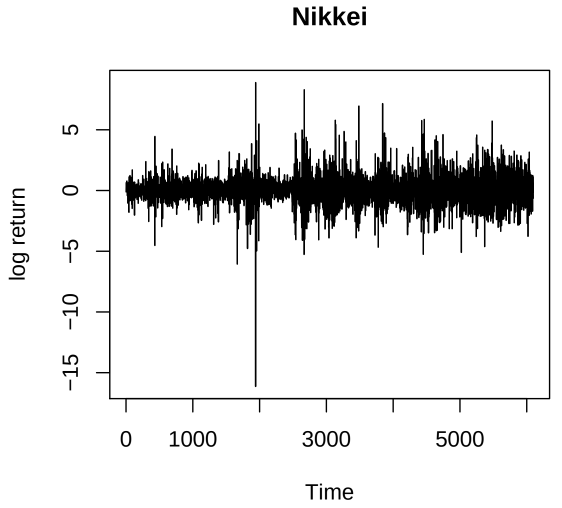

FIGURE 2.1: Evolution of daily log-returns of Nikkei 225 index

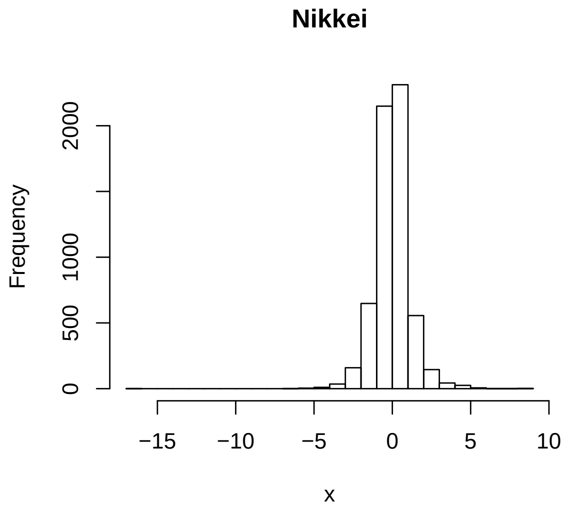

FIGURE 2.2: Histogram of daily log-returns of Nikkei 225 index

From the Table 2.1 and the Figure 2.1 and 2.2, the persistence of stylized facts in financial data can be easily understood.

The linear time series models are useful when we have the assumption of constant error variance. But in real life situation, particularly the information recorded in economics and financial areas, the conditional variances are non-constants. From Example 2.1, we can observe these characteristics of financial data. So volatility models are introduced to model the financial market variables.

2.3 Models for Volatility

A volatility model describes the evolution of conditional variance . The dynamics of volatility can be essentially described by two types of models. In the first category, volatility is described as a deterministic function of a given set of past values, that is, the observation driven models. This category includes the Autoregressive Conditional Heteroscedastic model/ Generalized ARCH models. In the second category of model, volatility is generated by a stochastic model, that is, the parameter driven models. It includes Stochastic Volatility models.

2.3.1 Observation Driven Models

The observation driven models assume that the conditional variances are the functions of past values of the series. The famous autoregressive conditional heteroscedastic (ARCH) model introduced by Engle (1982) is an example of such models. The simplest form of ARCH model assumes that the conditional variance of y_t given the past is a linear function of the squares of the past data.

2.3.1.1 Autoregressive Conditional Heteroscedastic model

The first model that provides a systematic framework for volatility modelling is the ARCH model of Engle (1982). The basic idea of ARCH models is as follows :

- The shock y_t of an asset return is serially uncorrelated but dependent, and
- The dependence of y_t can be described by a simple quadratic function of its lagged values.

ARCH models have been widely used in financial time series analysis and particularly in analysing the risk of holding an asset, evaluating the price of an option, forecasting time-varying confidence intervals and obtaining more efficient estimators under the existence of heteroscedasticity. An ARCH(p) model assumes that

$$y_t = \varepsilon_t \sqrt{h_t}\,, \tag{2.1}$$

$$h_t = \phi_0 + \sum_{i=1}^{p} \phi_i \, y_{t-i}^2,$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed symmetric random variables with mean zero and variance 1, $\phi_0 > 0$, and $\phi_i \geq 0$ for $i > 0$. If $\{\varepsilon_t\}$ has standardized Gaussian distribution, y_t is conditionally normal with mean 0 and variance h_t. We can also assume more heavy-tailed distributions, such as the Student's t-distribution, Generalized error distribution etc for $\{\varepsilon_t\}$. Now we describe the properties of a first order ARCH model in detail.

ARCH(1) model and Properties

The structure of the ARCH(1) model implies that the conditional variance h_t of y_t, evolves according to the most recent realizations of y_{t-1}^2 analogous to an $AR(1)$ model. Large past squared shocks, $\{y_{t-i}^2\}_{i=1}^p$, imply a large conditional variance, h_t, for y_t. As a consequence, y_t tends to assume a large value which in turn implies that a large shock tends to be followed by another large shock. To understand the ARCH models, let us now take a closer look at the ARCH(1) model,

$$y_t = \varepsilon_t \sqrt{h_t}, \;\; h_t = \phi_0 + \phi_1 y_{t-1}^2, \tag{2.2}$$

where $\phi_0 > 0$ and $\phi_1 \geq 0$.

1. The unconditional mean of y_t is zero, since

$$E\left(y_t\right) = E\left(E\left(y_t | y_{t-1}\right)\right) = E\left(\sqrt{h_t}\, E\left(\varepsilon_t\right)\right) = 0.$$

2. The conditional variance of y_t is

$$E\left(y_t^2 | y_{t-1}\right) = E\left(h_t \varepsilon_t^2 | y_{t-1}\right) = h_t E\left(\varepsilon_t^2 | y_{t-1}\right) = h_t = \phi_0 + \phi_1\, y_{t-1}^2.$$

3. The unconditional variance of y_t is $V\left(y_t\right) = E(y_t^2) = E(E(y_t^2 | y_{t-1})) = E(\phi_0 + \phi_1 y_{t-1}^2) = \phi_0 + \phi_1 E(y_{t-1}^2)$.

   Because y_t is a stationary process with $E(y_t) = 0,\; V(y_t) = V(y_{t-1}) = E(y_{t-1}^2)$. Therefore, $V(y_t) = \phi_0/(1 - \phi_1)$. Because the variance of y_t must be positive, we require $0 \leq \phi_1 < 1$.

4. Assuming that the fourth moment of y_t is finite, the kurtosis K_y of y_t, is given by
$$K_y = \frac{E\left(y_t^4\right)}{E\left(y_t^2\right)^2} = 3\frac{1-\phi_1^2}{1-3\phi_1^2} > 3 \text{ provided } \phi^2 < 1/3.$$
The ARCH model with a conditionally normally distributed y_t leads to heavy tails in the unconditional distribution. In other words, the excess kurtosis of y_t is positive and the tail distribution of y_t is heavier than that of the normal distribution.

5. The autocovariance of y_t is defined by

$$Cov(y_t, y_{t-k}) = E(y_t y_{t-k}) - E(y_t)E(y_{t-k}) \tag{2.3}$$
$$= E(\sqrt{h_t}\sqrt{h_{t-k}})E(\varepsilon_t \varepsilon_{t-k}) = 0. \tag{2.4}$$

Thus, the ACF of y_t is zero. The ACF of $\{y_t^2\}$ becomes $\rho_{y_t^2} = \phi_1^k$, and notice that $\rho_{y_t^2}(k) \geq 0$ for all k, a result which is common to all linear ARCH models.

Thus, the ARCH(1) process has a mean of zero, a constant unconditional variance, and a time-varying conditional variance. These properties continue to hold for general ARCH models, but the formulae become more complicated for higher order ARCH models.

Estimation

The most commonly used estimation procedure for ARCH models is the method of maximum likelihood. Under the normality assumption for ε_t, the likelihood function of an ARCH(p) model is

$$f(y_1, y_2, ..., y_n|\Phi) = \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi h_t}} \exp(-\frac{y_t^2}{2h_t}) f(y_1, y_2, ..., y_p|\Phi),$$

where $\Phi = (\phi_0, \phi_1, \ldots, \phi_n)'$ and $f(y_1, y_2, ..., y_p|\Phi)$ is the joint probability density function of $y_1, y_2, ..., y_p$. Since the exact form of $f(y_1, y_2, ..., y_p|\Phi)$ is complicated, it is commonly dropped from the prior likelihood function, especially when the sample size is sufficiently large. This results in using the conditional-likelihood function

$$f(y_{p+1}, y_{p+2}, ..., y_n|\Phi, y_1, y_2, ..., y_p) = \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi h_t}} \exp(-\frac{y_t^2}{2h_t}). \tag{2.5}$$

Model diagnosis

After estimating the model parameters, the standardized residuals are calculated by

$$\widetilde{\varepsilon}_t = \frac{y_t}{\sqrt{\hat{h_t}}} \qquad t = 1, 2, 3, \ldots,$$

where $\hat{h_t} = \hat{\phi}_0 + \sum_{i=1}^{p} \hat{\phi}_i \, y_{t-i}^2$. For a properly specified ARCH(p) model, $\widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2, \ldots, \widetilde{\varepsilon}_n$ are iid. The adequacy of the fitted ARCH model is checked by examining the $\widetilde{\varepsilon}_t$ series. One can utilise the Ljung-Box test, Q-Q-plot and the skewness and kurtosis measures to check the validity of the assumptions.

Forecasting

An important use of ARCH models is the evaluation of the accuracy of volatility forecasts. In standard time series methodology which uses conditionally homoscedastic ARMA processes, the variance of the forecast error does not depend on the past information. If the series being forecasted displays ARCH effect, the current information set can indicate the accuracy by which the series can be forecasted. Engle and Kraft (1983) were the first to consider the effect of ARCH on forecasting. As the conditional variance is a linear function of the squares of the past observations, one can use the minimum mean square error (MMSE) method for forecasting the volatility as in the case of classical AR models.

Using the MMSE method, the l-step-ahead forecast for

$$h_n(l) = \phi_0 + \sum_{i=1}^{p} \phi_i h_n(l-i),$$

where $h_n(l-i) = y^2_{n+l-i}$ if $l-1 \leq 0$.

Weakness of ARCH models

1. The model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks. In practice, it is well known that the price of a financial asset responds differently to positive and negative shocks.

2. The ARCH model is rather restrictive. For instance, ${\phi_1}^2$ of an ARCH(1) model must be in the interval $[0, \frac{1}{3})$ if the series is to have a finite fourth moment. The

constraint becomes complicated for higher order ARCH models. In practice, it limits the ability of ARCH models with Gaussian innovations to capture excess kurtosis.

3. The ARCH model does not provide any new insight for understanding the source of variations of financial time series. It merely provides a mechanical way to describe the behaviour of conditional variance. It gives no indication about what causes such behaviour to occur.

4. ARCH models are likely to over-predict the volatility because they respond slowly to large isolated shocks to the return series.

2.3.1.2 Generalized ARCH Models

Although the ARCH model is simple, it often requires many parameters to adequately describe the volatility process of an asset return. Sometimes an ARCH(p) model, where p is of higher order may be needed for the volatility process. The possibility that estimated parameters in ARCH model do not satisfy the stationarity condition increases with lag. So an alternative model must be sought. Bollerslev (1986) proposes a useful extension known as the generalized ARCH (GARCH) model. That is, the GARCH model is an extension that allows the conditional variance to depend on the previous conditional variance and the squares of previous returns. The GARCH(p,q) model is defined by

$$y_t = \varepsilon_t \sqrt{h_t}, \tag{2.6}$$

$$h_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2,$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with mean zero, variance 1, $\phi_0 > 0$, $\phi_i \geq 0$, $\beta_j \geq 0$ and $\sum_{i=1}^{max(p,q)}(\phi_i + \beta_i) < 1$. The constraint on $\phi_i + \beta_i$ implies that the unconditional variance of y_t is finite, whereas its conditional variance h_t evolves over time.

The GARCH model has several extensions like log-GARCH, Exponential GARCH, Integrated GARCH, Fractionally Integrated GARCH, Threshold GARCH etc. The model parameters may be estimated by the method of conditional maximum likelihood, for details see Tsay (2014), Chapter 4.

2.3.2 Parameter Driven Models

The parameter driven models assume that the volatilities are generated by some latent models, such as the an autoregressive process.

2.3.2.1 Stochastic Volatility Models

One of the empirical findings of asset returns is its volatility. It varies randomly with time. This type of changes in volatility (conditional variance) over time can be analysed by utilising the statistical models for time-dependent variances. For these models the volatility depends on some unobserved components or a latent structure. One interpretation for latent h_t is to represent the random and uneven flow of new information, which is very difficult to model directly, into financial markets. The most popular of these parameter-driven stochastic volatility models, is the one by Taylor (1986)

$$y_t = \varepsilon_t \exp(h_t/2), \qquad h_t = \phi_0 + \phi_1 h_{t-1} + \eta_t, \tag{2.7}$$

where ε_t and η_t are two independent Gaussian white noises, with variances 1 and ${\sigma_\eta}^2$, respectively. Due to the Gaussianity of η_t, this model is called a log-normal SV model. The properties of this model are discussed in Taylor (1986, 1994).

2.3.3 Some examples of non-Gaussian distributions in modelling volatility

Let us consider an example to illustrate the applicability of some non-Gaussian models for financial data.

Students t-distribution

A rv X is said to follow a Student t distribution with ν degrees of freedom if its probability density function is given by

$$f(x|\nu) = \frac{\Gamma((\nu+1)-2)}{\Gamma(\nu-2)\sqrt{(\nu-2)\pi}}(1+\frac{x^2}{\nu-2})^{(\nu+1)-2} \qquad \nu > 2, \tag{2.8}$$

where $\Gamma(x)$ is the gamma function (i.e., $\Gamma(x) = \int\limits_0^\infty y^{x-1}\exp(-y)dy$).

Generalized Error Distribution

A rv X is said to follow a GED(μ, σ, β) if its probability density function is given by

$$f(x) = \frac{\beta}{2\,\sigma\,\Gamma(1/\beta)} \exp\left\{-\left(\left|\frac{(x-\mu)}{\sigma}\right|\right)^\beta\right\}\;, \quad -\infty < x < \infty$$

$$-\infty < \mu < \infty, \;\; \sigma > 0, \; \beta > 0. \tag{2.9}$$

For $\beta = 1$ the above density reduces to Laplace distribution and for $\beta = 2$, it takes the form of Normal distribution.

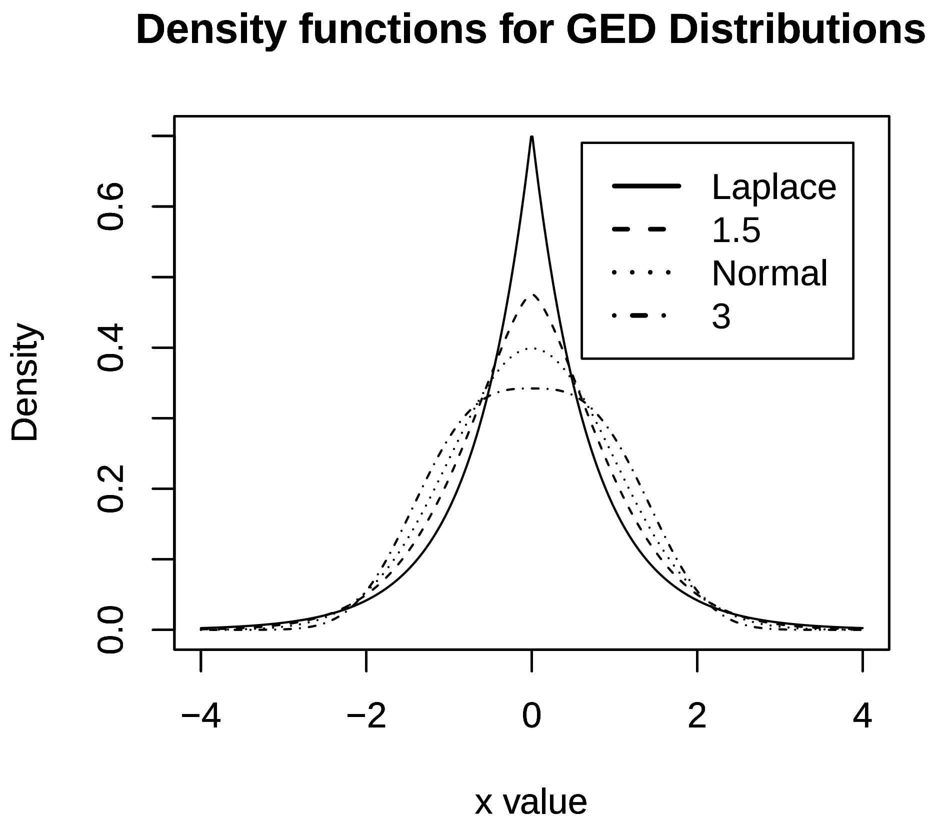

FIGURE 2.3: Density function for Generalized error distribution

Example 2.2. *We analyse the returns for the daily exchange rate between INR and US Dollar from January 1, 2015 to September 30, 2017. The time series plot of the data is given in Figure 2.4.*

The time plot of log-return for the daily exchange rate between INR and US Dollar is given in Figure 2.5. The summary of the data is given in Table 2.2.

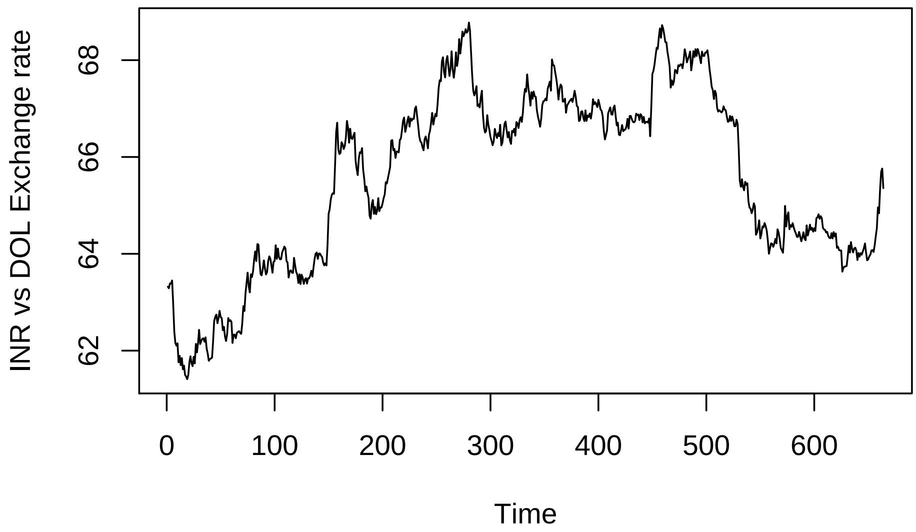

Figure 2.4: Time series plot of INR vs USD Exchange rate

Table 2.2: Summary statistics of INR vs USD log return series

Data	Summary
no: of obs	663
Minimum	-0.010105
Maximum	0.010269
Mean	0.000048
Median	0.000011
Variance	0.000008
Stdev	0.002913
Skewness	0.072711
Kurtosis	1.175012

In Figure 2.6 the sample ACF of the log return series and sample ACF and PACF of the squared series of daily log returns are given. Here, the sample ACF suggest no serial correlation whereas the sample ACF and PACF plots of the squared log return series are significant. It confirms that there is significant ARCH effect in the series of daily log returns. The value of Ljung-Box test statistic for squared

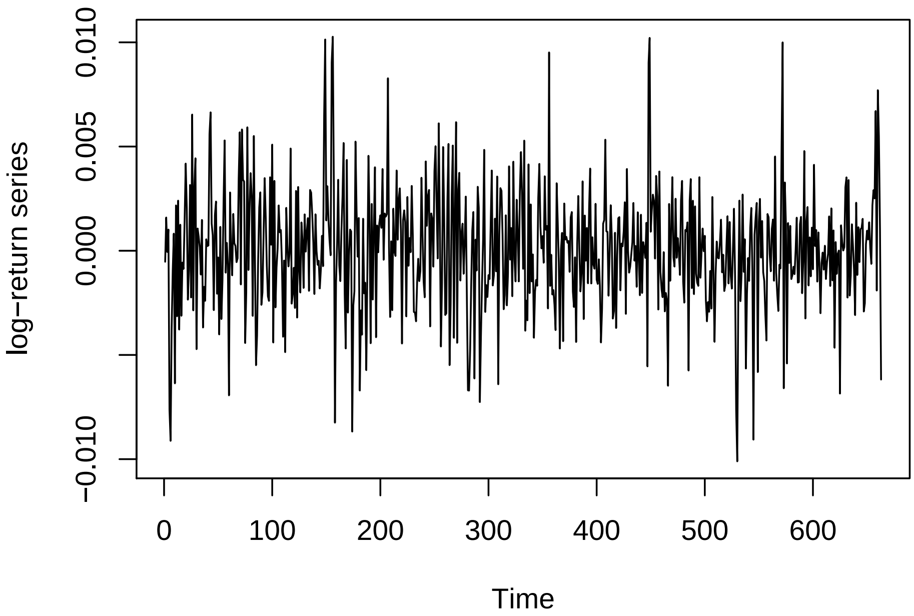

FIGURE 2.5: Daily log returns of the exchange rate between INR and USD from January 1, 2015, to September 30, 2017

series gives Q(20)= 37.681 with p-value close to zero and hence it confirm the ARCH effects in the exchange rate series. We model the series by utilizing the R commands of fGarch package. We propose a GARCH(1,1) model. In fGarch package the garchFit() allows for several conditional distributions. The edited output of GARCH modelling is given in Table 2.3. We use AIC and BIC values to select the model and results are tabulated in Table 2.4. From Table 2.4 the AIC value obtained for GARCH(1,1) with t distributed innovations is -8.88817 and for GED innovations is -8.88565. The corresponding BIC values are -8.85426 and -8.85174 respectively. We choose the model with minimum AIC/BIC value. The Q-Q plot of standardized residuals with different conditional distributions are given in Figure 2.7. From the Q-Q plot we can see that the normality assumptions are not satisfied

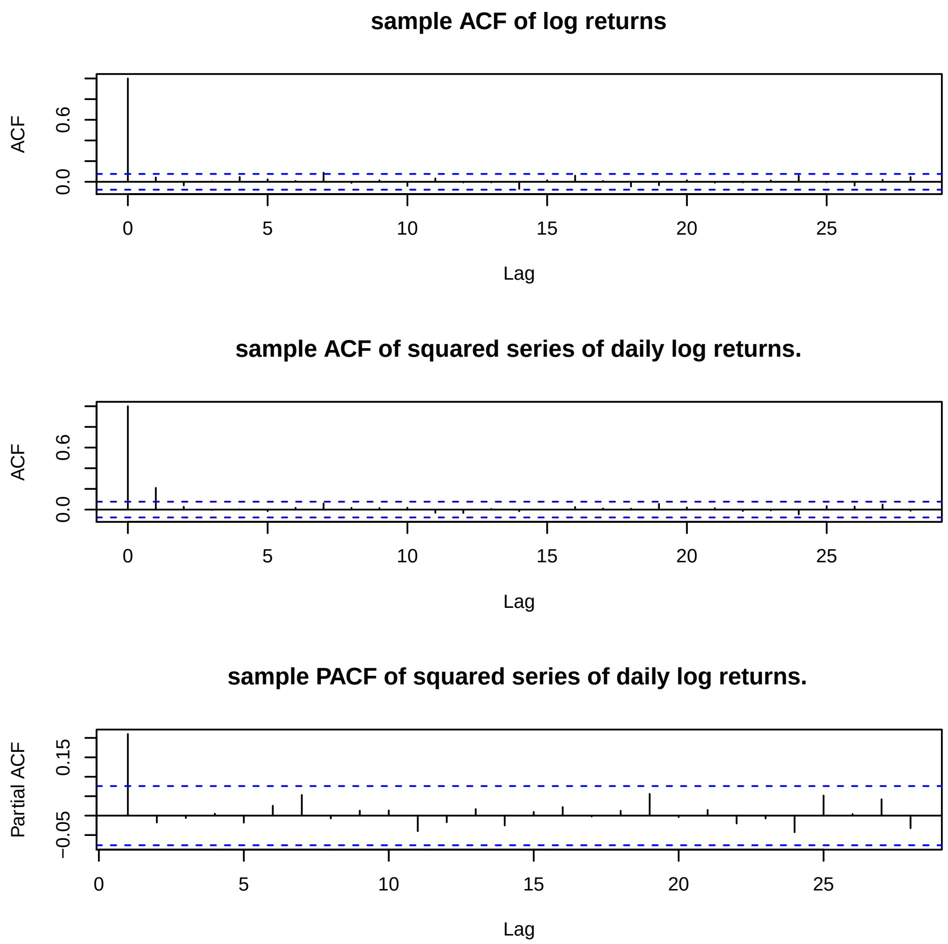

FIGURE 2.6: sample ACF of returns , squared series and sample PACF of squared series of daily log returns of the exchange rate between INR and USD from January 1, 2015, to September 30, 2017

by GARCH(1,1) model with normal and skewed normal innovations. The p-value obtained for testing no skewness is 0.9, which rejects the null hypothesis of skewness. So we choose the GARCH(1,1) model with t and GED innovations. The histogram of residuals with superimposed normal, t and GED density on the histogram of residuals is given in Figure 2.8. From the Figure the GARCH(1,1) model with t and GED innovations seems to be of good fit.

Table 2.3: Results of Garch Modelling : Standardised Residuals Tests

Conditional Distribution		Jarque-Bera Test	Ljung-Box Test R	Ljung-Box Test R^2
norm	Statistic	17.1099	18.8347	13.55707
	p-Value	0.000192	0.53259	0.85222
t	Statistic	18.5864	18.8764	14.3444
	p-Value	$9.204686e-05$	0.52986	0.81261
ged	Statistic	17.59493	18.85807	13.7775
	p-Value	0.000151	0.53106	0.84159
snorm	Statistic	17.1096	18.83608	13.54916
	p-Value	0.000192	0.53251	0.85259
sstd	Statistic	18.59724	18.87617	14.35559
	p-Value	$9.155071e-05$	0.52988	0.812001
sged	Statistic	17.59682	18.85766	13.78268
	p-Value	0.00015097	0.53109	0.84134

Table 2.4: The Information Criterion Statistics of GARCH model with different conditional distributions

Conditional Distribution	Information Criterion Statistics	
norm	AIC	-8.86904
	BIC	-8.84191
t	AIC	**-8.88817**
	BIC	**-8.85426**
ged	AIC	-8.88565
	BIC	-8.85174
snorm	AIC	-8.86604
	BIC	-8.83213
sstd	AIC	-8.88516
	BIC	-8.84446
sged	AIC	-8.88264
	BIC	-8.84194

2.4 Duration Models

The availability of the intraday databases consisting of detailed information on the complete trading process led to the modelling of high-frequency data. These are the transaction-by-transaction or tick-by-tick data in security markets. Here, let

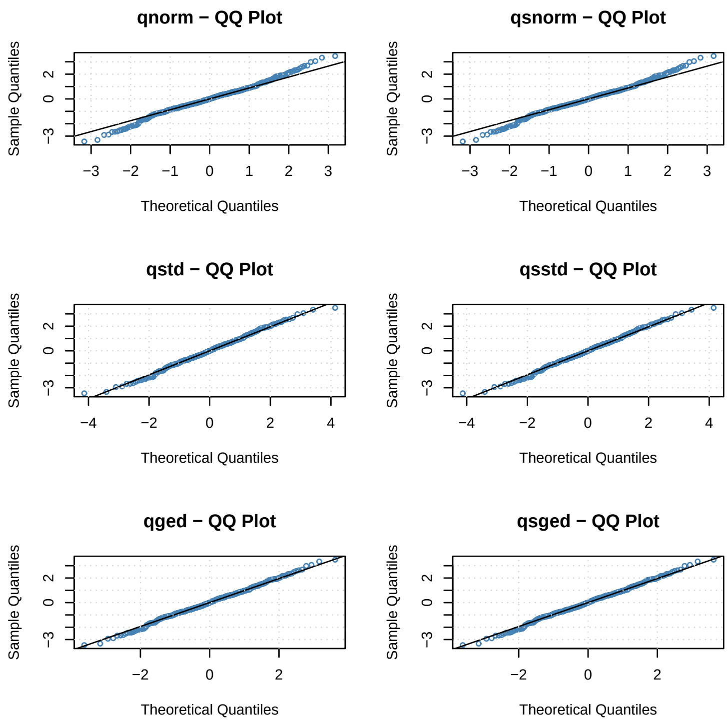

FIGURE 2.7: The Quantile-to-quantile plot

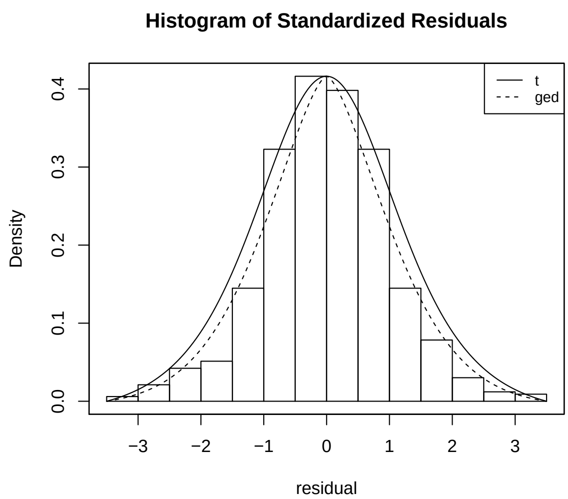

FIGURE 2.8: The superimposed t and GED density on the histogram of residuals

't_i' be the calender time which is measured in seconds from midnight, at which the i^{th} trade of an asset takes place. These data have some unique characteristics that do not appear in lower frequencies. The durations are subjected to strong intraday seasonality patterns, which is also reflected in the autocorrelation functions. In particular, the empirical characteristics of trading data are:

1. unequally spaced time intervals.
2. existence of a daily periodic or diurnal pattern.
3. multiple transactions within a single second.

To study the behaviour of the durations and to describe the evaluation of the durations of stocks Engle and Russell (1998) proposed the Autoregressive Conditional Duration model, described below.

2.4.1 Autoregressive Conditional Duration Models.

Let τ_i be the time of occurrence of an event(or transaction) of interest with $\tau_0 = 0$ and $X_i = \tau_i - \tau_{i-1}, i = 1, 2, \ldots, n$ be the i^{th} trade duration, which is defined as the waiting time between two consecutive transactions of an underlying asset from time i to $i+1$. Also, let ψ_i be the expected adjusted duration given $\mathcal{F}_i$, where $\mathcal{F}_{i-1}$ is the information set available at the $(i-1)^{th}$ trade. That is, $\psi_i = E(X_i|\mathcal{F}_{i-1})$ is the conditional expectation of the adjusted duration between the $(i-1)^{th}$ and i^{th} trades, where $\mathcal{F}_{i-1} = \sigma(X_1, X_2, \ldots, X_{i-1})$ is the sigma field generated by $(X_1, X_2, \ldots, X_{i-1})$ is the information known at time $i-1$. The basic ACD model is defined as

$$X_i = \psi_i \varepsilon_i \tag{2.10}$$

where ε_i is a sequence of independent and identically distributed non-negative random variables such that $E(\varepsilon_i) = 1$. The choice of different specifications for the expected duration ψ_i and different distributions for ε_i results in a variety of models. When the distribution of ε_i is exponential, the resulting model is called an Exponential autoregressive conditional duration ($EACD(r,s)$) model. Similarly, if ε_i follows a Weibull distribution, the model is a Weibull autoregressive conditional duration ($WACD(r,s)$) model. For further details on duration models, one can refer Grammig and Maurer (2000), Pacurar (2008) and the references therein. Engle and Russell (1998) proposed the $ACD(r,s)$ model specifying ψ as

$$\psi_i = \omega + \sum_{j=1}^{r} \alpha_j X_{i-j} + \sum_{j=1}^{s} \beta_j \psi_{i-j}, \tag{2.11}$$

where r and s are non-negative integers. The sequence is stationary if $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$, $\sum_{j=1}^{r} \alpha_j + \sum_{j=1}^{s} \beta_j < 1$. The model in (2.11) can also be formulated as an ARMA(r,s) model for durations. The ACD model and the GARCH model of Bollerslev (1986) defined in Section 2.3 share several common features.

2.4.2 Stochastic Conditional Duration Models.

Bauwens and Veredas (2004) introduced a class of models called the stochastic conditional duration models to study the evolution of durations driven by latent variables. Unlike ACD model, SCD model generates a double stochastic process, that is, a model with two stochastic innovations, one for the observed duration and the other for the latent variable. The SCD model treats the conditional mean of durations as a stochastic process captured by an appropriate distribution with positive support. The SCD model is defined by

$$\begin{aligned} X_i &= e^{\psi_i}\varepsilon_i, \\ \psi_i &= \omega + \phi\psi_{i-1} + \eta_i\,, \qquad i = 1, 2, \ldots, n, \\ \psi_0 &\sim N\left(0, \frac{\sigma^2}{1-\phi^2}\right), \end{aligned} \tag{2.12}$$

where $|\phi| < 1$ to ensure the stationarity of the process and η_i follows independent and identically distributed $N(0, \sigma^2)$ so that $\{\psi_i\}$ follows a Gaussian AR(1) sequence and $\{\varepsilon_i\}$ is an independent and identically distributed sequence on the positive support with common probability density function $f(\varepsilon_i)$ and η_j is independent of $\varepsilon_i \;\forall\; i, j$. Note that the model depends on the unobservable ψ_i, called the latent variable. One interpretation for the latent variable is that, it captures the random flow of information which is not directly observable.

The estimation of the parameters of the SCD model involves the evaluation of a multiple integral and hence find its difficult to compute the likelihood function. This requires computing an integral that has the dimension of the sample size. In view of this, several authors introduced different estimation methods. (Bauwens and Galli (2009) , Strickland et al. (2006) etc).

2.5 Bayesian Estimation of Duration Models

In Bayesian inference, the Bayes' theorem is used to update the information of the unknown model parameters. The continuous form of Bayes' theorem is

$$p(\Theta|y) \propto L(\Theta|y)\pi(\Theta) \tag{2.13}$$

where Θ=unknown parameter to be estimated.

$y = (y_1, y_2, \ldots, y_n)$.

$\pi(\Theta)$=prior distribution of Θ depending on one or more parameters, called hyper-parameters.

$L(\Theta|y)$=likelihood function for Θ.

$p(\Theta|y)$=posterior(updated) distribution of Θ.

The likelihood function for an SCD model is difficult to evaluate exactly. In view of this Strickland et al. (2006) developed a Bayesian Markov Chain Monte Carlo approach using Gibbs sampling and Metropolis-Hastings(M-H) algorithm for the estimation of parameters of SCD models. As the evaluation of integral is difficult the sampling is done through Monte Carlo simulation.

2.5.1 Markov Chain Monte Carlo Methods

Algorithms for simulating from the posterior distribution can be divided as independent simulation and dependent simulation categories. Rejection sampling and importance sampling are the representatives of the first category. These algorithms produce an independent and identically distributed sample from the posterior. In dependent simulation algorithms, the output is a sample of identically distributed(but not independent) draws from the posterior. All algorithms based on generation (simulation) of a Markov Chain belongs to this category.

Metropolis-Hastings Algorithm

The algorithm consists of two stages: first, a draw from the proposal density is obtained and second, that draw is either retained or rejected with certain probability. Let $p(\Theta|y)$ be the posterior density from which sampling is not possible. Let Θ be a K-dimensional parameter vector, $\Theta = (\theta_1, \theta_2, \ldots, \theta_k)$ and $q(\Theta|\Theta^{t-1})$ be the approximating density, called the proposal density or the candidate-generating density. This is to generate randomly a realization of Θ given the value at the previous iteration of the algorithm. The steps involved in the algorithm are given as follows:

1. Let $\Theta^{(0)}$ be the initial value from the parameter space of Θ.
2. At t^{th} iteration, draw a (multivariate) realization, Θ^*, from the proposal density, $q(\Theta|\Theta^{t-1})$, where Θ^{t-1} is the parameter value at the previous step.
3. Compute the acceptance probability, given by
$$a(\Theta^*, \Theta^{t-1}) = min\left\{1, \frac{\frac{p(\Theta^*)}{q(\Theta^*|\Theta^{t-1})}}{\frac{p(\Theta^{t-1})}{q(\Theta^{t-1}|\Theta^*)}}\right\},$$
4. Draw u from the uniform distribution on (0,1), U(0,1). Then,
   if $u \leq a(\Theta^*, \Theta^{t-1})$, set $\Theta^{(t)} = \Theta^*$.
   Otherwise, set $\Theta^{(t)} = \Theta^{(t-1)}$.
5. Go back to Step 2.

The algorithm is iterated(step 2 through step 5) a large number of times. After the convergence of the chain, a sample from the posterior distribution is obtained .

2.5.2 Slice Sampling

Neal (2003) introduced the class of slice sampling methods. In Slice sampling method, one can sample points uniformly from the region under the curve of its density function and then look only at the horizontal coordinates of the sample points. A Markov chain that converges to this uniform distribution can be constructed by alternately sampling uniformly from the vertical interval defined by the density at the current point and from the union of intervals that constitutes the horizontal "slice "though the plot of the density function that this vertical position defines. It makes use of the fact that drawing a sample from a distribution $p(x)$ is the same as uniformly sampling from the points underneath the curve of such a distribution. Let us consider the case of single-variable slice sampling method in detail.

Single-variable slice sampling method

Let x_i be the single real variable being updated(with subscripts denoting different such points). To update x_i, compute a function, $f_i(x_i)$, that is proportional to $p(x_i|\{x_j\}_{j\neq i})$, where $\{x_j\}_{j\neq i}$ are the values of the other variables. Let $f(x)$ be the function proportional to the probability density of x. In the single variable slice sampling method the current value, x_0 is replaced with a new value, x_1, found by a three-step procedure:

i Draw a real value, y, uniformly from $(0, f(x_0))$, thereby defining a horizontal slice: $S = \{x : y < f(x)\}$. Here, x_0 is always within S.

ii Find an interval, $I = (L, R)$, around x_0 that contains all, or much, of the slice.

iii Draw the new point, x_1, from the part of the slice within this interval.

Step (i) picks a value for the auxiliary variable that is characteristic of slice sampling. To avoid possible problems with floating-point underflow one can also compute $g(x) = log(f(x))$, and use the auxiliary variable $z = log(y) = g(x_0) - e$, where e is exponentially distributed with mean one, and define the slice by $S = \{x : z < g(x)\}$. Steps (ii) and (iii) are implemented in such a way that the resulting Markov chain leaves the distribution defined by $f(x)$ invariant.

Chapter 3

Autoregressive Moving Average Model with Generalized Error Distributed innovations

3.1 Introduction

The Box-Jenkins methodology of time series analysis focusses on stationary Autoregressive Moving Average models with Gaussian innovations. That is, the observed time series $\{Z_t\}$ is generated by a linear $ARMA(p, q)$ model defined as in (1.10). The likelihood based analysis of time series in classical set up assumes that, the innovations are independent and identically distributed normal random variables with mean 0 and constant variance σ^2. Even though the assumption of normality makes the analysis simpler, the model, fails to take care of many real life situations which are better explained by non-Gaussian distributions. This motivated several researchers to introduce classes of non-Gaussian time series models during the last three decades. One of the advantages of assuming normally distributed innovations in ARMA models is that, it leads to a normally distributed stationary marginal distribution for Z_t and vice-versa. This is not true for non-Gaussian ARMA models and as a result, there is no unified theory available in this case. Most of the non-Gaussian time series models are distribution specific, either in terms of a specified stationary marginal distribution or in terms of a specific innovation distribution.

Some of the examples of these type of models are discussed in Section 1.6 of Chapter 1. One of the difficulties in developing the former class of models is that the likelihood based inference becomes intractable due to a complicated form of the innovation distributions. In the latter class of models, a suitable distribution is specified for the innovation variables, but not bothered about the specific marginal distribution except that the time series is stationary. In this Chapter, we propose an $ARMA$ model whose errors follow a Generalized Error Distribution which includes, Normal and Laplace as special cases.

3.1.1 Generalized Error Distribution

Let us recall the probability density function of GED discussed in the previous Chapter. A random variable X is said to follow a GED(μ, σ, β) if its probability density function is given by

$$f(x) = \frac{\beta}{2\,\sigma\,\Gamma(1/\beta)} \exp\left\{-\left(\left|\frac{(x-\mu)}{\sigma}\right|\right)^{\beta}\right\}, \quad -\infty < x < \infty$$
$$-\infty < \mu < \infty, \;\; \sigma > 0, \, \beta > 0. \quad (3.1)$$

This is a symmetric distribution with mean μ, variance $\sigma^2\Gamma(3/\beta)/(\Gamma(1/\beta)$ and coefficient of kurtosis $(\Gamma(1/\beta)\Gamma(5/\beta))/(\Gamma(3/\beta)^2)$. One may refer Johnson et al. (1995), Chapter 24 for more details. The density function (3.1) is also known as the exponential power distribution and is a parametric family of symmetric distributions, which allows leptokurtic $(0 < \beta < 2)$ and platykurtic $(\beta > 2)$ distributions. In particular the probability density function in (3.1) reduces to that of Laplace distribution for $\beta = 1$ and it becomes $\mathcal{N}(\mu, \frac{\sigma^2}{2})$ for $\beta = 2$. Further the pointwise limit $\lim_{\beta\to\infty} f(x,\mu,\sigma,\beta)$ coincides with the probability density function of the uniform distribution $\mathcal{U}(\mu-\sigma,\mu+\sigma)$. Explicit expression for the characteristic function of GED(μ,σ,β) with pdf (3.1) is given by Pogány and Nadarajah (2010) as

$$\zeta_X(s) = \frac{\sqrt{\pi}e^{is\mu}}{\Gamma[\frac{1}{\beta}]} {}_1\Psi_1 \left[\begin{array}{c} (\frac{1}{\beta},\frac{2}{\beta}) \\ (\frac{1}{2},1) \end{array} ; -\frac{(\sigma s)^2}{4} \right],$$

provided $\beta > 1$, where ${}_p\Psi_q(.)$ denotes the Fox-Wright generalized hypergeometric function with p numerator and q denominator parameters, defined by

$${}_p\Psi_q = \left[\begin{array}{c} (\alpha_1, A_1), \ldots, (\alpha_p, A_p) \\ (\beta_1, B_1), \ldots, (\beta_q, B_q) \end{array} ; z \right] = \sum_{n=0}^{\infty} \frac{\prod_{j=1}^{p} \Gamma(\alpha_j + A_j n)}{\prod_{j=1}^{q} \Gamma(\beta_j + B_j n)} \frac{z^n}{n!},$$

where the series converges for $A_j, B_k > 0$ and $1 + \sum_{j=1}^{q} B_j - \sum_{j=1}^{p} A_j > 0$.

Some other forms of the distribution and related properties are discussed in D. B. Nelson (1991), Diananda (1949) and Nadarajah (2005). Agro (1995) addresses the problem of obtaining maximum likelihood estimates for the three parameters of GED and Varanasi and Aazhang (1989) discuss parameter estimation of GED using the method of moments and maximum likelihood.

The Chapter is organized as follows: In Section 3.2, we describe the stationary ARMA(p,q) model with GED innovations. In this Section, the ARMA(1,1), MA(1) and AR(1) models are discussed. In Section 3.3, the estimation procedures for the models discussed in Section 3.2 are considered. Section 3.4 establish the asymptotic properties of AR(1) model with GED innovations. Section 3.5 consists of the simulation studies for ARMA(1,1), MA(1) and AR(1) model. In Section 3.6, we discuss the applicability of this model to two financial data sets.

3.2 GED-ARMA(p,q) Model

Consider the $ARMA(p,q)$ model for the time series $\{Z_t\}$ defined by (1.10). Also, the innovation sequence $\{a_t\}$ is assumed to follow independent and identically distributed GED with common probability density function

$$f(a) = \frac{\beta}{2\,\sigma\,\Gamma(1/\beta)} \exp\left\{-\left(\left|\frac{a}{\sigma}\right|\right)^{\beta}\right\} \;,\; -\infty < a < \infty \;,\; \sigma > 0,\; \beta > 0. \tag{3.2}$$

Now, the model becomes

$$\Phi(B)Z_t = \Theta(B)a_t, \tag{3.3}$$

with $\{a_t\}$ given by (3.2). The time series $\{Z_t\}$ defined by (3.3) with GED errors is stationary when the roots of the polynomial $\phi(B)$ lie outside the unit circle. However, there is no explicit expression for the stationary distribution. Our focus is on the applicability of this model in describing the real life problems. The ARMA(p,q) model with GED innovations for $p = q = 1$ is defined as follows.

3.2.1 GED-ARMA(1,1) Model

The stationary $ARMA(1,1)$ model for $\{Z_t\}$ may be defined by

$$Z_t - \phi Z_{t-1} = a_t - \theta a_{t-1} \qquad t = 1, 2, \ldots \tag{3.4}$$

with $|\phi| < 1$, $|\theta| < 1$. The errors are assumed to be independent and identically distributed with common probability density function given by (3.2). All the moments of $\{Z_t\}$ exists. The first four moments are:

$$
\begin{aligned}
E(Z_t) &= 0 = E(Z_t^3). \\
E(Z_t^2) &= \frac{(1+\theta^2-2\theta\phi)(\sigma^2\Gamma(3/\beta))}{(1-\phi^2)(\Gamma(1/\beta))}.
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
E(Z_t^4) = \frac{\sigma^4}{1-\phi^4} \Bigg(6\,\phi^2\,\theta^2 \left(\left(\frac{\phi^2}{1-\phi^2} + \left(\theta^2 - 2\,\theta\,\phi\right) \right) {g_1}^2 + g_2 \right) \\
- \left(12\,\phi\,\theta + 6\,\theta^2 + \frac{6\,\phi^2}{1-\phi^2} \right) {g_1}^2 + \left(1 - 4\,\phi\,\theta^3 + \theta^4\right) g_2 \Bigg),
\end{aligned}
\tag{3.6}
$$

where $g_1 = \frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}$ and $g_2 = \frac{\Gamma(5/\beta)}{\Gamma(1/\beta)}$.

The Autocorrelation function, ACF is given by

$$
\rho_Z(k) = \frac{(1-\phi\theta)(\phi-\theta)}{1-2\theta\phi+\theta^2}\phi^{k-1} \qquad k \geq 1, \tag{3.7}
$$

which is the standard form of ACF for any stationary $ARMA(1,1)$ model.

3.2.2 GED-MA(1) Model

The stationary MA(1) model for $\{Z_t\}$ with GED innovations is defined by

$$
Z_t = a_t - \theta a_{t-1} \qquad t = 1, 2, ..., \tag{3.8}
$$

with $|\theta| < 1$ and $\{a_t\}$ is an independent and identically distributed GED with common probability density function (3.2). The first four moments of $\{Z_t\}$ are

$$
\begin{aligned}
E(Z_t) &= 0. \\
E(Z_t{}^2) = V(Z_t) &= (1+\theta^2)\,\sigma^2 \frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}. \qquad (3.9) \\
E(Z_t{}^3) &= 0. \\
E(Z_t{}^4) &= (1-4\theta+6\theta^2-4\theta^3+\theta^4)\,\sigma^4 \frac{\Gamma(5/\beta)}{\Gamma(1/\beta)}. \qquad (3.10)
\end{aligned}
$$

Also $Cov(Z_t, Z_{t-1}) = -\theta\,\sigma^2\,\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}$

$$
\text{or } \gamma_Z(k) = \begin{cases} -\theta\,\sigma^2\,\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)} & if\ |k| < 1, \\ 0 & if\ |k| = 1. \end{cases}
$$

The ACF is given by

$$
\rho_Z(k) = \begin{cases} \frac{-\theta}{1+\theta^2} & if \quad |k| = 1. \\ 0 & if \quad |k| > 1. \end{cases}
$$

3.2.3 GED-AR(1) Model

Let $\{Z_t\}$ be a stationary AR(p) model defined by

$$
Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t, \qquad t = 1, 2, \ldots \cdot
$$

For p=1, we get the AR(1) process. The AR(1) model is defined by

$$Z_t = \phi Z_{t-1} + a_t, \qquad t = 1, 2, \ldots, \tag{3.11}$$

where $|\phi| < 1$. The a_t's are independent and identically distributed with common probability density function (3.2). The mean, variance , third and fourth moments are

$E(Z_t) = 0.$

$V(Z_t) = \frac{\sigma^2}{(1-\phi^2)} \frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}.$

$E(Z_t^3) = 0$, and

$E(Z_t^4) = \sigma^4 \left(\frac{(6\phi^2)\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \right) / (1 - \phi^4).$

The autocorrelation function is $\rho_Z(k) = \phi^k, \; k = 1, 2, ...$ The characteristic function is given by

$$\begin{aligned}
\zeta_z(s) &= \zeta_z(\phi^t s) \prod_{j=1}^{t-1} (\phi^{(j-1)s} \\
&= \zeta_{z_0}(\phi^t s) \prod_{j=1}^{t-1} \frac{\sqrt{\pi}}{\Gamma[\frac{1}{\beta}]} {}_1\Psi_1 \left[\begin{array}{c} (\frac{1}{\beta}, \frac{2}{\beta}) \\ (\frac{1}{2}, 1) \end{array} ; \frac{-(\sigma\phi^j s)^2}{4} \right].
\end{aligned}$$

Under stationarity, $|\phi| < 1$. Then

$$\zeta_z(s) = \zeta_{z_0}(0) \prod_{j=1}^{\infty} \frac{\sqrt(\pi)}{\Gamma[\frac{1}{\beta}]} {}_1\Psi_1 \left[\begin{array}{c} (\frac{1}{\beta}, \frac{2}{\beta}) \\ (\frac{1}{2}, 1) \end{array} ; \frac{-(\sigma\phi^j s)^2}{4} \right].$$

3.3 Estimation of Parameters

To explain the applicability of the model, we need to estimate the parameters appearing in the model. The usual estimation methods such as Yule-Walker, Method of moments, Maximum likelihood may be applied in this case, as well. The likelihood function of the parameters based on a realization $(z_1, z_2, \ldots, z_n)$ can be obtained from the joint probability density function of the innovations $a_1, a_2, \ldots, a_n$ using an explicit form of the model (3.3). In this case the joint density function of $a_1, a_2, \ldots, a_n$ is given by

$$f(a_1, a_2, \ldots, a_n) = \frac{\beta^n}{2^n \sigma^n \Gamma(1/\beta)^n} \exp\left(-\sum_{t=1}^{n} \left|\frac{a_t}{\sigma}\right|^{\beta}\right), \quad -\infty < a_t < \infty \;, \sigma > 0, \; \beta > 0. \tag{3.12}$$

In order to express the likelihood function in terms of the data, we can represent (3.12) in terms of data by replacing each of a_t's as $a_t = z_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}$ for $t = 1, 2, \ldots, n$. Conditioning on $a_{t-q} = 0$ for $(t-q) \leq 0$ and $z_{t-p} = 0$ for $(t-p) \leq 0$ the term in the exponential $\sum_{t=1}^{n} \left|\frac{a_t}{\sigma}\right|^{\beta}$ can be represented in terms of z_t's. The parameter vector to be estimated is $\Lambda = (\Phi, \Theta, \sigma, \beta)'$, where $\Phi = (\phi_1, \phi_2, \ldots, \phi_p)'$ and $\Theta = (\theta_1, \theta_2, \ldots, \theta_q)'$ and we consider the method of ML and method of moments. The ML method provides an estimate of Λ that maximize the likelihood function, $L(\Lambda | z_1, z_2, \ldots, z_n)$ or equivalently the log-likelihood function. Here the ML estimators do not have closed forms and they have to be obtained by numerical methods (Newton-Raphson). The sample autocorrelation and moment estimates can be used as initial guess for estimating the parameters in the iteration procedures. Now a detailed inference for $p = q = 1$ is considered.

3.3.1 GED-ARMA(1,1) model

As a preliminary analysis, we estimate the parameters by the method of moment using $E(Z_t^2)$, $E(Z_t^4)$, ρ_1 and ρ_2. Replacing ρ_1 and ρ_2 by the corresponding sample ACF's r_1 and r_2 respectively in (3.7), we get $\hat{\phi} = (r_2/r_1)$. Then solve for $\hat{\theta}$ in $r_1 = ((1-\hat{\phi}\theta)(\hat{\phi}-\theta))/(1-2\theta\hat{\phi}+\theta^2)$ and for β using the equation $\frac{E[Z_t^4]}{E[Z_t^2]^2} = \frac{\sum Z_t^4}{(\sum Z_t^2)^2}$ using some iterative root finding technique, with $\beta > 0$. Finally, obtain the estimate of σ using the moment equation (3.5) after substituting for the estimates of ϕ, θ, β. From (3.5), we have $\sigma = \left(\frac{(1-\phi^2)(\Gamma(1/\beta))E(Z_t^2)}{(1+\theta^2-2\theta\phi)\Gamma(3/\beta)}\right)^{\frac{1}{2}}$. We can obtain the estimate of σ by replacing ϕ, θ and β with their corresponding estimates in the above explicit form. We apply the method of conditional likelihood to estimate the parameter vector $\Lambda = (\phi, \theta, \sigma, \beta)'$ by taking $a_0 = 0$ and $z_j = 0$ for $j \leq 0$ by using the inverted form of the model (cf. Box et al. (1994)):

$$a_t = z_t - \sum_{j=1}^{\infty} \pi_j z_{t-j}, \quad where\ \pi_j = (\phi-\theta)\theta^{j-1}, \quad j = 1, 2, \ldots; t = 2, 3, \ldots$$

and $a_1 = z_1$.

Under the condition that, $z_t = 0$ for $t \leq 0$, we have $a_t = z_t + (\theta-\phi)\sum_{j=1}^{t-1} \theta^{(j-1)} z_{t-j}$ for $\quad t = 2, 3, \ldots, n$.

So the resulting conditional likelihood function via (3.3) can be expressed as

$$L(\phi, \theta, \sigma, \beta) = \frac{\beta^n}{2^n\, \sigma^n\, \Gamma(1/\beta)^n} \exp\left\{-\frac{1}{\sigma^\beta}\left(|a_1|^\beta + |a_2|^\beta + \ldots + |a_n|^\beta\right)\right\} \qquad (3.13)$$

$$= \frac{\beta^n}{2^n\,\sigma^n\,\Gamma(1/\beta)^n}\exp\left\{-\frac{1}{\sigma^\beta}\left(|z_1|^\beta + |z_2+(\theta-\phi)z_1|^\beta + \ldots \right.\right.$$
$$\left.\left. + \left|z_t + (\theta-\phi)\sum_{j=1}^{t-1}\theta^{(j-1)}z_{t-j}\right|^\beta\right)\right\}.$$

Then the log likelihood function is given by

$$\log\,L = c + n\,\log\,\beta - n\log\sigma - n\log\Gamma(1/\beta) - \frac{1}{\sigma^\beta}\left(\sum_{t=1}^{n}\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{(j-1)}z_{t-j}\right|^\beta\right),$$

where c is a constant. The ML estimates of parameters can be obtained by maximizing this likelihood form. However, to obtain the ML estimate as a solution of the likelihood equation we consider the following likelihood equations. The log likelihood function is differentiable for $\beta > 1$, and the first derivative of the log likelihood function w.r.t Λ can be expressed as

$$\frac{\partial}{\partial\phi}logL = \frac{\beta}{\sigma^\beta}\sum_{t=1}^{n}\left\{\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{(j-1)}z_{t-j}\right|^{\beta-2}\left(\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right)\right.$$
$$\left.\left(z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right)\right\},$$

$$\frac{\partial}{\partial\theta}logL = -\frac{\beta}{\sigma^\beta}\sum_{t=1}^{n}\left\{\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right|^{\beta-2}\left(z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right)\right.$$

$$\left.\left((\theta-\phi)\sum_{j=1}^{n}(j-1)\,\theta^{j-2}\,z_{t-j} + \sum_{j=1}^{t}(\theta^{j-1}z_{t-j})\right)\right\},$$

$$\frac{\partial}{\partial\sigma}logL = -\frac{n}{\sigma} + \frac{\beta}{\sigma^{(1+\beta)}}\sum_{t=1}^{n}\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right|^\beta, \qquad \text{and}$$

$$\frac{\partial}{\partial\beta}logL = \frac{n}{\beta} + \frac{n\,\Psi(1/\beta)}{\beta^2} + \frac{1}{\sigma^\beta}\sum_{t=1}^{n}\left\{\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right|^\beta\right.$$
$$\left.\left(log\,\sigma - log\left(\left|z_t + (\theta-\phi)\sum_{j=1}^{t}\theta^{j-1}z_{t-j}\right|\right)\right)\right\},$$

where $\Psi(z) = \frac{d}{dz}\log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function.

The ML estimators of ϕ, θ, β, σ are obtained as the solutions of the likelihood equations. As the likelihood equations do not lead to closed form solutions, we have used numerical methods for obtaining the ML estimates. The computations based on simulated data are presented in Section 3.5.

3.3.2 GED-MA(1) model

In order to express the likelihood function in terms of the data, we again consider the inverted form of the model given by

$$a_t = z_t + \sum_{j=1}^{t} \theta^j z_{t-j}, \qquad t = 1, 2, \ldots, n,$$

with the condition $z_t = 0$ for $t \leq 0$ and $a_0 = 0$.

So the likelihood function can be expressed as

$$\begin{aligned} L(\theta,\, \sigma, \beta) = & \frac{\beta^n}{2^n\, \sigma^n\, \Gamma(1/\beta)^n} \\ & \exp\left\{ \frac{-1}{\sigma^\beta} \left(|z_1|^\beta + |z_2 + \theta\, z_1|^\beta + \ldots + \left| z_n + \theta z_{n-1} + \ldots + \theta^{n-1} z_1 \right|^\beta \right) \right\} \\ = & \frac{\beta^n}{2^n\, \sigma^n\, \Gamma(1/\beta)^n} \exp\left\{ -\frac{1}{\sigma^\beta} \left(\sum_{l=1}^{n} \left(\left| \sum_{k=0}^{l-1} \theta^k z_{l-k} \right|^\beta \right) \right) \right\}. \end{aligned} \tag{3.14}$$

The log likelihood function (without the constant) is given by

$$\log\, L = n\, \log\, \beta - n \log \sigma - n \log \Gamma(1/\beta) - \frac{1}{\sigma^\beta} \left(\sum_{l=1}^{n} \left(\left| \sum_{k=0}^{l-1} \theta^k z_{l-k} \right|^\beta \right) \right). \tag{3.15}$$

By maximizing the above log-likelihood form, we can obtain the ML estimates. In terms of the likelihood equations, the derivative of the log likelihood equations for $\beta > 1$ is obtained as

$$\frac{\partial}{\partial\theta}\log\, L = -\frac{\beta}{\sigma^{\beta}}\left(\sum_{l=1}^{n}\left(\left|\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\right|^{\beta-2}\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\sum_{k=0}^{l-1}k\theta^{k-1}z_{l-k}\right)\right),$$

$$\frac{\partial}{\partial\beta}\log\, L = \frac{n}{\beta}+\frac{n}{\beta^{2}}\Psi(\tfrac{1}{\beta})-\left(\frac{1}{\sigma^{\beta}}\sum_{l=1}^{n}\left(\left|\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\right|^{\beta}\log\left(\left|\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\right|\right)\right)\right)$$

$$+\frac{\log\sigma}{\sigma^{\beta}}\left(\sum_{l=1}^{n}\left(\left|\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\right|^{\beta}\right)\right),$$

$$\frac{\partial}{\partial\sigma}\log\, L = -\frac{n}{\sigma}+\frac{\beta}{\sigma^{\beta+1}}\left(\sum_{l=1}^{n}\left(\left|\sum_{k=0}^{l-1}\theta^{k}z_{l-k}\right|^{\beta}\right)\right),$$

where $\Psi(z) = \frac{d}{dz}\log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$, is the digamma function. The ML estimators of θ, β, σ are the solutions of the likelihood equations. As the likelihood equations do not lead to closed form solutions, we have used numerical methods for obtaining ML estimates. The computations based on simulated data are presented in Section 3.5. The sample autocorrelation and the moment estimate are used as the initial guess for the estimates in the iteration procedures.

We have $\rho(1) = \frac{-\theta}{1+\theta^2}$. Equating ρ_1 to r_1, we are led to solve a quadratic in θ, where r_1 is the sample autocorrelation given by $r_1 = \frac{\sum_{t=2}^{n}(z_t-\bar{z})(z_{t-1}-\bar{z})}{\sum_{t=1}^{n}(z_t-\bar{z})^2}$. Then the two real roots are given by $\frac{-1\pm\sqrt{1-4{r_1}^2}}{2r_1}$. However, only one of the solution satisfy the invertibility condition $|\theta| < 1$ and is obtained as $\hat{\theta}_m = \frac{-1+\sqrt{1-4{r_1}^2}}{2r_1}$. If $r_1 = \pm 0.5$, unique, real solution exist, namely ∓ 1, but neither is invertible. If $|r_1| > 0.5$, no real solution exist, and so the method of moments failed to yield an estimator of θ. On squaring (3.9) and dividing by (3.10), the resulting equation is

$$\frac{\left(\frac{1}{n}\sum_{i=1}^{n} {z_i}^2\right)^2}{\frac{1}{n}\sum_{i=1}^{n} {z_i}^4} = \frac{\left(\theta^2+1\right)^2\Gamma\left(\frac{3}{\beta}\right)^2}{(\theta^4+1)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{5}{\beta}\right)+6\theta^2\Gamma\left(\frac{3}{\beta}\right)^2}.$$

As an initial estimate of θ is obtained, it is possible to rewrite the above equation as

$$\frac{\Gamma(5/\beta)(\Gamma(1/\beta))}{(\Gamma(3/\beta))^2} = \frac{-6\,\hat{\theta}_m^2\left(\frac{1}{n}\sum_{i=1}^{n} {z_i}^2\right)^2+\frac{1}{n}\sum_{i=1}^{n} {z_i}^4\left(1+\hat{\theta}_m^2\right)^2}{(1+\hat{\theta}_m^4)}.$$

In the above equation the RHS can be evaluated as $\hat{\theta}_m$ is obtained. By equating it to the LHS, we can find the value of β, ie., an initial guess for β, $\hat{\beta}_m$ is obtained. The likelihood equation for σ has an explicit form. By substituting the values of $\hat{\theta}$ and $\hat{\beta}$ in the equation
$\sigma = \left(\frac{\beta\left(\sum_{l=1}^{n}\left(\left|\sum_{k=0}^{l-1}\theta^k z_{l-k}\right|^{\beta}\right)\right)}{n}\right)^{\frac{1}{\beta}}$, we can obtain the estimate of σ.

3.3.3 GED-AR(1) model

Let us consider the AR(p) model with GED innovations. The conditional likelihood function of the general AR(p) model with parameters $\Lambda = (\Phi, \sigma, \beta)$ can be expressed

as

$$L(\Lambda|z_1, z_2, ..., z_n) \propto f(z_1)f(z_2|z_1)\dots f(z_p|z_1,\dots,z_{p-1})$$
$$\prod_{t=p+1}^{n} f_{t|t-1,t-2\dots,t-p}(z_t|z_{t-1}, z_{t-2},\dots,z_{t-p})$$
$$= \frac{\beta^n}{2^n\,\sigma^n\,\Gamma(1/\beta)^n}\exp\left\{-\sum_{t=p+1}^{n}\left(\frac{|\,(z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2}-\cdots-\phi_p z_{t-p})\,|}{\sigma}\right)^{\beta}\right\}.$$

The ML estimators are obtained by maximizing the above log likelihood function. For $\beta > 1$, the likelihood equations can be obtained as

$$\frac{\partial}{\partial\phi_i}\log L = \frac{\beta}{\sigma^\beta}\sum_{t=p+1}^{n}(|z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2}-\dots-\phi_p z_{t-p}|)^{\beta-2}$$
$$(z_t-\phi z_{t-1}-\dots-\phi_p z_{t-p})\,(z_{t-i}) = 0 \quad (i=1,2,\dots,p),$$

$$\frac{\partial}{\partial\beta}\log\,L = \frac{n}{\beta}+\frac{n}{\beta^2}\Psi\left(\frac{1}{\beta}\right)\sum_{t=p+1}^{n}(|z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2}-\dots-\phi_p z_{t-p}|)^{\beta}$$
$$log\left(\frac{|z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2}-\dots-\phi_p z_{t-p}|}{\sigma}\right),$$

$$\frac{\partial}{\partial\sigma}\log\,L = \frac{-n}{\sigma}+\frac{\beta}{\sigma}\sum_{t=p+1}^{n}\left(\frac{|z_t-\phi z_{t-1}-\dots-\phi_p z_{t-p}|}{\sigma}\right)^{\beta},$$

where $\Psi(\cdot)$ is the digamma function defined by $\Psi(z) = \frac{d}{dz}log(\Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)}$. The ML estimators of ϕ_i, $\beta\,\text{and}\,\sigma$ are obtained as the the solutions of the likelihood equations. As the ML estimates of Φ, β, σ do not possess explicit forms, we need to obtain them by numerical methods.

By taking p=1, 2,. . . k, $k \leq p$, in an AR(p) model, we get the corresponding autoregressive model of lag 'k' with GED innovations. However, we discuss the estimation problem of AR(1) process in view of its Markov property, which leads to some asymptotic properties.

The $\{Z_t\}$ defined by (3.1) is a Markov sequence with transition density

$$f(z_t|z_{t-1}) = \frac{\beta}{2\,\sigma\,\Gamma(1/\beta)} \exp\left\{-\left(\frac{|z_t - \phi z_{t-1}|}{\sigma}\right)^{\beta}\right\}, \quad -\infty < z_t < \infty, \quad \sigma > 0,\ \beta > 0. \tag{3.16}$$

Then the conditional likelihood function of $\Lambda = (\phi, \sigma, \beta)$ is given by

$$\begin{aligned} L(\Lambda|z_1, z_2, ..., z_n) &\propto \prod_{t=2}^{n} f_{t|t-1}(z_t|z_{t-1}) \\ &= \frac{\beta^n}{2^n\,\sigma^n\,\Gamma(1/\beta)^n} \exp\left\{-\sum_{t=2}^{n}\left(\frac{|\,(z_t - \phi z_{t-1})\,|}{\sigma}\right)^{\beta}\right\}. \end{aligned} \tag{3.17}$$

The ML estimates of the parameters are obtained by maximizing this likelihood form. To study the properties and check the regularity conditions of the estimators, we obtain the ML estimators as the solution of likelihood equations. The log-likelihood function (without the constant) becomes

$$\log L = n\log\beta - n\log\,\sigma - n\log\,\Gamma(1/\beta) - \sum_{t=2}^{n}\left(\frac{|z_t - \phi z_{t-1}|}{\sigma}\right)^{\beta}. \tag{3.18}$$

As the likelihood function is differentiable for $\beta > 1$, the first derivative of the likelihood function w.r.t Θ can be expressed as

$$\frac{\partial}{\partial\phi}\log L = \frac{\beta}{\sigma^{\beta}}\sum_{t=2}^{n}\left(|z_t - \phi\, z_{t-1}|\right)^{\beta-2}(z_t - \phi\, z_{t-1})\,(z_{t-1}), \tag{3.19}$$

$$\frac{\partial}{\partial\beta}\log L = \frac{n}{\beta} + \frac{n}{\beta^2}\Psi\left(\frac{1}{\beta}\right) - \sum_{t=2}^{n}\left(\frac{|z_t - \phi\, z_{t-1}|}{\sigma}\right)^{\beta}\log\left(\frac{|z_t - \phi\, z_{t-1}|}{\sigma}\right), \tag{3.20}$$

$$\frac{\partial}{\partial\sigma}\log L = \frac{-n}{\sigma} + \frac{\beta}{\sigma}\sum_{t=2}^{n}\left(\frac{|z_t - \phi\, z_{t-1}|}{\sigma}\right)^{\beta}, \tag{3.21}$$

where $\Psi(\cdot)$ is the digamma function defined by $\Psi(z) = \frac{d}{dz}log(\Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)}$ for $\beta > 1$. The ML estimators of ϕ, β, σ are the solutions of the likelihood equations $\frac{\partial logL}{\partial\phi} = 0$, $\frac{\partial logL}{\partial\beta} = 0$, $\frac{\partial logL}{\partial\sigma} = 0$ for $\beta > 1$. We solve the resulting likelihood equations iteratively, using the moment estimates as the initial values, and the computation results based on simulated data is summarised in Section 3.5. However, in our case, the regularity conditions hold only when $\beta = 2$ and $\beta \geq 3$. Problem here is the non-differentiability of the likelihood function with respect to ϕ for other values of β. So we study the properties of the MLE of σ and β by considering ϕ to be known. The estimates of β and σ can be obtained by solving $\partial logL/\partial\beta = 0$ and $\partial logL/\partial\sigma = 0$.

Generalized Method of Moments

The method of maximum likelihood discussed in the earlier Section helps in studying the ML estimators properties only for two parameters σ and β. So in this Section, we propose the method of moments to estimate all the three parameters, and establish their asymptotic properties. To estimate the parameter vector Λ' we use the following moment equations

$E(Z_t^2) = (\sigma^2 \Gamma\left(\frac{3}{\beta}\right))/((1-\phi^2)\,\Gamma\left(\frac{1}{\beta}\right)).$

$E(Z_t^4) = (\sigma^4 \left(\frac{(6\phi^2)\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right))/(1-\phi^4).$

$E\left(Z_t Z_{t-1}\right) = (\sigma^2 \phi \Gamma\left(\frac{3}{\beta}\right))/((1-\phi^2)\,\Gamma\left(\frac{1}{\beta}\right)).$

In order to establish the properties of the estimators, we define the estimating function:

$$g(z_t, z_{t-1}, \Lambda) = \begin{pmatrix} z_t^2 - \frac{\sigma^2\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \\ z_t^4 - \frac{\sigma^4\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4} \\ z_t z_{t-1} - \frac{\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \end{pmatrix}.$$

Then the moment estimator $\hat{\Lambda}' = (\hat{\phi}, \hat{\sigma}, \hat{\beta})$ of Λ' may be obtained by solving the equation

$$\frac{1}{n}\sum_{t=1}^{n} g\left(z_t, z_{t-1}, \Lambda\right) = 0.$$

The resulting moment equations for estimating ϕ, σ and β are expressed as

$$\hat{\phi} = (\overline{Y_{11}})/(\overline{Y_2}); \quad \frac{\overline{Y_4}}{\overline{Y_2^2}} = \frac{\left(1-\phi^2\right)^2\Gamma\left(\frac{1}{\beta}\right)^2\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{(1-\phi^4)\Gamma\left(\frac{3}{\beta}\right)^2}; \quad \sigma = \sqrt{\frac{\overline{Y_2}\left(1-\hat{\phi}^2\right)\Gamma\left(\frac{1}{\beta}\right)}{\Gamma\left(\frac{3}{\beta}\right)}},$$

where $\overline{Y_2} = (\sum\limits_{t=1}^{n} z_t^2)/(n), \quad \overline{Y_{11}} = (\sum\limits_{t=1}^{n} z_t z_{t-1})/(n), \quad \overline{Y_4} = (\sum\limits_{t=1}^{n} z_t^4)(n).$

3.4 Asymptotic properties of the estimators for GED-AR(1) Model.

3.4.1 Maximum Likelihood Estimates

Under certain regularity conditions listed below, Billingsley (1961) (pp. 10-14) proved that the MLE is Consistent and Asymptotically Normal for Markov sequence. Let $\{Z_t\}$ be a stationary Markov sequence with one-step transition density function $f(z_t;\theta|z_{t-1})$ and the initial density $f_Z(z_1;\theta)$ and Θ be the parameter space.

i) $\log f(z_t;\theta|z_{t-1})$ is thrice differentiable with respect to θ for all θ in a neighbourhood I of θ_0.

ii) $E|\partial^2 logf(z_t;\theta_0|z_{t-1})/\partial\theta_i\theta_j| < \infty$;
$E|\partial^2 logf(z_t;\theta_0|z_{t-1})/\partial\theta_i\theta_j| + (\partial logf(z_t;\theta_0|z_{t-1})/\partial\theta_i)^2| < \infty$.

iii) There exist sequences $\{K(t)\}$ and $\{M(t)\}$ of positive constants with $K(T) \to \infty$ and $M(T) \to \infty$, as $T \to \infty$ such that

  (a) $M(T)\{K(T)\}^{-1}\sum_{t=1}^{T} \partial logf(z_t;\theta_0|z_{t-1}/\partial\theta_i \xrightarrow{L} N(0, B(\theta_0))$ for some non-random function $B(\theta_o) > 0$,

  (b) $\{K(T)\}^{-1}|\sum_{t=1}^{T} \partial^2 logf(z_t;\theta_0|z_{t-1}/\partial\theta_i\partial\theta_j| \xrightarrow{p} A(\theta_0)$, and

  (c) for all $\epsilon > 0$ and for all $\nu > 0$, there exists $\delta = \delta(\epsilon, \nu)$ and $N = N(\epsilon, \nu)$ such that for all $T > N$,

$$P\left[\{K(T)\}^{-1}\left|\sum_{t=1}^{T}\left(\frac{\partial^2 logf(z_t;\theta^*|z_{t-1})}{\partial\theta_i\theta_j} - \frac{\partial^2 logf(z_t;\theta_0|z_{t-1})}{\partial\theta_i\theta_j}\right)\right| > \nu\right] < \epsilon,$$

whenever, $|\theta^* - \theta_0| < \delta$, where $\theta^* = \theta_0 + r(\theta - \theta_0)$ with $r = r(T, \theta_0)$ satisfying $|r| < 1$.

Hence, there exists a root $\hat{\theta}$ of the likelihood equation with P_{θ_0}-probability approaching one which is consistent for θ_0 as $n \longrightarrow \infty$. Under the conditions (i)-(iii), any consistent solution of the maximum likelihood equation is asymptotic normal (CAN). That is,

$$M(T)(\hat{\theta} - \theta_0) \xrightarrow{L} N(0, C(\theta_0)),$$

where $C(\theta_0) = ((C_{ij}(\theta_0))) = A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1}$ and
$A_{ij}(\theta_0) = E(\partial^2 logf(z_t; \theta_0 | z_{t-1} / \partial\theta_i \partial\theta_j))$,
$B_{ij}(\theta_0) = E((\partial logf(z_t; \theta_0 | z_{t-1} / \partial\theta_i)(\partial logf(z_t; \theta_0 | z_{t-1} / \partial\theta_j)$.

In our case, the above regularity conditions hold only when $\beta = 2$ and $\beta > 3$ (Varanasi and Aazhang (1989)). It is due to the non-differentiability of the likelihood function with respect to ϕ for other values of β. So, considering ϕ to be known, we obtain the MLE of σ and β. The MLE $\sqrt{n} \begin{pmatrix} \hat{\sigma} & -\sigma \\ \hat{\beta} & -\beta \end{pmatrix} \rightarrow N_2\left(0, I^{-1}\right)$

where $I_{11} = \frac{\beta}{\sigma^2}$
$I_{12} = I_{21} = \frac{1+\beta+\Psi(\frac{1}{\beta})}{\beta\,\sigma}$ and
$I_{22} = \frac{\beta^2+\beta(\Psi(1+\frac{1}{\beta}))^2+2\beta\Psi(\frac{1}{\beta})+\beta\Psi'(1+\frac{1}{\beta})+\Psi'(\frac{1}{\beta})}{\beta^4}$,
where $\Psi(z) = \frac{d}{dz}\log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$, is the digamma function. One can construct asymptotic confidence interval based on the diagonal elements of I^{-1}.

3.4.2 Generalized Method of Moment estimates

The asymptotic properties for generalized method of moments are established using the results of Hansen (1982). According to Hansen (1982) GMM estimators are obtained using a large number of moment equations. Under the following assumptions, Hansen (1982) proved that the generalized moment estimators are consistent and asymptotically normal. We state the result below.

1. $\{z_t : -\infty < t < \infty\}$ is stationary and ergodic sequence.

2. The parameter space Θ is an open subset of R^q that contains the true parameter θ_0.

3. $g(.,\theta)$ and $\partial g(.,\theta)/\partial\theta$ are Borel-measurable for each $\theta \in \Theta$ and $\partial g(z,.)/\partial\theta$ is continuous on Θ for each $z \in R^q$.

4. $\partial g_1/\partial\theta$ is first moment continuous at θ_0, $D = E[\partial g(z_t, \theta_0)/\partial\theta]$ exists, is finite, and has full rank.

5. Let $\omega_t = g(z_t, \theta_0),\ -\infty < t < \infty$ and $\nu_j = E(\omega_0|\omega_{t-j}, \omega_{-j-1}, \ldots) - E(\omega|\omega_{-j-1}, \omega_{-j-2}, \ldots)$, $j \geq 0$. The assumptions are that $E(\omega_0\omega_0')$ exists and is finite, $E(\omega_0|\omega_{-j}, \omega{-j} - 1, \ldots)$ converges in mean square to zero and $\sum_{j=0}^{\infty} E(\nu_j'\nu_j)^{1/2}$ is finte.

Now, we have the following Theorem proved by Hansen (1982).

Theorem 3.1. Suppose the sequence $\{z_t : -\infty < t < \infty\}$ satisfies the assumptions stated in Hansen (1982). Then $\left\{\sqrt{T}(\hat{\Lambda} - \Lambda), T \geq 1\right\}$ converges in distribution to a normal random vector with mean 0 and dispersion matrix $[DS^{-1}D']^{-1}$, where D =

$\mathrm{E}\left(\frac{\partial g(z_t,\Lambda_0)}{\partial \Lambda}\right)$ exists, is finite, and has full rank and $\frac{\partial g}{\partial \Lambda}$ is the first moment continuous at Λ_0 , $S = \sum_{k=-\infty}^{\infty} \Gamma_{(k)}$, $\Gamma_k = E\left(\omega_t \omega'_{t-k}\right)$ where $\omega_t = g(z_t, \Lambda_0), -\infty < t < \infty$.

The sequence $\{z_t\}$ defined in (3.11) is stationary, ergodic and has finite moments. The partial derivatives of f w.r.t Λ are

$$\frac{\partial g}{\partial \sigma} = \begin{pmatrix} -\frac{2\sigma\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \\ \\ -\frac{4\sigma^3\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4} \\ \\ -\frac{2\sigma\phi\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \end{pmatrix};$$

$$\frac{\partial g}{\partial \beta} = \begin{pmatrix} \frac{3\sigma^2\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{\sigma^2\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \\ \\ -\frac{\sigma^4\left(\frac{12\phi^2\Gamma\left(\frac{3}{\beta}\right)^2\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}-\frac{36\phi^2\Gamma\left(\frac{3}{\beta}\right)^2\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\Gamma\left(\frac{5}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2\Gamma\left(\frac{1}{\beta}\right)}-\frac{5\Gamma\left(\frac{5}{\beta}\right)\Psi\left(\frac{5}{\beta}\right)}{\beta^2\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4} \\ \\ \frac{3\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} \end{pmatrix};$$

and

$$\frac{\partial g}{\partial \phi} = \begin{pmatrix} -\frac{2\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)} \\ -\frac{\sigma^4\left(\frac{12\phi\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{12\phi^3\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)^2}\right)}{1-\phi^4} - \frac{4\sigma^4\phi^3\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{(1-\phi^4)^2} \\ -\frac{\sigma^2\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{2\sigma^2\phi^2\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)} \end{pmatrix}.$$

where $\Psi(.)$ is the digamma function defined by $\Psi(z) = \frac{d}{dz}log(\Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)}$. Thus, $\partial g/\partial\theta$ exists and continuous for all θ. Similarly, we can show that $E\left(\frac{\partial g}{\partial \Lambda}\right)$ where $E\left(\omega_0\omega_0'\right)$ exists and finite. Hence, the regularity conditions of Hansen hold good for our model. The derivation of the elements of the dispersion matrix which are required to compute the asymptotic standard errors of the estimators are given as follows.

$$\text{Let} \quad \Gamma_{(k)} = \begin{pmatrix} \gamma_{11}^{(k)} & \gamma_{12}^{(k)} & \gamma_{13}^{(k)} \\ \gamma_{21}^{(k)} & \gamma_{22}^{(k)} & \gamma_{23}^{(k)} \\ \gamma_{31}^{(k)} & \gamma_{32}^{(k)} & \gamma_{33}^{(k)} \end{pmatrix}; \quad k = 0, \pm 1, \pm 2, \ldots$$

and $\Gamma_{(k)}$=$\Gamma_{(-k)}$, $k = 1, 2, \ldots$. Then the 3×3 matrix S is given by $S = \Gamma_{(0)} + 2\sum_{k=1}^{\infty}\Gamma_{(k)}$.

When $k = 0$, the elements of $\Gamma_{(0)}$= $E\left(\omega_t\omega_t'\right)$ are obtained as

$$\gamma_{11}^{(0)} = \frac{\sigma^4\left(\left(9\phi^2+3\right)\Gamma\left(\frac{3}{\beta}\right)^2-\left(\phi^2-1\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{5}{\beta}\right)\right)}{(\phi^2-1)^2(\phi^2+1)\Gamma\left(\frac{1}{\beta}\right)^2};$$

$$\gamma_{12}^{(0)} = \frac{\sigma^6\left(6\phi^2\left(-14\phi^4+\phi^2+1\right)\Gamma\left(\frac{3}{\beta}\right)^3+\left(29\phi^6-15\phi^4-15\phi^2+1\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{5}{\beta}\right)\Gamma\left(\frac{3}{\beta}\right)+\left(-\phi^6+\phi^4+\phi^2-1\right)\Gamma\left(\frac{1}{\beta}\right)^2\Gamma\left(\frac{7}{\beta}\right)\right)}{(\phi^2-1)^3(\phi^2+1)(\phi^4+\phi^2+1)\Gamma\left(\frac{1}{\beta}\right)^3};$$

$$\gamma_{13}^{(0)}=\frac{\sigma^4\left(\phi\left(3\phi^4-\phi^2+2\right)\Gamma\left(\frac{3}{\beta}\right)^2-\phi^3\left(\phi^2-1\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{5}{\beta}\right)\right)}{(\phi^2-1)^2(\phi^2+1)\Gamma\left(\frac{1}{\beta}\right)^2};$$

$$\gamma_{21}^{(0)}=\gamma_{12}^{(0)};$$

$$\begin{aligned}\gamma_{22}^{(0)}=\Bigg(\sigma^8\Bigg(&36\phi^4\left(69\phi^8-\phi^6-2\phi^4-\phi^2-1\right)\Gamma\left(\frac{3}{\beta}\right)^4\\&-12\phi^2\left(104\phi^{10}-35\phi^8-36\phi^6-34\phi^4+1\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{3}{\beta}\right)^2\Gamma\left(\frac{5}{\beta}\right)\\&+28\phi^2\left(2\phi^{10}-\phi^8-2\phi^6+1\right)\Gamma\left(\frac{1}{\beta}\right)^2\Gamma\left(\frac{3}{\beta}\right)\Gamma\left(\frac{7}{\beta}\right)\\&-\left(\phi^2-1\right)^2\left(\phi^4+\phi^2+1\right)\Gamma\left(\frac{1}{\beta}\right)^2\left(\left(1-69\phi^4\right)\Gamma\left(\frac{5}{\beta}\right)^2+\left(\phi^4-1\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{9}{\beta}\right)\right)\Bigg)\Bigg)\\&\left(\left(\phi^2-1\right)^4\left(\phi^2+1\right)^2\left(\phi^4+1\right)\left(\phi^4+\phi^2+1\right)\Gamma\left(\frac{1}{\beta}\right)^4\right)^{-1};\end{aligned}$$

$$\begin{aligned}\gamma_{23}^{(0)}=-\sigma^6\phi\Bigg(&6\phi^2\left(5\phi^8-\phi^4+9\phi^2-1\right)\Gamma\left(\frac{3}{\beta}\right)^3\\&+\left(-15\phi^{10}+5\phi^8+11\phi^6-15\phi^4+10\phi^2+4\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{3}{\beta}\right)\Gamma\left(\frac{5}{\beta}\right)\\&+\phi^4\left(\phi^2-1\right)^2\left(\phi^2+1\right)\Gamma\left(\frac{1}{\beta}\right)^2\Gamma\left(\frac{7}{\beta}\right)\Bigg)\left(\left(\phi^2-1\right)^3\ \left(\phi^2+1\right)\left(\phi^4+\phi^2+1\right)\Gamma\left(\frac{1}{\beta}\right)^3\right)^{-1};\end{aligned}$$

$$\gamma_{31}^{(0)}=\gamma_{13}^{(0)};$$

$$\gamma_{32}^{(0)}=\gamma_{23}^{(0)};$$

$$\gamma_{33}^{(0)}=\frac{\sigma^4\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}-\frac{\sigma^4\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\sigma^4\phi^2\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4}.$$

Similarly, these following are the elements of $\Gamma_{(k)}$ for $k = 1, 2, \ldots$.

$$\gamma_{11}^{(k)} = c_1 \left(\sum_{i=1}^{k} \phi^{2(i-1)} \right) c_3 + \phi^{2k} c_4 - c_3^2;$$

$$\gamma_{12}^{(k)} = c_1 \left(\sum_{i=1}^{k} \phi^{2(i-1)} \right) c_4 + \phi^{2k} c_5 - c_3 c_4;$$

$$\gamma_{13}^{(k)} = c_1 \left(\sum_{i=1}^{k} \phi^{2i-1} \right) c_3 + 3\phi^{2k+1} c_1 c_3 + \phi^{2k+3} c_4 - c_7 c_3;$$

$$\gamma_{21}^{(k)} = 6c_1^2 c_3 \left(\sum_{j=0}^{k} \sum_{i=2j+1}^{j+k-1} \phi^{2i} \right) + 6c_1 \left(\sum_{i=1}^{k} \phi^{2((i-1)+k)} \right) c_4 + c_2 \left(\sum_{i=1}^{k} \phi^{4(i-1)} \right) c_3 + \phi^{4k} c_5 - c_3 c_4;$$

$$\gamma_{22}^{(k)} = 6c_1^2 c_4 \left(\sum_{j=0}^{k} \sum_{i=2j+1}^{j+k-1} \phi^{2i} \right) + 6c_1 \left(\sum_{i=1}^{k} \phi^{2((i-1)+k)} \right) c_5 + c_2 \left(\sum_{i=1}^{k} \phi^{4(i-1)} \right) c_4 + \phi^{4k} c_6 - c_3 c_4;$$

$$\gamma_{23}^{(k)} = 18c_1^2 \left(\sum_{i=1}^{k} \phi^{2((i-1)+k)+1} \right) c_3 + 6c_1 \left(\sum_{i=1}^{k} \phi^{2((i-1)+k)+3} \right) c_4 + c_2 \left(\sum_{i=1}^{k} \phi^{4(i-1)+1} \right) c_3 + \sum_{j=0}^{k} \sum_{i=2j+1}^{j+k-1} \phi^{2i+1} + 10\phi^{4k+3} c_1 c_4 + 5\phi^{4k+1} c_2 c_3 + \phi^{4k+5} c_5 - c_7 c_4;$$

$$\gamma_{31}^{(k)} = \left(c_1 \left(\sum_{i=1}^{k-1} \phi^{2i-1} \right) c_3 + \phi^{2k-1} c_4 \right) - c_7 c_3;$$

$$\gamma_{32}^{(k)} = \left(c_1 \left(\sum_{i=1}^{k-1} \phi^{2i-1} \right) c_4 + \phi^{2k-1} c_5 \right) - c_7 c_4;$$

$$\gamma_{33}^{(k)} = \left(c_1 \left(\sum_{i=1}^{k-1} \phi^{2i} \right) c_3 + 3\phi^{2k} c_1 c_3 + \phi^{2k+2} c_4 \right) - c_7^2,$$

where $c_1 = \frac{\sigma^2\Gamma(3/\beta)}{\Gamma(1/\beta)}$, $c_2 = \frac{\sigma^2\Gamma(5/\beta)}{\Gamma(1/\beta)}$, $c_3{=}E(Z_t^2)$, $c_4 = E(Z_t^4)$, $c_5{=}E(Z_t^6)$, $c_6{=}E(Z_t^8)$ and $c_7{=}Cov(z_1, z_2)$.

The 3×3 matrix D is evaluated using the form $D = E\left(\frac{dg(z_t,\Lambda)}{\mathrm{d}\Lambda}\right)$ and its elements are :

$$D_{11} = -\frac{2\sigma\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)};$$

$$D_{12} = -\frac{4\sigma^3\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4};$$

$$D_{13} = -\frac{2\sigma\phi\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)};$$

$$D_{21} = \frac{3\sigma^2\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{\sigma^2\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)};$$

$$D_{22} = -\frac{\sigma^4\left(\frac{12\phi^2\Gamma\left(\frac{3}{\beta}\right)^2\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} - \frac{36\phi^2\Gamma\left(\frac{3}{\beta}\right)^2\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2} + \frac{\Gamma\left(\frac{5}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2\Gamma\left(\frac{1}{\beta}\right)} - \frac{5\Gamma\left(\frac{5}{\beta}\right)\Psi\left(\frac{5}{\beta}\right)}{\beta^2\Gamma\left(\frac{1}{\beta}\right)}\right)}{1-\phi^4};$$

$$D_{23} = \frac{3\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{3}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)\Psi\left(\frac{1}{\beta}\right)}{\beta^2(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)};$$

$$D_{31} = -\frac{2\sigma^2\phi\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)};$$

$$D_{32} = -\frac{\sigma^4\left(\frac{12\phi\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{12\phi^3\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)^2}\right)}{1-\phi^4} - \frac{4\sigma^4\phi^3\left(\frac{6\phi^2\Gamma\left(\frac{3}{\beta}\right)^2}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)^2}+\frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\right)}{(1-\phi^4)^2};$$

$$D_{33} = -\frac{\sigma^2\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)\Gamma\left(\frac{1}{\beta}\right)} - \frac{2\sigma^2\phi^2\Gamma\left(\frac{3}{\beta}\right)}{(1-\phi^2)^2\Gamma\left(\frac{1}{\beta}\right)}.$$

Hence the asymptotic dispersion matrix becomes $\frac{1}{n}[DS^{-1}D']^{-1}$. The diagonal elements of this matrix are used to compute the asymptotic standard errors of the estimators.

3.5 Simulation study

3.5.1 Conditional Maximum Likelihood Estimates

A simulation study is carried out to assess the performance of the estimation procedure for ARMA(1,1), MA(1) and AR(1) models by the method of maximum likelihood for $\beta > 3$ with the sample size 500, 1000. This was repeated 100 times. The parameter estimates were obtained for each simulated series. The average and standard errors of the resulting estimates are presented in Tables 3.1, 3.2, 3.3. From the tables we can see that the estimates of ARMA(1,1), MA(1) and AR(1) are close to the true values. Thus, the simulation results suggest that ML estimates behave reasonably well for large samples. A simulation study is carried out with sample size 3000 to obtain the ML estimates of β and σ for specified values of ϕ. The average of the estimates and the corresponding standard errors of the estimates are presented in Table 3.4

TABLE 3.1: The average estimates and the corresponding standard errors of the maximum likelihood estimates based on simulated sample of size n=500 and 1000 for ARMA(1,1)

True Values(n=500)				Maximum Likelihood Estimates							
ϕ	θ	σ	β	$\hat{\phi}$	sd($\hat{\phi}$)	$\hat{\theta}$	sd($\hat{\theta}$)	$\hat{\sigma}$	sd($\hat{\sigma}$)	$\hat{\beta}$	sd($\hat{\beta}$)
0.6	0.2	2	3.1	0.59375	0.07709	0.19336	0.09200	1.98833	0.08211	3.07965	0.24503
			3.3	0.61003	0.06345	0.20522	0.08516	2.00623	0.07687	3.29812	0.27960
			3.5	0.61634	0.06762	0.22063	0.08535	2.00302	0.08662	3.51649	0.25124
-0.7	0.2		3.1	-0.69794	0.04039	0.20192	0.04887	1.99941	0.09020	3.09581	0.26828
			3.3	-0.69158	0.03910	0.20358	0.05350	1.99106	0.09726	3.35463	0.28364
			3.5	-0.69317	0.03705	0.20973	0.04959	2.00704	0.08487	3.46460	0.26621
-0.6	-0.3		3.1	-0.58970	0.09995	-0.29105	0.11302	2.00092	0.09750	3.13889	0.26306
			3.3	-0.59910	0.09104	-0.30108	0.11019	2.01911	0.10425	3.29637	0.27342
			3.5	-0.59746	0.08511	-0.30518	0.10503	2.00713	0.08970	3.49720	0.27030
0.2	-0.7		3.1	0.21255	0.05463	-0.68996	0.04502	2.00501	0.08880	3.04239	0.28078
			3.3	0.19473	0.04869	-0.70054	0.03350	1.99517	0.09371	3.24452	0.28924
			3.5	0.19746	0.04854	-0.69591	0.03351	2.00571	0.09479	3.49831	0.27427
n=1000											
0.6	0.2	2	3.1	0.59419	0.05488	0.19562	0.07003	1.99809	0.06589	3.12764	0.19348
			3.3	0.59633	0.04905	0.19553	0.05952	1.99455	0.07274	3.33283	0.22187
			3.5	0.59881	0.04906	0.19891	0.06327	1.99290	0.06461	3.46590	0.22039
-0.7	0.2		3.1	-0.70002	0.02568	0.20156	0.03725	1.99742	0.06917	3.08627	0.20719
			3.3	-0.70032	0.02717	0.19882	0.03985	2.00879	0.07348	3.32635	0.20311
			3.5	-0.70267	0.02653	0.19911	0.03417	1.99007	0.07273	3.50906	0.23380
-0.6	-0.3		3.1	-0.59674	0.06276	-0.29397	0.07812	2.00244	0.06630	3.13741	0.24829
			3.3	-0.58927	0.07179	-0.28669	0.08621	1.99800	0.06135	3.32393	0.22693
			3.5	-0.59192	0.06153	-0.29026	0.07387	2.00314	0.07436	3.48030	0.22915
0.2	-0.7		3.1	0.20244	0.03652	-0.69951	0.02723	2.00687	0.06612	3.09970	0.19558
			3.3	0.20203	0.03601	-0.69831	0.02450	1.99447	0.06764	3.27086	0.20868
			3.5	0.19675	0.04061	-0.70062	0.02582	1.99309	0.07138	3.47421	0.24635

3.5.2 Generalized method of moment estimates

In Tables 3.5 and 3.6 the GMM estimators of Λ are calculated with the sample size 3000 and 5000. For each simulated series the parameter estimates were obtained. We also calculate the asymptotic standard errors based on the theoretical values of the parameter. They are obtained as the square root of the diagonal elements of

TABLE 3.2: The average estimates and the corresponding standard deviation of the maximum likelihood estimates based on simulated observations of sample size n=500 and 1000 for MA(1) model.

True Values			Maximum Likelihood Estimates					
θ	σ	β	$\hat{\theta}$	sd($\hat{\theta}$)	$\hat{\sigma}$	sd($\hat{\sigma}$)	$\hat{\beta}$	sd($\hat{\beta}$)
-0.7	2	3.1	-0.70280	0.03165	1.99365	0.05453	3.10463	0.06259
		3.3	-0.70247	0.02999	1.99425	0.05510	3.31550	0.07233
		3.5	-0.69702	0.02738	1.98920	0.05354	3.50744	0.07363
-0.3		3.1	-0.29575	0.04082	1.99923	0.05539	3.10964	0.07525
		3.3	-0.30072	0.03944	1.99487	0.05497	3.32058	0.07061
		3.5	-0.29887	0.04243	1.99541	0.04585	3.50758	0.07677
0.2		3.1	0.20058	0.03874	1.99039	0.05592	3.10867	0.06911
		3.3	0.19535	0.04248	1.99187	0.05239	3.30207	0.06805
		3.5	0.19634	0.03834	1.99012	0.05133	3.49968	0.07708
0.7		3.1	0.70031	0.02543	1.99276	0.05071	3.11378	0.06897
		3.3	0.71017	0.04853	1.99107	0.06235	3.26445	0.15322
		3.5	0.69996	0.03055	1.99757	0.05520	3.51590	0.08005
-0.7	2	3.1	-0.70011	0.01873	1.99184	0.03035	3.10837	0.05017
		3.3	-0.70096	0.02045	1.99945	0.03589	3.31925	0.05019
		3.5	-0.69765	0.02297	1.99985	0.03599	3.5068	0.05462
-0.3		3.1	-0.30090	0.03091	1.99987	0.03713	3.10394	0.04706
		3.3	-0.29953	0.02849	1.99243	0.03465	3.30183	0.05042
		3.5	-0.29970	0.02846	2.00463	0.04114	3.50287	0.05964
0.2		3.1	0.20116	0.02718	2.00478	0.03643	3.10065	0.04802
		3.3	0.2036	0.03283	1.99209	0.03591	3.30054	0.05816
		3.5	0.20602	0.02922	1.99964	0.03780	3.50450	0.04995
0.7		3.1	0.69728	0.02321	1.99470	0.03783	3.09729	0.04729
		3.3	0.70317	0.03398	1.99680	0.03720	3.27854	0.09702
		3.5	0.69966	0.02078	1.99687	0.03738	3.50058	0.05893

the corresponding asymptotic dispersion matrix. These values are compared with the estimated standard errors of respective elements.

TABLE 3.3: The average estimates and the corresponding standard errors of the maximum likelihood estimates based on simulated observations of sample size n=500 and 1000 for AR(1) model.

	True Values			Maximum Likelihood Estimates					
n	ϕ	σ	β	$\hat{\phi}$	sd($\hat{\phi}$)	$\hat{\sigma}$	sd($\hat{\sigma}$)	$\hat{\beta}$	sd($\hat{\beta}$)
500	-0.7	2	3.1	-0.69574	0.02576	1.98617	0.06913	3.13188	0.29644
			3.3	-0.69667	0.02166	1.99449	0.05673	3.33743	0.297789
			3.5	-0.69348	0.02564	1.99903	0.05370	3.59775	0.301358
	-0.3		3.1	-0.29469	0.02638	1.99428	0.05451	3.09561	0.25252
			3.3	-0.30196	0.02361	1.99962	0.05071	3.34939	0.26267
			3.5	-0.29677	0.02980	2.00787	0.05186	3.57451	0.29966
	0.2		3.1	0.20003	0.02656	2.00150	0.06316	3.12273	0.28295
			3.3	0.20496	0.03069	1.99659	0.05062	3.33680	0.28813
			3.5	0.19799	0.03082	1.99929	0.04994	3.55290	0.283429
	0.6		3.1	0.60366	0.02505	1.99667	0.05914	3.13311	0.26461
			3.3	0.60080	0.02426	1.99356	0.05972	3.29678	0.29236
			3.5	0.59233	0.02310	1.99401	0.05346	3.48860	0.29171
1000	-0.7	2	3.1	-0.69931	0.01875	1.99274	0.04823	3.07917	0.207264
			3.3	-0.70044	0.01544	1.99419	0.04294	3.31003	0.22374
			3.5	-0.69952	0.01810	2.00480	0.04207	3.55943	0.254925
	-0.3		3.1	-0.29683	0.02179	1.99638	0.04733	3.08192	0.200464
			3.3	-0.30276	0.02278	2.00224	0.04314	3.32579	0.24673
			3.5	-0.29785	0.022834	2.003742	0.042997	3.515516	0.249564
	0.2		3.1	0.19992	0.02504	1.99960	0.04953	3.13211	0.20576
			3.3	0.19687	0.02282	1.99727	0.04289	3.30485	0.21901
			3.5	0.20434	0.02397	1.99793	0.04349	3.55918	0.27322
	0.6		3.1	0.60036	0.02022	1.99576	0.04625	3.11241	0.20227
			3.3	0.59349	0.02087	1.99703	0.04364	3.31851	0.214171
			3.5	0.59836	0.01813	2.00048	0.04199	3.56516	0.273019

TABLE 3.4: Average of estimates and the corresponding standard errors of the mle's of β and σ based on simulated sample size n=3000, $\sigma = 1.5$ and for specified values of ϕ. for an AR(1)

ϕ	β	$\hat{\sigma}$	se($\hat{\sigma}$)	$\hat{\beta}$	se($\hat{\beta}$)
	1.3	1.4922	0.0550	1.2929	0.0439
0.3	1.7	1.4983	0.0379	1.7009	0.0575
	2.1	1.5029	0.0322	2.1135	0.0805
	1.3	1.5030	0.0477	1.3026	0.0397
0.5	1.7	1.4943	0.0349	1.6909	0.0632
	2.1	1.5015	0.0316	2.1084	0.0741
	1.3	1.5031	0.0538	1.3021	0.0449
0.7	1.7	1.5050	0.0302	1.7067	0.0618
	2.1	1.4983	0.0326	2.1004	0.0787
	1.3	1.5022	0.0437	1.3053	0.0353
0.85	1.7	1.4967	0.0396	1.7013	0.0563
	2.1	1.5011	0.0309	2.1085	0.0737

TABLE 3.5: The average estimates and the corresponding standard errors of the moment estimates based on simulated observations of sample size n=3000. The estimates of asymptotic standard errors are also given for AR(1) model

True Values			$\hat{\phi}$			$\hat{\sigma}$			$\hat{\beta}$		
σ	β	ϕ	Mean	Std Dev	Asymp sd	Mean	Std Dev	Asymp sd	Mean	Std Dev	Asymp sd
2	1.3	0.3	0.29949	0.01784	0.02727	2.00618	0.12927	0.07328	1.30645	0.07811	0.03728
		0.5	0.49937	0.01531	0.02727	2.02097	0.16051	0.07328	1.31795	0.09799	0.03728
		0.7	0.69908	0.01358	0.05893	2.03429	0.25203	0.07135	1.33575	0.15948	0.03312
		0.85	0.84909	0.00918	0.02393	2.18336	0.58132	0.08981	1.59367	0.71321	0.03209
	1.5	0.3	0.29951	0.01767	0.02178	2.00936	0.09846	0.06535	1.51283	0.08837	0.03958
		0.5	0.49979	0.01595	0.05846	2.01339	0.11702	0.07595	1.51685	0.10407	0.04489
		0.7	0.69944	0.01315	0.04952	2.03335	0.23112	0.11292	1.55654	0.21851	0.01859
		0.85	0.85007	0.00976	0.02469	2.13914	0.51309	0.09828	1.95850	1.00912	0.04839
	1.7	0.3	0.30143	0.01773	0.01023	2.00077	0.07830	0.05903	1.70596	0.09001	0.03555
		0.5	0.49985	0.01542	0.05418	2.00525	0.10201	0.07723	1.71349	0.12086	0.06737
		0.7	0.69827	0.01251	0.03081	2.02773	0.19383	0.03768	1.77535	0.26582	0.05313
		0.85	0.84896	0.00981	0.02393	2.09693	0.45059	0.09634	2.26266	1.14557	0.06196
	1.85	0.3	0.30063	0.01823	0.03026	2.00258	0.06852	0.09656	1.86143	0.10193	0.01469
		0.5	0.50049	0.01539	0.04793	2.00092	0.09202	0.07616	1.86668	0.14091	0.09274
		0.7	0.69873	0.01318	0.02615	2.02739	0.18596	0.02304	1.95419	0.35314	0.07703
		0.85	0.84918	0.00998	0.02248	2.05779	0.41003	0.08878	2.46909	1.24752	0.06634

TABLE 3.6: The average estimates and the corresponding standard errors of the moment estimates based on simulated observations of sample size n=5000. The estimates of asymptotic standard errors are also given for AR(1) model

True Values			$\hat{\phi}$			$\hat{\sigma}$			$\hat{\beta}$		
σ	ϕ	β	Mean	Std Dev	Asymp sd	Mean	Std Dev	Asymp sd	Mean	Std Dev	Asymp sd
2	0.3	1.3	0.30081	0.01458	0.01219	1.98701	0.09828	0.03277	1.2934	0.05787	0.01667
	0.5		0.49928	0.01208	0.00993	2.02245	0.12473	0.03087	1.31826	0.07561	0.01709
	0.7		0.69876	0.00999	0.00781	2.01905	0.19648	0.03298	1.32310	0.12786	0.01927
	0.85		0.84902	0.00710	0.00563	2.03641	0.44063	0.03817	1.37042	0.30720	0.01732
	0.3	1.5	0.29850	0.01341	0.00974	1.99210	0.05986	0.02922	1.49629	0.05215	0.01770
	0.5		0.50159	0.01171	0.01835	2.01603	0.10416	0.02711	1.51949	0.09146	0.02547
	0.7		0.69860	0.00892	0.00742	2.02962	0.16107	0.03182	1.53588	0.15075	0.02567
	0.85		0.84940	0.00767	0.00492	2.10235	0.38802	0.03809	1.71109	0.49808	0.03153
	0.3	1.7	0.30265	0.01301	0.00457	2.00378	0.06307	0.02639	1.71071	0.07530	0.01590
	0.5		0.50068	0.01220	0.02287	2.00870	0.06803	0.03047	1.71235	0.08168	0.03568
	0.7		0.70013	0.01038	0.00779	2.05017	0.15455	0.02045	1.78649	0.20506	0.05209
	0.85		0.85056	0.00813	0.00541	2.06041	0.37724	0.03665	2.06986	0.95352	0.03292
	0.3	1.85	0.30092	0.01273	0.02088	1.99711	0.04866	0.03165	1.85483	0.06960	0.06725
	0.5		0.49975	0.01248	0.02945	2.01420	0.06496	0.04026	1.88191	0.10081	0.04131
	0.7		0.69828	0.01015	0.00561	2.00850	0.14723	0.02701	1.89760	0.24964	0.03669
	0.85		0.84946	0.00707	0.01005	1.98510	0.34376	0.03970	2.11583	0.94585	0.02967

3.6 Data Analysis

3.6.1 BSE Index

To illustrate the models discussed in the Chapter, we consider two sets of financial data. The first set of data consists of daily maximum of BSE Index from May 18^{th} 2006 to $June$ 27^{st} 2007. There are 277 observations. The plots of the actual time series, logarithmic differences, ACF and PACF of the log returns are given in Figure 3.1. One can see from the plots that the original series is not stationary, where as the logarithmic difference of the series is stationary. The ACF and PACF in Figure 3.1 give an indication that the log difference series might follow an AR(1) model. As a starting point we fitted an AR(1) model with Gaussian innovations and found

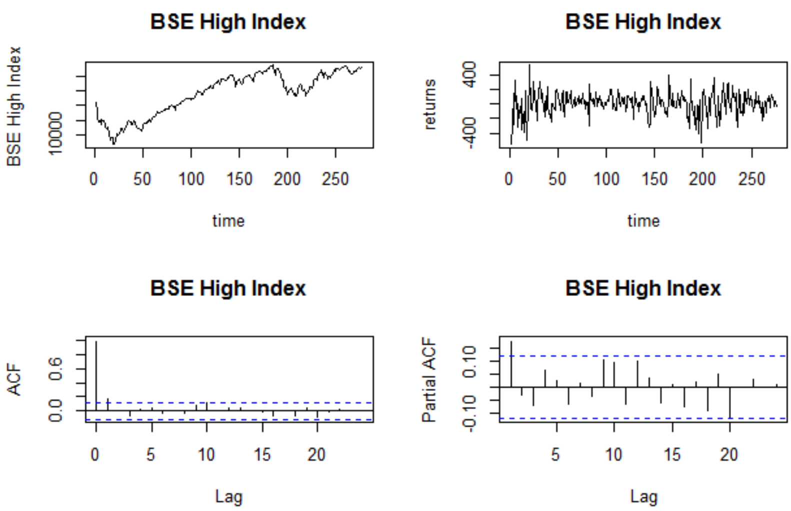

FIGURE 3.1: Time series plot of BSE Index and returns(top panel) and ACF and PACF plot of returns(bottom panel)

that the estimate of ϕ, $\hat{\phi}$ as 0.1791. However, the Q-Q plot which is shown in Figure 3.2 and the p-value(<0.0001) based on KolmogorovSmirnov(KS) test lead to the rejection of the null hypothesis that the residuals are normal.

Then we fit an AR(1) model with GED innovations to the data and obtained the estimates as $\hat{\phi}$=0.18156, $\hat{\sigma}$=.00928 and $\hat{\beta}$=1.1 by the method maximum likelihood. Now to check the validity of the model we consider the KS test. The KS statistic is obtained as 0.07524 and the p-value is 0.01508. So we cannot reject at 5% level of significance. The ACF plot of residuals given in Figure 3.2 shows that they are uncorrelated. The superimposition of the histogram of the residuals on the pdf of GED in the last panel of Figure 3.2 shows that there is a close agreement.

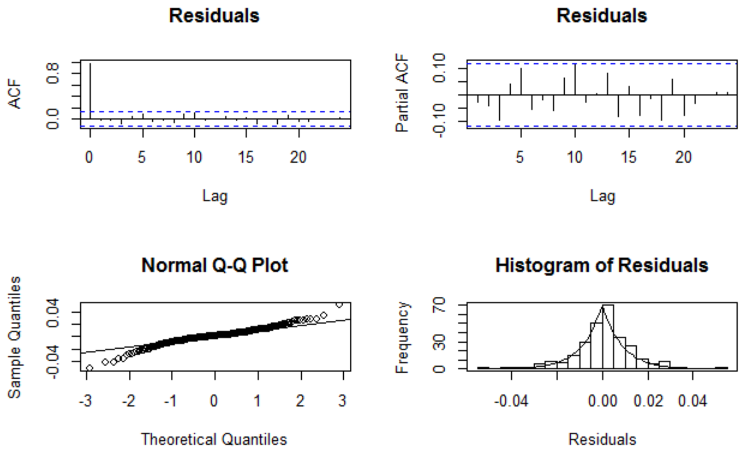

FIGURE 3.2: ACF and PACF plot of residuals(top panel) and Q-Q plot and histogram of residuals with superimposed GED density(bottom panel)

3.6.2 BSE 500 Index

Another set of data we considered is a time series of daily opening of BSE 500 INDEX from *July* 7^{th} 2010 to *December* 31^{st} 2010. There are 125 observations. The time series plot, logarithmic difference plot, ACF and PACF of the log returns are given in Figure 3.3.

The ACF/PACF plots in Figure 3.3 suggest that an MA(1) or an ARMA(1,1) model may be suitable for the data. Initially an MA(1) model and an ARMA(1,1) model with Gaussian innovations was fitted to the data. We obtained the estimate of θ as -0.435 for the MA(1) model and $\hat{\phi} = -0.334$ and $\hat{\theta} = 0.9$ for the ARMA(1,1) model. But the Q-Q plot and the p-value lead to the rejection of normality. Now we fit an MA(1) model with GED innovations to the data and obtained the estimates as $\hat{\theta}$=-0.41387, $\hat{\sigma}$=0.003652 and $\hat{\beta}$=0.75601 by GMM method. Also an ARMA(1,1) model with GED innovations is fitted and obtained the estimates as $\hat{\phi}$=0.1, $\hat{\theta}$=0.4518,

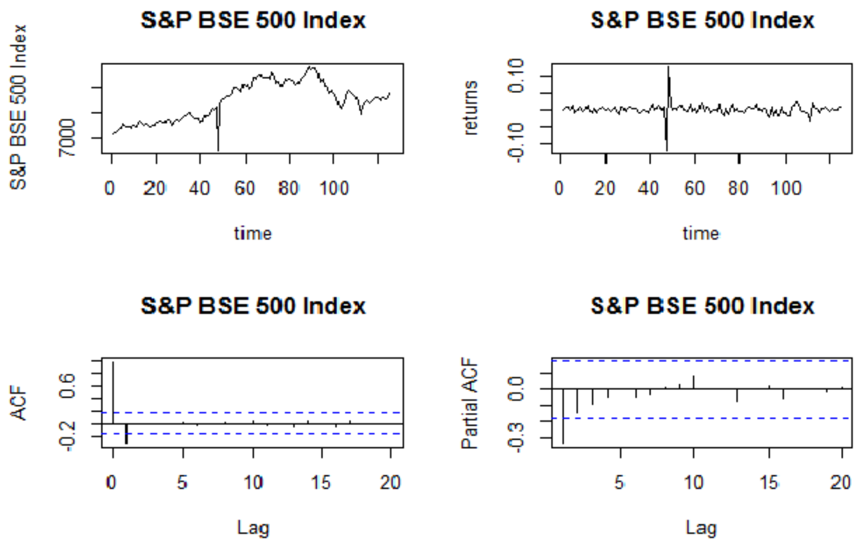

FIGURE 3.3: Time series plot of BSE 500 Index and returns(top panel) and ACF and PACF plot of returns(bottom panel)

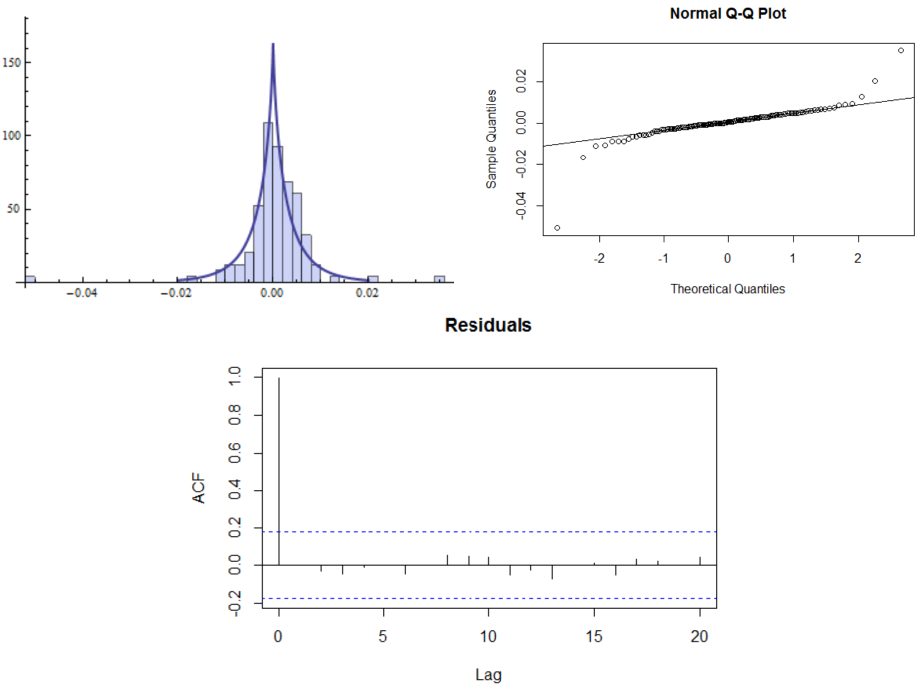

FIGURE 3.4: Histogram of residuals with superimposed GED density and Q-Q plot(top panel) and ACF plot of residuals(bottom panel)

$\hat{\sigma}$=0.00378 and $\hat{\beta}$=0.6. To fix the model, we use Akaike information criterion. The AIC values obtained for MA(1) and ARMA(1,1) model are -1074.3 and -951.89 respectively. The criterion identifies MA(1) model for the series. The KS statistic for MA(1) model is obtained as 0.11031 and the p-value is 0.21595. Also, the ACF of residuals given in Figure 3.4(c) are not significant. In Figure 3.4(a) we superimpose the GED density on the histogram of the residuals.

The results of this Chapter are reported in the paper Balakrishna and Sri Ranganath (2015).

Chapter 4

Hurwicz Estimator for Autoregressive model with Generalized Error Distributed innovations

4.1 Introduction

In the literature, the statistical problem of estimating the parameters of linear time series models has gained the interests of many researchers. The classical analysis of time series assumes that the sequence of observations is a realization from some Gaussian sequence. But the real life situations can be explained by non-Gaussian time series models, as seen in the previous Chapter. Majority of the literature deals with the linear autoregressive models in the study of Gaussian and non-Gaussian time series. So a contribution to the inference and estimation of AR(1) models may help to handle many real life situations. Several authors have proposed different estimation methods to estimate the coefficient of an AR(1) process as it is of great importance. For an AR(1) model, the least squares method is chosen frequently to estimate the autoregressive parameter. Sen (1968), Fox (1972) and Denby and Martin (1979) pointed out the estimation problem associated with the least square estimator in a first order autoregressive process. In literature, several authors study the problem of obtaining a robust estimate for the parameter of a stationary first-order AR process. In that direction, Hurwicz (1950) considered the problem of least square bias in time series of the AR(1) model $Z_t = \phi Z_{t-1} + \varepsilon_t$, t=1,2,...,n , where $|\phi| < 1$ is the autoregressive parameter. He observed that every ratio, $\{(Z_t)/(Z_{t-1})\}$ is an unbiased estimator of ϕ and proposed the estimator $\hat{\phi} = med(Z_t/Z_{t-1}, t = 1, 2, 3, \ldots, n)$, where med is the median.

Andrews (1993) studied the problem of LS estimator and proposed exactly median-unbiased estimator of ϕ. Fellag and Zieliński (1996) studied the bias of the LS estimator in a contaminated Gaussian model. Andel (1988) derived the distribution of an approximation to the Maximum Likelihood estimator $\hat{\phi} = min_{2\leq t\leq n}(Z_t/Z_{t-1})$ where the innovations are distributed according to exponential distributions. Zielinski (1999) proved that the median of the ratios of the consecutive observations of a stationary first order autoregressive process $Z_t = \phi Z_{t-1} + a_t$ with $P(a_t \geq 0) = P(a_t \leq 0) = 1/2$ and $P(Z_t = 0) = 0$ is a median unbiased estimator of ϕ. Also he showed that, it is true not only in the Gaussian case but whenever the medians of independent (not necessarily identically distributed) innovations $a_1, a_2, \ldots, a_n$ are equal to zero. Haddad (2000) developed a robust estimation method for the stationary Gaussian AR(1) process with autoregressive parameter ϕ. The method was based on the median of a product of two correlated normal variates. Guo (2000) proposed a robust estimator for ϕ, the median of ratio's of observations. Also, he compare the performance of the proposed estimator with Theil's , Hussain's and LS estimator. Berkoun et al. (2003) studied the testing problem of serial correlation in AR(1) model. Provost and Sanjel (2005) considered four estimates of ϕ, that can be expressed as the ratio of two quadratic forms in terms of observations, namely, a modified LS ,the Yule-Walker, Burg's estimator and the ordinary

LS estimators. They proposed bias-corrected estimators and obtained the bounds for the supports of the YW and Burgs estimators. The bias associated with certain time series parameters is also discussed in Kendall (1954), Marriott and Pope (1954) and Tjøstheim and Paulsen (1983). Breton and Pham (1989) determined the bias of the LS estimator under Gaussian white noise and provided asymptotic results. Luger (2006) shows that the Hurwicz estimator remains median-unbiased for the AR model with ARCH innovations. Fellag (2010) consider the problem of stability of estimation in AR models for the finite sample case . He compare the performance of the LS estimator and the Hurwicz estimator and showed that the Hurwicz estimator performs better when the distribution of innovations is heavy tailed. Berkoun and Douki (2011) extend the result of Haddad (2000) by taking the innovations to be in the domain of attraction of $\alpha-$stable symmetric distributions. Berkoun and Fellag (2011) established the asymptotic result of Hurwicz estimator for the coefficient of autoregressive in a linear process with innovations belonging to the domain of attraction of an α-stable law$(1 < \alpha < 2)$.

In view of the above studies, in this Chapter, we propose the Hurwicz's estimator for the coefficient of $AR(1)$ model whose errors follow a Generalized Error Distribution which includes, normal and Laplace as special cases. Balakrishna and Sri Ranganath (2015) discussed the estimation methods such as method of moments and maximum likelihood and the difficulties involved in estimating the parameters. Here, we consider the the Hurwicz's estimator for the autoregressive coefficient and analyse the asymptotic properties for the autoregressive parameter ϕ. A Monte Carlo simulation is carried out to study the nature of LS and Hurwicz estimators when the innovations follow GED. A comparison study of bias, mean square error and Mean

Absolute Deviation(MAD) of $\hat{\phi}_{Hur}$ and $\hat{\phi}_{LS}$ is done. The performance of two estimators are analysed with respect to Pitman-Closeness Criterion (PCC) (Wenzel (2002)). Also, we study the coverage level of 95% and 90% bootstrap prediction intervals and length of interval when using the LS estimate and Hurwicz estimator.

Definitions

Hurwicz Estimator:- An estimator $\theta = \theta^*(x)$ of a real-valued parameter θ is called median-unbiased if Prob$[\theta^*(X) < \theta|\theta]$ =Prob $[\theta^*(X) > \theta|\theta]$ for each θ; that is, if for each θ, the median of the estimator's distribution is θ (Birnbaum (1964)).

A random variable X is said to follow GED(μ, σ, β) if its probability density function is given by (3.1). The properties of GED are discussed in Chapter 3 and the references therein.

In Section 4.2, we give the details of an $AR(1)$ model with GED innovations and the Hurwicz's estimation method. Section 4.3 deals with the asymptotic properties of Hurwicz's estimator for the GED-AR(1) model. In Section 4.4, the bootstrap prediction interval for the GED-AR(1) model is discussed. The simulation studies are carried out in Section 4.5 to check the performance of the proposed estimator. Also, simulation studies are done to compare the performance of LS and Hurwicz estimator. Data analysis is considered in Section 4.6.

4.2 GED-AR(1) model and Hurwicz Estimation

Let $\{Z_t\}$ be a stationary AR(1) model defined by

$$Z_t = \phi Z_{t-1} + a_t, \qquad t = 1, 2, \ldots, \tag{4.1}$$

where $|\phi| < 1$. Here we consider the innovation sequence $\{a_t\}$ as independent and identically distributed GED with common probability density function defined as in (3.2). The properties of autoregressive model with GED innovations are detailed in Section 3.3.

The joint density function of the innovation random variables $a_1, a_2, \ldots, a_n$ is given by

$$f(a_1, a_2, \ldots, a_n) = \frac{\beta^n}{2^n \sigma^n \Gamma(1/\beta)^n} \exp\left(-\sum_{t=1}^{n} \left|\frac{a_t}{\sigma}\right|^{\beta}\right). \tag{4.2}$$

The conditional likelihood function can be written as (Balakrishna and Sri Ranganath (2015))

$$L(\phi, \sigma, \beta) = \frac{\beta^n}{2^n\, \sigma^n\, \Gamma(1/\beta)^n} \exp\left\{-\frac{1}{\sigma^\beta}\left(\sum_{t=1}^{n} |z_t - \phi z_{t-1}|^{\beta}\right)\right\}, \tag{4.3}$$

which can be expressed as

$$L(\phi, \sigma, \beta) = \frac{\beta^n}{2^n\, \sigma^n\, \Gamma(1/\beta)^n} \exp\left\{-\frac{1}{\sigma^\beta}\left(\sum_{t=1}^{n} |z_{t-1}|^{\beta}\left|\frac{z_t}{z_{t-1}} - \phi\right|^{\beta}\right)\right\}. \tag{4.4}$$

Here we propose the Hurwicz estimator to estimate ϕ. We take

$$\hat{\phi}_{Hur} = med\left(\frac{Z_2}{Z_1}, \frac{Z_3}{Z_2}, \ldots, \frac{Z_n}{Z_{n-1}}\right) \tag{4.5}$$

as an estimator of ϕ where $med(\xi_1, \xi_2, \ldots, \xi_n)$ denotes the sample median of the observations. For further details on Hurwicz estimator refer Zielinski (1999). On replacing ϕ by its Hurwicz estimator the estimators of σ and β are obtained by the method of moments. Also, we can obtain $\hat{\sigma}$ and $\hat{\beta}$ by solving the likelihood function iteratively by taking the Hurwicz estimator as $\hat{\phi}$.

4.3 Asymptotic Properties

The asymptotic distribution of the Hurwicz estimator for the coefficient of autoregression in a linear process with innovations belonging to the domain of attraction of an α-stable law $(1 < \alpha < 2)$ is studied by Berkoun and Fellag (2011). Here we study the asymptotic distribution of Hurwicz estimator when the innovations follow GED. The main result in Berkoun and Fellag (2011) can be applied to the model (4.1) to establish the asymptotic property, provided the linear process in (4.1) satisfy some necessary conditions.

Let $U_t = \frac{Z_{t+1}}{Z_t}$, $t = 1, 2, \ldots, n$, and let G be their common distribution function with density function g. Let $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ be the corresponding ordered random variables. Let u_p be the pth quantile of G defined by $u_p = G^{-1}(p) = inf\{u : G(u) \geq p\}$. We define the pth sample quantile by $G_n^{-1}(p) = inf\{u : G_n(u) \geq p\}$.

Then

$$G_n{}^{-1}(p) = \hat{U}_{(np)} = \begin{cases} U_{(np)}, & \text{if np is an integer,} \\ U_{([np]+1)} & \text{if not,} \end{cases}$$

where $[np]$ denotes the integer part of np and G_n is the empirical distribution based on the random variables $U_1, U_2, \ldots, U_n$. In particular, $\phi_{Hur} = med\left(\frac{Z_2}{Z_1}, \frac{Z_3}{Z_2}, \ldots, \frac{Z_n}{Z_{n-1}}\right) = G_n{}^{-1}(\frac{1}{2})$, which we denote by $\hat{U}_{(np)}$. (Here p=1/2.)

Let $Y_t = p - I_{(U_t \leq u_p)}$. Also let $S_n = \sum_{t=1}^{n} Y_t$ and $\sigma_n^2 = V(n^{-1/2} S_n)$. Also, consider the condition (A) given as follows

$$\left.\begin{aligned} & g = dG \text{ is bounded in some neighbourhood } V_0\, of\; u_p \text{ with } 0 < u_p < \infty \\ & \quad \text{and } 0 < g(u_p) < \infty, \\ & g' \text{ is bounded in} V_0. \end{aligned}\right\} \quad (A)$$

Result 1: Assume condition (A) holds. Then,

$$(\hat{U}_{(np)} - u_p)g(u_p) = (p - G_n(u_p)) + O_{a.s}(n^{\frac{-3}{4+\delta}} log(n)), (\delta > 0).$$

In addition, if $\inf_n \sigma_n^2 > 0$, then

$$\frac{\sqrt{n}(\hat{U}_{(np)} - u_p)g(u_p)}{\sigma_n} \xrightarrow{L} N(0,1),$$

where $\sigma_n^2 = E(Y_1^2) + 2\sum_{k=1}^{n-1}(1 - \frac{k}{n})E(Y_1 Y_{1+k})$.

In order to apply the Result 1, the following conditions should be satisfied

i) the process defined in (4.1) satisfies the strong mixing property.

ii) The Bahadur representation of sample quantiles for strong mixing random variables should hold for (4.1).

In particular, we need to verify whether condition (A) holds. The necessary and sufficient conditions for a linear process to be strong mixing is developed in Gorodestkii (1977), Withers (1981), Andrews (1983) and Athreya and Pantula (1986). By verifying the conditions (see notes at the end of this Chapter) of Athreya and Pantula (1986) we show that the process defined in (4.1) satisfies the strong mixing property. Now, to show that the Bahadur representation of sample quantiles for strong mixing rv holds for (4.1), we need to verify condition (A). For that, we need to obtain the density function of ratio of $U_t = Z_{t+1}/Z_t$, where Z_t is the process defined in (4.1). The process Z_t defined by (4.1) has no explicit expression for the stationary marginal distribution. However, Niehsen (1999) has approximated the pdf of first order AR process driven by generalized Gaussian white noise by a generalized Gaussian pdf. However, the ratio of two correlated GED has no explicit known forms. So, we analyse the distribution of $\hat{\phi}_{Hur}$ empirically. We carry out a Monte Carlo simulation to study the empirical distribution of $\hat{\phi}_{Hur}$. A parametric bootstrap procedure is carried out to check the performance of the $\hat{\phi}_{Hur}$. The steps in this procedure are indicated below:

1. We create the bootstrap sample by first estimating the residuals $\hat{a}_t$.
2. For that, obtain residuals $\hat{a}_t = z_t - \hat{\phi} z_{t-1}$ by substituting $\hat{\phi}_{Hur}$, the Hurwicz's estimator in (4.1).
3. Obtain n independent replicate samples from $\hat{a}_t$, which is the bootstrap sample $(\hat{a_1}^*, \hat{a_2}^*, \ldots, \hat{a_n}^*)$.
4. For each sample obtain the Hurwicz's estimator $\hat{\phi}^*_{Hur}$.

The histogram of the bootstrap distribution of Hurwicz's estimator are plotted and the normality is confirmed by Jarque-Bera test. The Figure 4.1 corresponds to the simulated bootstrap distributions .

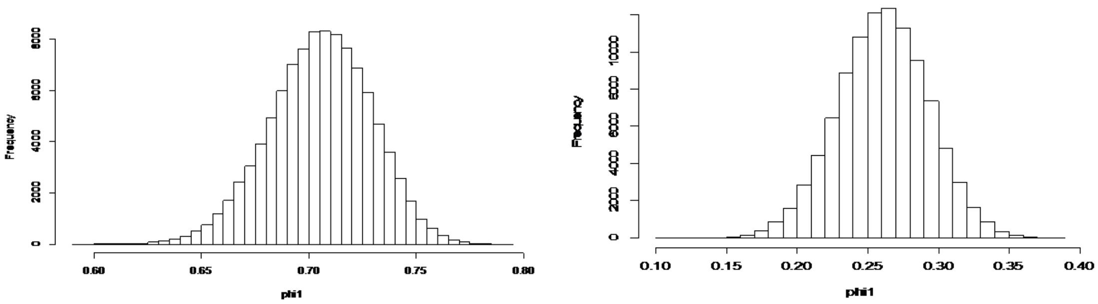

FIGURE 4.1: Histogram of the bootstrap distribution of ϕ_{Hur} for $\phi = 0.7$(left panel) and $\phi = 0.3$(right panel).

Hence the empirical distribution of Hurwicz estimator is analysed using the Monte Carlo simulation and found to be normal.

Mean Absolute Deviation and Pitman-Closeness Criterion

Here we compare the proposed estimator with the LS estimate. A comparison of bias, MSE and MAD are done. Also we analyse the performance of two estimators with respect to PCC.

We define **Mean Absolute Deviation** of $\hat{\phi}_{Hur}$ and $\hat{\phi}_{LS}$ as follows :

$$\hat{\phi}_{LS} = E_{\phi}|\hat{\phi}_{LS} - \hat{\phi}| \qquad \text{and} \qquad \hat{\phi}_{Hur} = E_{\phi}|\hat{\phi}_{Hur} - \hat{\phi}|.$$

Pitman-Closeness:- Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimators for unknown parameter $\theta \in \Theta$, where Θ denotes the parameter space. Then $\hat{\theta}_1$ is Pitman-closer(with respect ot θ) than $\hat{\theta}_2$ if and only if

$$P\{|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|\} \geq \frac{1}{2},$$

with strict inequality for at least one $\theta \in \Theta$.

4.4 Bootstrap Prediction Interval for GED-AR(1) model

Forecasting a time series is one of the main objectives of analysing it. By observing up to time t, the observation at a future time should be predictable. The uncertainty of the predicted value is indicated by computing the prediction interval. Due to the dependency and non-Gaussian nature, some form of resampling techniques are to be implemented to construct prediction intervals. We follow the algorithm of Pan and Politis (2016) for constructing bootstrap prediction interval. The pseudo- series can be generated by either a forward or backward bootstrap, using either fitted or predictive residuals. Here, by simulation we construct forward bootstrap prediction intervals using fitted residuals algorithm to construct bootstrap prediction interval. The detailed algorithm is given below.

Algorithm : Forward bootstrap with fitted residuals.

1. Use all observations $z_1, \ldots, z_n$ to obtain the LS estimator and Hurwicz estimator $\hat{\phi}_{Hur}$.

2. For $t = 2, 3, \ldots, n$, compute the fitted value and fitted residuals :

   $\hat{z}_t = \hat{\phi}_1 z_{t-1} \qquad \hat{a}_t = z_t - \hat{z}_t.$

3. Center the fitted residuals : let $z_t = \hat{a}_t - \bar{\hat{a}}$ for $t = 2, 3, \ldots, n$ and $\bar{\hat{a}} = (n-1)^{-1} \sum_2^n \hat{a}_t$; let the empirical density of z_t be denoted by $\hat{F}_n$.

    (i) Draw bootstrap pseudo-residuals $\{a_t^*, t \geq 1\}$ i.i.d from $\hat{F}_n$.

    (ii) To ensure stationarity of bootstrap series, generate $n + m$ pseudo data for some large positive m and then discard the first m data. Let (u_1^*) be chosen at random from the $\{(z_1, \ldots, z_n)\}$; then generate $\{u_t^*, t \geq 2\}$ by recursion :
    $u_t^* = \hat{\phi}_1 u_{t-1}^* + \hat{a}^*$, for $t = 2, 3, \ldots, n+m$. Then define $z_t^* = u_{m+t}^*$ for $t = 1, 2, \ldots, n$.

    (iii) Based on the pseudo-data $\{z_1^*, z_2^*, \ldots, z_n^*\}$, re-estimate the coefficient ϕ by the LS estimator/Hurwicz estimator. Then compute the future bootstrap one-step ahead predicted value $\hat{z}_{tn+1}^*$ by recursion:
    $\hat{z}_{n+1}^* = \hat{\phi}_1{}^* \hat{z}_n{}^*$ where $\hat{z}_{n+1}^* = z_n$.

    (iv) In order to conduct conditionally validate predictive inference, re-define the last observation to match the original observed value, i.e., let $z_n^* = z_n$. Then generate the future observation z_{n+1}^* by the recursion :
    $z_{n+1}^* = \hat{\phi} z_n^* + a_n^*$.

    (v) Calculate a bootstrap root replicate as $z_{n+1}^* - \hat{z}_{n+1}^*$.

4. Steps (i) to (v) above are repeated B times, and the B bootstrap replicates are collected in the form of an empirical distribution whose $\alpha - quantile$ is denoted $q(\alpha)$.

5. Compute the one step ahead predicted future value $\hat{z}_{n+1}$ by the following recursion :
    $\hat{z}_{n+1} = \hat{\phi} \hat{z}_n$ where $\hat{z}_n = z_n$.

6. Construct the $(1-\alpha)100\%$ equal-tailed prediction interval for Z_{n+1} as $[\hat{z}_{n+1} + q(\alpha/2), \hat{z}_{n+1} + q(1-\alpha/2)]$.

4.5 Simulation study

A simulation study is carried out to assess the performance of the estimation procedure with the sample size 200 and 500. We consider different specified values for $\phi(= 0.3, 0.5, 0.75)$ and $\beta(= 1.3, 1.5, 1.7, 1.9, 2.1, 2.3)$. The value of σ is taken to be 1.5. This was repeated 100 times. A comparison of LS and the proposed estimator is done. The average and the mean square error of the estimates are presented in Table 4.1 and Table 4.2. Also the corresponding bias is shown in bold fonts below each average estimate and mse. The bias is lower for Hurwicz estimator for different values of ϕ.

4.5.1 Comparison of Least Square and Hurwicz estimators

A simulation study is carried out to compare the performance of ϕ_{LS} and ϕ_{Hur} estimators for the AR(1) model with GED innovations. We consider the simulation study for β values(1 , 1.1) for sample size 50 and 100. Also, a comparison of the bias, Mean square error is also done. The results are given in Table 4.3-4.6 and it can be concluded that the proposed estimator performs better. The performance of both the estimators is also checked through PCC and mean absolute deviation. The computational results are shown in Table 4.7 and Table 4.8. From the results we can observe that MAD_{Hur} values are small for heavy tailed distributions and

TABLE 4.1: AR(1) with GED innovations(Comparison of Hurwicz and LS estimates of ϕ),$\sigma = 1.5$,n=200. Moment estimates of σ and β and their corresponding bias and mse are also obtained.

	LS Estimtes of ϕ and moment estimates of σ and β and their corresponding mse and bias						Hurwicz Estimate of ϕ and moment estimates of σ and β and their corresponding mse and bias					
	$\hat{\phi}$	mse($\hat{\phi}$)	$\hat{\sigma}$	mse($\hat{\sigma}$)	$\hat{\beta}$	mse($\hat{\beta}$)	$\hat{\phi_m}$	mse($\hat{\phi_m}$)	$\hat{\sigma_m}$	mse($\hat{\sigma_m}$)	$\hat{\beta_m}$	mse($\hat{\beta_m}$)
ϕ=0.3	0.28927	0.00256	1.50881	0.02386	1.31056	0.01667	0.29231	0.00152	1.52399	0.02739	1.32343	0.02012
β=1.3	**-0.01073**		**0.00881**		**0.01056**		**-0.00769**		**0.02399**		**0.02343**	
ϕ=0.3	0.28848	0.00199	1.50183	0.01304	1.52112	0.02303	0.30170	0.00328	1.50520	0.01560	1.52813	0.02436
β=1.5	**-0.01152**		**0.00183**		**0.02112**		**0.00170**		**0.00520**		**0.02813**	
ϕ=0.3	0.30442	0.00374	1.49650	0.01119	1.72769	0.02828	0.30264	0.00198	1.50948	0.01056	1.74332	0.02747
β=1.7	**0.00442**		**-0.00350**		**0.02769**		**0.00264**		**0.00948**		**0.04332**	
ϕ=0.3	0.29001	0.00423	1.49613	0.01027	1.91420	0.04089	0.29742	0.00182	1.50293	0.00713	1.92169	0.03377
β=1.9	**-0.00999**		**-0.00387**		**0.01420**		**-0.00582**		**0.00293**		**0.02169**	
ϕ=0.3	0.29050	0.02903	1.51526	0.00766	2.18256	0.06405	0.29373	0.04648	1.51114	0.00830	2.19261	0.09635
β=2.1	**-0.00950**		**0.01526**		**0.08256**		**-0.00627**		**0.01114**		**0.09261**	
ϕ=0.3	0.29325	0.00214	1.50127	0.00578	2.33204	0.05845	0.29959	0.00511	1.50057	0.00621	2.36002	0.08334
β=2.3	**-0.00675**		**0.00127**		**0.03204**		**-0.00041**		**0.00057**		**0.06002**	
ϕ=0.5	0.49706	0.00132	1.50475	0.02531	1.30680	0.01548	0.49740	0.00196	1.52496	0.03128	1.32911	0.02011
β=1.3	**-0.00294**		**0.00475**		**0.00680**		**-0.00260**		**0.02496**		**0.02911**	
ϕ=0.5	0.50513	0.00280	1.51052	0.01385	1.53229	0.02157	0.50028	0.00158	1.54177	0.01973	1.57735	0.03466
β=1.5	**0.00513**		**0.01052**		**0.03229**		**0.00028**		**0.04177**		**0.07735**	
ϕ=0.5	0.49482	0.02629	1.48209	0.01121	1.70751	0.02649	0.49573	0.03968	1.51595	0.01951	1.74798	0.04468
β=1.7	**-0.00518**		**-0.01791**		**0.00751**		**-0.00427**		**0.01595**		**0.04798**	
ϕ=0.5	0.50341	0.02464	1.52057	0.00876	1.96695	0.04571	0.50315	0.04221	1.52873	0.00974	2.00679	0.09052
β=1.9	**0.00341**		**0.02057**		**0.06695**		**0.00315**		**0.02873**		**0.10679**	
ϕ=0.5	0.49180	0.00171	1.49510	0.00921	2.14383	0.06579	0.49648	0.00398	1.51071	0.01655	2.22519	0.12676
β=2.1	**-0.00820**		**-0.00490**		**0.04383**		**-0.00352**		**0.01071**		**0.12519**	
ϕ=0.5	0.49180	0.03753	1.51745	0.00815	2.37747	0.08475	0.49559	0.02600	1.52181	0.00986	2.44762	0.14420
β=2.3	**-0.00820**		**0.01745**		**0.07747**		**-0.00441**		**0.02181**		**0.14762**	
ϕ=0.75	0.74437	0.00143	1.50020	0.01817	1.31840	0.01429	0.74548	0.00101	1.58547	0.08127	1.41109	0.07425
β=1.3	**-0.00563**		**0.00020**		**0.01840**		**-0.00452**		**0.08547**		**0.11109**	
ϕ=0.75	0.74848	0.01911	1.49778	0.01047	1.50205	0.01394	0.74962	0.02721	1.52161	0.08192	1.61713	0.18651
β=1.5	**0.00152**		**-0.00222**		**0.00205**		**-0.00038**		**0.02161**		**0.11713**	
ϕ=0.75	0.74412	0.02324	1.50007	0.01005	1.73675	0.02853	0.74469	0.03494	1.47725	0.05545	1.83647	0.19672
β=1.7	**-0.00588**		**0.00007**		**0.03675**		**-0.00531**		**-0.02275**		**0.13647**	
ϕ=0.75	0.75563	0.00229	1.49010	0.00830	1.89984	0.03050	0.74598	0.00089	1.48244	0.05637	2.05795	0.21894
β=1.9	**0.00563**		**-0.00990**		**-0.00016**		**-0.00402**		**-0.01756**		**0.15795**	
ϕ=0.75	0.74026	0.00080	1.49254	0.00652	2.13180	0.05180	0.74045	0.00210	1.45297	0.03592	2.09974	0.23261
β=2.1	**-0.00974**		**-0.00746**		**0.03180**		**-0.00955**		**-0.04703**		**-0.00026**	
ϕ=0.75	0.74258	0.00090	1.50754	0.00534	2.39662	0.07955	0.74649	0.00205	1.46049	0.02699	2.33360	0.25316
β=2.3	**-0.00742**		**0.00754**		**0.09662**		**-0.00351**		**-0.03951**		**0.03360**	

for large ϕ values. Also, the PCC suggests that ϕ_{Hur} performs better than ϕ_{LS} for heavy tailed distributions.

Monte Carlo studies for bootstrap prediction interval

The performance of the Forward bootstrap with fitted residuals (Ff) method is carried out as follows:

TABLE 4.2: AR(1) with GED innovations (Comparison of Hurwicz and LS estimates of ϕ), $\sigma = 1.5$, n=500. Moment estimates of σ and β and their corresponding bias and mse are also obtained.

	LS Estimate of ϕ and moment estimates of σ and β and their corresponding mse and bias						Hurwicz Estimate of ϕ and moment estimates of σ and β and their corresponding mse and bias					
	$\hat{\phi}$	mse($\hat{\phi}$)	$\hat{\sigma}$	mse($\hat{\sigma}$)	$\hat{\beta}$	mse($\hat{\beta}$)	$\hat{\phi_m}$	mse($\hat{\phi_m}$)	$\hat{\sigma_m}$	mse($\hat{\sigma_m}$)	$\hat{\beta_m}$	mse($\hat{\beta_m}$)
$\phi = 0.3$	0.30096	0.03388	1.51348	0.01713	1.31008	0.07960	0.30031	0.04056	1.53652	0.01479	1.33201	0.09496
$\beta = 1.3$	**0.00095**		**0.01348**		**0.01008**		**0.00031**		**0.03652**		**0.03201**	
$\phi = 0.3$	0.29964	0.02874	1.50140	0.08464	1.51244	0.09585	0.29651	0.04457	1.51197	0.08906	1.52397	0.09653
$\beta = 1.5$	**-0.00036**		**0.00140**		**0.01244**		**-0.00349**		**0.01197**		**0.02397**	
$\phi = 0.3$	0.30401	0.04331	1.49526	0.07499	1.70055	0.10997	0.30113	0.02782	1.49361	0.07885	1.70509	0.12396
$\beta = 1.7$	**0.00401**		**-0.00474**		**0.00055**		**0.00113**		**-0.00639**		**0.00509**	
$\phi = 0.3$	0.29185	0.05631	1.51394	0.06759	1.90866	0.12254	0.29797	0.03471	1.51505	0.07443	1.92155	0.14510
$\beta = 1.9$	**-0.00815**		**0.01394**		**0.00866**		**-0.00203**		**0.01505**		**0.02155**	
$\phi = 0.3$	0.30486	0.00388	1.50355	0.06237	2.11415	0.15970	0.30114	0.00207	1.50547	0.06254	2.12417	0.16483
$\beta = 2.1$	**0.00486**		**0.00355**		**0.01415**		**0.00114**		**0.00547**		**0.02417**	
$\phi = 0.3$	0.29738	0.04316	1.50211	0.05765	2.33637	0.20513	0.29895	0.03204	1.50938	0.04836	2.35222	0.18624
$\beta = 2.3$	**-0.00262**		**0.00211**		**0.03637**		**-0.00105**		**0.00938**		**0.05222**	
$\phi = 0.5$	0.50088	0.02860	1.49003	0.10182	1.29445	0.07668	0.49915	0.03263	1.50762	0.14455	1.30934	0.11467
$\beta = 1.3$	**0.00088**		**-0.00997**		**-0.00555**		**-0.00085**		**0.00762**		**0.00934**	
$\phi = 0.5$	0.49596	0.02660	1.48786	0.08741	1.49610	0.10156	0.49792	0.03611	1.50987	0.11406	1.52738	0.14168
$\beta = 1.5$	**-0.00404**		**-0.01214**		**-0.00390**		**-0.00208**		**0.00987**		**0.02738**	
$\phi = 0.5$	0.49542	0.00127	1.48967	0.07579	1.69764	0.11283	0.49590	0.00324	1.52031	0.09461	1.75167	0.15782
$\beta = 1.7$	**-0.00458**		**-0.01033**		**-0.00236**		**-0.00410**		**0.02031**		**0.05167**	
$\phi = 0.5$	0.49969	0.00162	1.50508	0.06004	1.90989	0.12670	0.50021	0.00371	1.51978	0.09326	1.96221	0.16221
$\beta = 1.9$	**-0.00031**		**0.00508**		**0.00989**		**0.00021**		**0.01978**		**0.06221**	
$\phi = 0.5$	0.49920	0.04455	1.49986	0.06122	2.11024	0.15600	0.49959	0.02694	1.49768	0.07960	2.11993	0.19178
$\beta = 2.1$	**-0.00080**		**-0.00014**		**0.01024**		**-0.00041**		**-0.00232**		**0.01993**	
$\phi = 0.5$	0.49364	0.00189	1.50188	0.06141	2.32702	0.18366	0.49628	0.00355	1.50921	0.07577	2.36201	0.21194
$\beta = 2.3$	**-0.00636**		**0.00188**		**0.02702**		**-0.00372**		**0.00921**		**0.06201**	
$\phi = 0.75$	0.74616	0.02216	1.49884	0.09996	1.30155	0.08151	0.74911	0.02578	1.57540	0.10277	1.40967	0.08524
$\beta = 1.3$	**-0.00384**		**-0.00116**		**0.00155**		**-0.00089**		**0.07540**		**0.10967**	
$\phi = 0.75$	0.74866	0.00091	1.51169	0.07850	1.52189	0.09260	0.74996	0.00151	1.54528	0.06421	1.58861	0.13798
$\beta = 1.5$	**-0.00134**		**0.01169**		**0.02189**		**-0.00034**		**0.04528**		**0.08861**	
$\phi = 0.75$	0.74688	0.00083	1.51447	0.07207	1.73709	0.12410	0.75025	0.00194	1.48689	0.09310	1.76582	0.14806
$\beta = 1.7$	**-0.00312**		**0.01447**		**0.03709**		**0.00025**		**-0.01311**		**0.06582**	
$\phi = 0.75$	0.75302	0.02087	1.49310	0.06170	1.89670	0.12108	0.75252	0.03323	1.51009	0.08272	2.01633	0.17967
$\beta = 1.9$	**0.00302**		**-0.00690**		**-0.00330**		**0.00252**		**0.01009**		**0.11633**	
$\phi = 0.75$	0.74301	0.03523	1.49867	0.06043	2.11514	0.14555	0.74553	0.02034	1.47415	0.14465	2.12289	0.19318
$\beta = 2.1$	**-0.00699**		**-0.00133**		**0.01514**		**-0.00447**		**-0.02585**		**0.02289**	
$\phi = 0.75$	0.74383	0.03673	1.49566	0.05661	2.31394	0.18255	0.74540	0.01797	1.45497	0.15150	2.25700	0.22774
$\beta = 2.3$	**-0.00617**		**-0.00434**		**0.01394**		**-0.00460**		**-0.04503**		**-0.04300**	

1. AR(1) model with iid GED innovations rescaled to unit variance.

2. 500 'true' datasets each of size $n = 50$ or $n = 100$, and for each 'true' dataset creating $B = 1000$ bootstrap pseudo-series.

3. prediction intervals with nominal coverage levels of 95% and 90%.

For the i^{th} 'true' dataset, we use one of the bootstrap methods to create B=1000 sample paths (step 4 of the algorithm), and construct the prediction interval (step 6

TABLE 4.3: Comparison of ϕ_{LS} and ϕ_{Hur} estimators for $\beta = 1$ (Laplace) for n=50. Moment estimates of σ and β and their corresponding bias and mse are also obtained with $\sigma = 1.5$.

ϕ value	Estimate	$\hat{\phi}$ (Bias,Mse)	$\hat{\sigma}$ (Bias,Mse)	$\hat{\beta}$ (Bias,Mse)
0.3	ϕ_{LS}	0.27506	1.62029	1.08701
		(-0.02493, 0.00960)	(0.12029, 0.14927)	(0.08700, 0.04351)
	ϕ_{Hur}	0.29011	1.59284	1.06976
		(-0.00988, 0.01301)	(0.09284, 0.14900)	(0.06976, 0.04039)
0.5	ϕ_{LS}	0.46884	1.62219	1.08752
		(-0.03116,0.00859)	(0.12219,0.151)	(0.08752, 0.04389)
	ϕ_{Hur}	0.48953	1.58542	1.06615
		(-0.01047,0.00913)	(0.08541,0.13866)	(0.06615, 0.03728)
0.7	ϕ_{LS}	0.66134	1.63075	1.09088
		(-0.03866,0.00717)	(0.13075,0.15639)	(0.09087, 0.04536)
	ϕ_{Hur}	0.68471	1.58805	1.06759
		(-0.01529,0.00622)	(0.08805,0.14672)	(0.06759, 0.03892)
0.8	ϕ_{LS}	0.75608	1.64096	1.09469
		-0.04392, 0.00647	0.14096, 0.16325	0.09469, 0.04684
	ϕ_{Hur}	0.78637	1.59238	1.07013
		-0.01363, 0.00495	0.09238, 0.15783	0.07013, 0.04195
0.9	ϕ_{LS}	0.84833	1.67152	1.08250
		(-0.05167,0.00598)	(0.12836,0.19517)	(0.10644, 0.05101)
	ϕ_{Hur}	0.88018	1.59586	1.06973
		(-0.01982,0.00389)	(0.09586,0.15621)	(0.06973, 0.03928)
0.95	ϕ_{LS}	0.89176	1.74996	1.13704
		(-0.05823,0.006004)	(0.24995,0.25443)	(0.13703,0.06765)
	ϕ_{Hur}	0.92699	1.62837	1.08706
		(-0.02300,0.00301)	(0.09635,0.16623)	(0.08249, 0.04619)
-0.5	ϕ_{LS}	-0.49234	1.61994	1.08763
		(0.00765,0.00775)	(0.11994,0.15198)	(0.08736, 0.04477)
	ϕ_{Hur}	-0.49180	1.58514	1.06708
		(-0.00719,0.01043)	(0.08514,0.14889)	(0.06708, 0.04092)
-0.95	ϕ_{LS}	-0.92272	1.63192	1.09170
		(0.02727,0.00231)	(0.13192,0.16535)	(0.09169, 0.04715)
	ϕ_{Hur}	-0.94349	1.58152	1.06472
		(0.00650,0.00145)	(0.0815,0.14414)	(0.06472, 0.03824)

TABLE 4.4: Comparison of ϕ_{LS} and ϕ_{Hur} estimators for $\beta = 1.1$ for n=50. Moment estimates of σ and β and their corresponding bias and mse are also obtained with $\sigma = 1.5$.

ϕ value	Estimate	$\hat{\phi}$ (Bias,Mse)	$\hat{\sigma}$ (Bias,Mse)	$\hat{\beta}$ (Bias,Mse)
0.3	ϕ_{LS}	0.28113	1.56547	1.18335
		(-0.01887,0.01071)	(0.06547,0.13244)	(0.08335, 0.06379)
	ϕ_{Hur}	0.29695	1.54506	1.16544
		(-0.00305,0.01467)	(0.04506,0.13128)	(0.06544, 0.05842)
0.5	ϕ_{LS}	0.47270	1.56833	1.18504
		(-0.0273,0.0097)	(0.06833,0.13666)	(0.08504, 0.06495)
	ϕ_{Hur}	0.48864	1.54150	1.16383
		(-0.01136,0.01154)	(0.0415,0.13453)	(0.06383, 0.05985)
0.7	ϕ_{LS}	0.66341	1.57404	1.18746
		(-0.03659,0.00779)	(0.07404,0.14019)	(0.08745, 0.06518)
	ϕ_{Hur}	0.68519	1.54422	1.16445
		(-0.01481,0.00817)	(0.04422,0.13445)	(0.06445, 0.05711)
0.8	ϕ_{LS}	0.75818	1.58258	1.19113
		-0.04182, 0.00654	0.08258, 0.14165	0.09113, 0.06509
	ϕ_{Hur}	0.78382	1.54138	1.16208
		-0.01618, 0.00537	0.04138, 0.13244	0.06208, 0.05397
0.9	ϕ_{LS}	0.85145	1.60250	1.19945
		(-0.04855,0.00537)	(0.1025,0.15264)	(0.09945, 0.06797)
	ϕ_{Hur}	0.88019	1.54803	1.16508
		(-0.01981,0.00371)	(0.04803,0.13974)	(0.06508, 0.0561)
0.95	ϕ_{LS}	0.89485	1.65685	1.22375
		(-0.05515,0.00525)	(0.15685,0.1969)	(0.12375, 0.08205)
	ϕ_{Hur}	0.92727	1.56513	1.17169
		(-0.02273,0.00303)	(0.06513,0.16527)	(0.07169, 0.06347)
-0.5	ϕ_{LS}	-0.48621	1.55773	1.17727
		(0.01379,0.00778)	(0.05773,0.1277)	(0.07727, 0.06046)
	ϕ_{Hur}	-0.49152	1.52891	1.15530
		(0.00848,0.01008)	(0.02891,0.12403)	(0.05529, 0.05542)
-0.95	ϕ_{LS}	-0.92197	1.56307	1.17835
		(0.02803,0.00248)	(0.06307,0.12813)	(0.07835, 0.05711)
	ϕ_{Hur}	-0.94160	1.52709	1.15460
		(0.0084,0.00197)	(0.02709,0.12342)	(0.0546, 0.05317)

Table 4.5: Comparison of ϕ_{LS} and ϕ_{Hur} estimators for $\beta = 1$ (Laplace) for n=100. Moment estimates of σ and β and their corresponding bias and mse are also obtained with $\sigma = 1.5$.

ϕ value	Estimate	estimate (Bias,Mse)	estimate (Bias,Mse)	estimate (Bias,Mse)
0.3	ϕ_{LS}	0.27656	1.58902	1.12418
		(-0.02344, 0.03225)	(0.08902, 0.30401)	(0.12418, 0.18545)
	ϕ_{Hur}	0.25430	1.63388	1.15717
		(-0.0457, 0.02038)	(0.13388, 0.30242)	(0.15717, 0.20787)
0.5	ϕ_{LS}	0.43858	1.64947	1.16266
		(-0.06142, 0.0202)	(0.14947, 0.31448)	(0.16266, 0.19906)
	ϕ_{Hur}	0.46461	1.58145	1.11318
		(-0.03539, 0.02659)	(0.08145, 0.29944)	(0.11318, 0.1526)
0.7	ϕ_{LS}	0.62087	1.66953	1.16435
		(-0.07913, 0.01963)	(0.16953, 0.3428)	(0.16435, 0.193)
	ϕ_{Hur}	0.65551	1.58727	1.12370
		(-0.04449, 0.01838)	(0.08727, 0.31697)	(0.1237, 0.19939)
0.8	ϕ_{LS}	0.70916	1.70146	1.18090
		(-0.09084, 0.01993)	(0.20146, 0.38924)	(0.1809, 0.20907)
	ϕ_{Hur}	0.75682	1.59381	1.12303
		(-0.04318, 0.01506)	(0.09381, 0.33258)	(0.12303, 0.20964)
0.9	ϕ_{LS}	0.79115	1.84202	1.24189
		(-0.10885, 0.02196)	(0.34202, 0.67816)	(0.24189, 0.29328)
	ϕ_{Hur}	0.85158	1.65334	1.14630
		(-0.04842, 0.01296)	(0.15334, 0.44823)	(0.1463, 0.21187)
0.95	ϕ_{LS}	0.82715	2.14973	1.40998
		(-0.12285, 0.02386)	(0.64973, 1.66405)	(0.40998, 0.77704)
	ϕ_{Hur}	0.89907	1.78592	1.20089
		(-0.05093, 0.01088)	(0.28592, 0.89797)	(0.20089, 0.33918)
-0.5	ϕ_{LS}	-0.48230	1.61986	1.14510
		(0.0177, 0.01558)	(0.11986, 0.30722)	(0.14509, 0.16945)
	ϕ_{Hur}	-0.49140	1.56318	1.09800
		(0.0086, 0.02388)	(0.06318, 0.29947)	(0.098, 0.13612)
-0.95	ϕ_{LS}	-0.89647	1.65032	1.14679
		(0.05353, 0.00765)	(0.15032, 0.3111)	(0.14679, 0.16667)
	ϕ_{Hur}	-0.93399	1.56044	1.10494
		(0.01601, 0.00517)	(0.06044, 0.29993)	(0.10494, 0.18753)

TABLE 4.6: Comparison of ϕ_{LS} and ϕ_{Hur} estimators for $\beta = 1.1$ for n=100. Moment estimates of σ and β and their corresponding bias and mse are also obtained with $\sigma = 1.5$.

ϕ value	Estimate	estimate (Bias,Mse)	estimate (Bias,Mse)	estimate (Bias,Mse)
0.3	ϕ_{LS}	0.25463 (-0.04537, 0.0192)	1.55685 (0.05685, 0.29644)	1.24514 (0.14514, 0.22563)
	ϕ_{Hur}	0.27208 (-0.02792, 0.02958)	1.52049 (0.02049, 0.30539)	1.20161 (0.10161, 0.17181)
0.5	ϕ_{LS}	0.43887 (-0.06113, 0.01898)	1.55997 (0.05997, 0.29852)	1.23425 (0.13425, 0.18785)
	ϕ_{Hur}	0.46637 (-0.03363, 0.02317)	1.51003 (0.01003, 0.29685)	1.19208 (0.09208, 0.15755)
0.7	ϕ_{LS}	0.61941 (-0.08059, 0.01883)	1.57469 (0.07469, 0.3069)	1.24504 (0.14504, 0.20597)
	ϕ_{Hur}	0.65559 (-0.04441, 0.02087)	1.51645 (0.01645, 0.2933)	1.20215 (0.10215, 0.19919)
0.8	ϕ_{LS}	0.70679 (-0.09321, 0.01948)	1.60490 (0.1049, 0.33339)	1.25330 (0.1533, 0.20906)
	ϕ_{Hur}	0.75194 (-0.04806, 0.0169)	1.53778 (0.03778, 0.31682)	1.20540 (0.1054, 0.17863)
0.9	ϕ_{LS}	0.78991 (-0.11009, 0.02096)	1.70854 (0.20854, 0.49688)	1.32075 (0.22075, 0.3873)
	ϕ_{Hur}	0.85225 (-0.04775, 0.01128)	1.57095 (0.07095, 0.388)	1.22408 (0.12408, 0.20716)
0.95	ϕ_{LS}	0.82883 (-0.12117, 0.02227)	1.91414 (0.41414, 1.01009)	1.47378 (0.37378, 1.10653)
	ϕ_{Hur}	0.88860 (-0.0614, 0.01285)	1.69097 (0.19097, 0.84172)	1.30337 (0.20337, 0.69424)
-0.5	ϕ_{LS}	-0.49212 (0.00788, 0.01285)	1.55325 (0.05325, 0.2902)	1.23769 (0.13769, 0.21748)
	ϕ_{Hur}	-0.50251 (-0.00251, 0.02268)	1.49774 (-0.00226, 0.28373)	1.18557 (0.08557, 0.16881)
-0.95	ϕ_{LS}	-0.89930 (0.0507, 0.00628)	1.57680 (0.0768, 0.29511)	1.24898 (0.14898, 0.25909)
	ϕ_{Hur}	-0.93806 (0.01194, 0.00433)	1.50354 (0.00354, 0.28364)	1.19371 (0.09371, 0.18781)

TABLE 4.7: Simulated values of MAD_{LS},MAD_{Hur} and PCC for GED with $\beta(=1, 1.1, 1.3, 1.5, 1.7, 2)$ values and n=50.

$\beta=1$	MAD_{LS}			MAD_{Hur}			PCC		
ϕ	ϕ	σ	β	ϕ	σ	β	ϕ	σ	β
0.30	**0.11340**	0.43642	0.26277	0.14126	0.43599	0.24947	0.434	0.504	0.514
0.50	**0.11122**	0.44373	0.26516	0.12392	0.43295	0.24031	0.468	0.514	0.546
0.70	0.10626	0.45933	0.26491	**0.10108**	0.44220	0.24956	**0.526**	0.510	0.552
0.80	0.10540	0.48031	0.27639	**0.09073**	0.44686	0.24991	**0.538**	0.546	0.528
0.90	0.11196	0.57471	0.32737	**0.07945**	0.48597	0.26704	**0.644**	0.594	0.606
0.95	0.12287	0.80742	0.46963	**0.07280**	0.60082	0.31708	**0.760**	0.590	0.584
-0.50	**0.09948**	0.43180	0.25549	0.11774	0.42304	0.23182			
-0.95	0.05847	0.43440	0.24560	**0.04788**	0.41926	0.23433			
$\beta=1.1$									
0.30	**0.11340**	0.43642	0.26277	0.14126	0.43599	0.24947	0.434	0.504	0.514
0.50	**0.11122**	0.44373	0.26516	0.12392	0.43295	0.24031	0.468	0.514	0.546
0.70	0.10626	0.45933	0.26491	**0.10108**	0.44220	0.24956	**0.526**	0.510	0.552
0.80	0.10540	0.48031	0.27639	**0.09073**	0.44686	0.24991	**0.538**	0.546	0.528
0.90	0.11196	0.57471	0.32737	**0.07945**	0.48597	0.26704	**0.644**	0.594	0.606
0.95	0.12287	0.80742	0.46963	**0.07280**	0.60082	0.31708	**0.760**	0.590	0.584
-0.50	**0.09948**	0.43180	0.25549	0.11774	0.42304	0.23182			
-0.95	0.05847	0.43440	0.24560	**0.04788**	0.41926	0.23433			
$\beta=1.3$									
0.30	**0.11695**	0.36634	0.37081	0.15397	0.36883	0.34746	0.376	0.498	0.492
0.50	**0.11531**	0.36165	0.37032	0.14819	0.36527	0.35935	0.404	0.488	0.468
0.70	**0.10976**	0.35938	0.37247	0.12468	0.37256	0.37147	0.460	0.468	0.476
0.80	0.10709	0.36186	0.36272	**0.10281**	0.38440	0.37028	**0.520**	0.462	0.486
0.90	0.10975	0.40151	0.42162	**0.09277**	0.41166	0.37533	**0.592**	0.504	0.502
0.95	0.11816	0.51102	0.52916	**0.08231**	0.47874	0.46844	**0.720**	0.516	0.520
-0.50	**0.09580**	0.36850	0.37454	0.12582	0.37257	0.34596			
-0.95	0.05301	0.36309	0.38866	**0.05005**	0.37929	0.35721			
$\beta=1.5$									
0.30	**0.11247**	0.27527	0.38425	0.16322	0.29232	0.36647	0.362	0.438	0.470
0.50	**0.10854**	0.27964	0.40594	0.15219	0.29702	0.39516	0.374	0.470	0.494
0.70	**0.10373**	0.28595	0.42360	0.12812	0.29974	0.39486	0.416	0.446	0.506
0.80	**0.10124**	0.28915	0.43562	0.11147	0.30462	0.38575	0.444	0.468	0.526
0.90	0.10472	0.30497	0.46773	**0.09895**	0.30634	0.40149	**0.528**	0.482	0.520
0.95	0.11353	0.37754	0.56012	**0.08708**	0.34478	0.44431	**0.648**	0.506	0.530
-0.50	**0.09768**	0.27959	0.39763	0.13762	0.29651	0.37097			
-0.95	0.06448	0.28442	0.41504	**0.06039**	0.29377	0.37898			
$\beta=1.7$									
0.30	**0.10762**	0.26592	0.47303	0.17395	0.27940	0.43476	0.298	0.442	0.492
0.50	**0.10462**	0.26476	0.47993	0.15695	0.28021	0.44991	0.306	0.452	0.498
0.70	**0.10152**	0.26460	0.47581	0.12792	0.28168	0.46731	0.362	0.448	0.482
0.80	**0.10040**	0.26793	0.48980	0.11450	0.28189	0.46996	0.436	0.434	0.482
0.90	0.10373	0.28162	0.53893	**0.10107**	0.30213	0.53512	**0.522**	0.480	0.442
0.95	0.11249	0.32987	0.63634	**0.09265**	0.33819	0.59458	**0.638**	0.482	0.464
-0.50	**0.10145**	0.26384	0.45918	0.14640	0.28322	0.45584			
-0.95	**0.05651**	0.27249	0.51240	0.05745	0.29006	0.48862			
$\beta=2$									
0.30	**0.10976**	0.24064	0.76749	0.16948	0.25117	0.70547	0.300	0.424	0.438
0.50	**0.10555**	0.24185	0.75103	0.14979	0.25292	0.70219	0.368	0.456	0.472
0.70	**0.10063**	0.24744	0.73419	0.13521	0.25247	0.67902	0.372	0.452	0.474
0.80	**0.09999**	0.24704	0.69654	0.12218	0.25671	0.68572	0.414	0.462	0.490
0.90	**0.10565**	0.25526	0.70217	0.10734	0.27340	0.63179	0.498	0.460	0.478
0.95	0.11595	0.30091	0.77565	**0.10556**	0.31602	0.65696	**0.588**	0.472	0.498
-0.50	**0.09274**	0.24353	0.82208	0.15317	0.25536	0.66929			
-0.95	**0.05230**	0.24222	0.74135	0.06225	0.25499	0.69239			

TABLE 4.8: Simulated values of MAD_{LS},MAD_{Hur} and PCC for GED with $\beta(=1, 1.1, 1.3, 1.5, 1.7, 2)$ values and n=100.

$\beta=1$	MAD_{LS}			MAD_{Hur}			PCC		
ϕ	ϕ	σ	β	ϕ	σ	β	ϕ	σ	β
0.30	**0.07893**	0.30951	0.15514	0.08722	0.30433	0.14631	0.488	0.508	0.546
0.50	0.07293	0.31208	0.15521	**0.07256**	0.29742	0.14238	**0.514**	0.584	0.596
0.70	0.06483	0.31657	0.15586	**0.05900**	0.30367	0.14518	**0.552**	0.548	0.574
0.80	0.06026	0.32252	0.15760	**0.05290**	0.31093	0.14986	**0.520**	0.540	0.524
0.90	0.05688	0.33008	0.16292	**0.04600**	0.31398	0.14888	**0.572**	0.530	0.532
0.95	0.05920	0.38219	0.18571	**0.03940**	0.33422	0.15835	**0.660**	0.582	0.598
-0.50	**0.06896**	0.31124	0.15614	0.08021	0.30752	0.14847			
-0.95	0.03406	0.32518	0.15886	**0.02828**	0.30437	0.14462			
$\beta=1.1$									
0.30	**0.08029**	0.28954	0.18230	0.09137	0.28773	0.17498	0.454	0.524	0.526
0.50	**0.07552**	0.29426	0.18472	0.08311	0.29164	0.17505	0.452	0.502	0.528
0.70	0.06780	0.29845	0.18595	**0.06412**	0.29454	0.17526	**0.502**	0.524	0.540
0.80	0.06231	0.30006	0.18743	**0.05661**	0.28915	0.17179	**0.528**	0.534	0.574
0.90	0.05608	0.30911	0.19263	**0.04538**	0.29676	0.17592	**0.568**	0.530	0.536
0.95	0.05660	0.33935	0.20773	**0.03890**	0.30741	0.18138	**0.656**	0.572	0.564
-0.50	**0.06959**	0.28277	0.17666	0.07827	0.27589	0.16678			
-0.95	0.03502	0.28470	0.17514	**0.03139**	0.27708	0.16621			
$\beta=1.3$									
0.30	**0.07542**	0.24127	0.21167	0.10299	0.24456	0.21038	0.380	0.480	0.504
0.50	**0.07190**	0.24114	0.21291	0.09123	0.24326	0.20476	0.392	0.490	0.504
0.70	**0.06549**	0.24200	0.21463	0.07360	0.24687	0.20947	0.438	0.516	0.534
0.80	0.06186	0.24238	0.21491	**0.06016**	0.24777	0.20859	**0.502**	0.504	0.536
0.90	0.05699	0.24389	0.21681	**0.05198**	0.24695	0.20926	**0.544**	0.494	0.504
0.95	0.05727	0.26157	0.22909	**0.04237**	0.25779	0.21677	**0.634**	0.476	0.506
-0.50	**0.06455**	0.24279	0.21490	0.09443	0.24914	0.21335			
-0.95	0.03492	0.23842	0.21441	**0.03468**	0.24733	0.20838			
$\beta=1.5$									
0.30	**0.07333**	0.20808	0.25664	0.10334	0.21839	0.26291	0.382	0.460	0.438
0.50	**0.06899**	0.20859	0.25690	0.10119	0.21676	0.25503	0.364	0.466	0.458
0.70	**0.06195**	0.21090	0.25774	0.08704	0.21876	0.25458	0.360	0.466	0.480
0.80	**0.05786**	0.21277	0.25989	0.07450	0.21949	0.25631	0.380	0.452	0.456
0.90	**0.05395**	0.21648	0.25767	0.05822	0.22021	0.24930	0.440	0.478	0.468
0.95	0.05522	0.22863	0.26686	**0.04809**	0.22965	0.25926	**0.564**	0.500	0.496
-0.50	**0.07054**	0.20552	0.25211	0.09442	0.21335	0.25065			
-0.95	0.03689	0.20527	0.25522	**0.03910**	0.21374	0.25150			
$\beta=1.7$									
0.30	**0.07579**	0.19081	0.30438	0.11284	0.19867	0.29862	0.344	0.482	0.498
0.50	**0.07049**	0.19078	0.29677	0.10653	0.19946	0.30751	0.322	0.472	0.468
0.70	**0.06404**	0.18914	0.29481	0.08622	0.20128	0.30136	0.352	0.434	0.454
0.80	**0.06009**	0.18772	0.29283	0.07363	0.20111	0.30332	0.422	0.450	0.460
0.90	**0.05604**	0.18831	0.29488	0.06265	0.19942	0.29522	0.424	0.458	0.486
0.95	0.05689	0.19733	0.29428	**0.05232**	0.20539	0.29903	**0.520**	0.478	0.458
-0.50	**0.07029**	0.18791	0.30177	0.10279	0.19486	0.28851			
-0.95	**0.03539**	0.19042	0.29662	0.04272	0.19678	0.28529			
$\beta=2$									
0.30	**0.07112**	0.16975	0.44157	0.11463	0.17099	0.42056	0.324	0.486	0.482
0.50	**0.06644**	0.17050	0.43812	0.10401	0.17181	0.42494	0.364	0.488	0.492
0.70	**0.06151**	0.17169	0.43729	0.08940	0.17201	0.40820	0.388	0.502	0.542
0.80	**0.05822**	0.17154	0.43717	0.08111	0.17050	0.40258	0.360	0.484	0.490
0.90	**0.05421**	0.17090	0.43766	0.06012	0.17090	0.40898	0.488	0.486	0.508
0.95	0.05449	0.17641	0.43816	**0.05364**	0.18334	0.41726	**0.510**	0.476	0.464
-0.50	**0.06469**	0.16729	0.43827	0.10804	0.17202	0.42061			
-0.95	**0.03567**	0.16432	0.42159	0.04132	0.17182	0.40248			

of the algorithm)$[L_i, U_i]$. To assess the corresponding empirical coverage level (cvr) and average length (len) of the constructed interval, we also generate 1000 one-step ahead future values $Y_{n+1,j} = \hat{\phi}_1 z_{ni} + a_j^*$ for $j = 1, 2, \ldots, 1000$ where $\hat{\phi}_1$ is the estimate from the i^{th} 'true' dataset and z_{ni} is the i^{th} dataset's last value. Then the empirical coverage level and length from the i^{th} dataset are given by $cvr_i = \frac{1}{1000}\sum_{j=1}^{1000} 1_{|L_i,U_i|}(Y_{n+1,j})$ and $len_i = U_i - L_i$ where $1_A(z)$ is the indicator function of set A. The coverage level and length is calculated by the average $\{cvr_i\}$ and $\{len_i\}$ over the 500 'true' datasets, i.e. $cvr = \frac{1}{500}\sum_{i=1}^{500} cvr_i$ and $len = \frac{1}{500}\sum_{i=1}^{500} len_i$. Table 4.9 summarize the empirical coverage level (cvr), average length (len) and standard error associated with each average length (serr) of prediction interval. From the table we can infer that, for the heavy tailed distribution ($\beta \leq 1.5$) the Hurwicz estimator yields a better coverage level which is shown in bold fonts.

Table 4.9: Simulation results of AR(1) with GED innovation (for $\beta = 1, 1.1, 1.3, 1.5, 1.7, 2$ values) and $\phi = 0.9$.

$\phi = 0.9$		nominal coverage 95%			nominal coverage 90%		
$\sigma = 1$		cvr	len	serr	cvr	len	serr
$\beta = 2$	LS	**0.952**	3.27438	0.01325	**0.900**	2.55926	0.00991
	HUR	0.950	3.29504	0.01334	0.898	2.57248	0.00996
1.7	LS	**0.948**	3.52452	0.01499	**0.920**	2.71718	0.01152
	HUR	0.946	3.55480	0.01534	0.910	2.70497	0.01137
1.5	LS	0.942	3.71255	0.01863	0.902	2.79499	0.01170
	HUR	**0.944**	3.73374	0.01866	**0.906**	2.79621	0.01179
1.3	LS	0.936	3.99018	0.01990	0.890	2.96035	0.01542
	HUR	**0.940**	4.00729	0.02015	**0.892**	2.96968	0.01542
1.1	LS	0.938	4.38040	0.02454	0.878	3.13251	0.01712
	HUR	**0.940**	4.38839	0.02488	**0.884**	3.13031	0.01712
1	LS	0.934	4.61106	0.02774	0.886	3.23719	0.01826
	HUR	**0.940**	4.60540	0.02761	**0.888**	3.23100	0.01860

4.6 Data Analysis

4.6.1 BSE Index

The model is applied to financial data sets. The first set of data consists of daily maximum of BSE Index from $May\ 5^{th}$ 2015 to $July\ 1^{st}$ 2016. There are 290 observations. The plots of the actual time series, logarithmic differences, ACF and PACF of the log returns are given in Figure 4.2. The original series is not stationary and hence by taking the logarithmic difference the series becomes stationary and is confirmed by augmented Dickey Fuller test. The Ljung-Box test lead to the

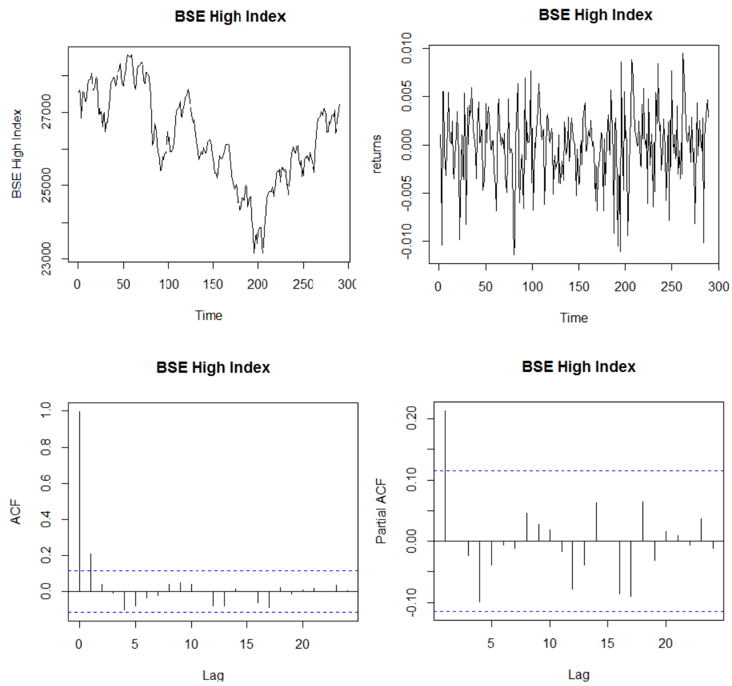

FIGURE 4.2: Time series plot of BSE index and returns(top panel) and ACF and PACF plots of return series(bottom panel)

conclusion that the logarithmic difference series is not independently distributed.

From Figure 4.2 we can infer that the ACF is exponentially decaying and PACF cuts off after lag 1, so the log difference series may follow an AR(1) model. As a starting point we fitted an AR(1) model with Gaussian innovations and found that the estimate of ϕ as 0.2131. However, the Q-Q plot which is shown in Figure 4.3 and the p-value(<0.0001) based on Kolmogorov-Smirnov test lead to the rejection of the null hypothesis that the residuals are normal. Then we fit an AR(1) model with GED innovations to the data and obtained the estimates as $\hat{\phi}$=0.2979, $\hat{\sigma}$=.003862 and $\hat{\beta}$=1.3979 by taking Hurwicz estimator and method of moments respectively. The K-S test statistic is obtained as 0.067381 and the p-value is 0.1463. The ACF plot of residuals and the superimposition of the histogram of the residuals on the pdf of GED given in Figure 4.3 gives that the model is of good fit.

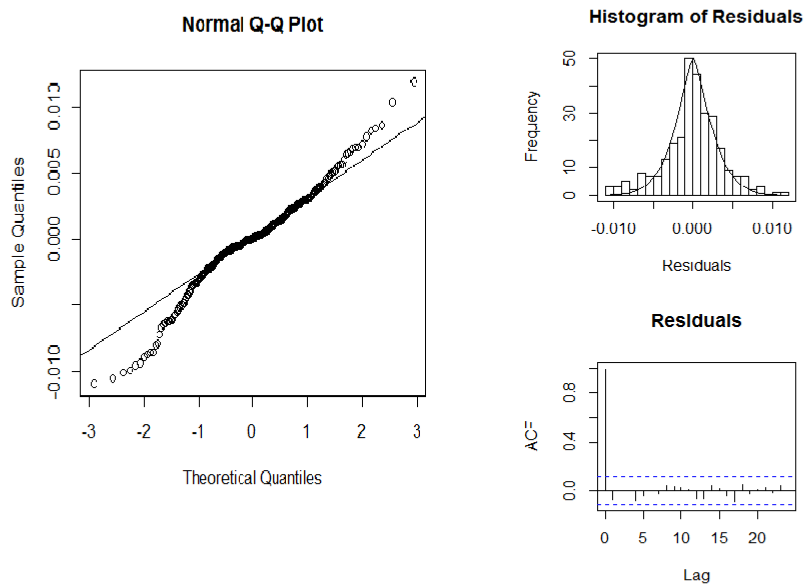

FIGURE 4.3: Q-Q plot (left panel) and histogram of the residuals superimposed with GED density and ACF plot of residuals (right panel)

4.6.2 NIFTY 50 INDEX

Another set of data considered is a time series of daily maximum of NIFTY 50 INDEX $May\ 5^{th}$ 2015 to $July\ 1^{st}$ 2016. There are 290 observations. The plots of the actual time series, logarithmic differences, ACF and PACF of the log returns are given in Figure 4.4. The ACF is exponentially decaying and PACF cuts off after lag 1 so the ACF/PACF of stationary series suggest an AR(1) model for the data.

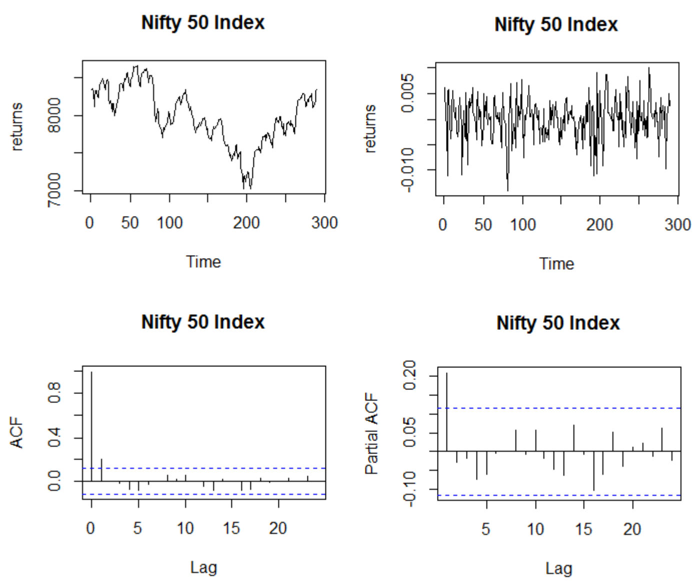

FIGURE 4.4: Time series plot of Nifty index and returns (top panel) and ACF and PACF plots of return series (bottom panel)

We fit an AR(1) model with Gaussian innovations and found that the estimate of ϕ as 0.2107. From the Q-Q plot shown in Figure 4.4 and the p-value(<0.0001) obtained based on K-S test lead to the rejection of the null hypothesis that the residuals are normal.

An AR(1) model with GED innovations is fitted to the data and obtained the estimates as $\hat{\phi}$=0.243, $\hat{\sigma}$=.00464 and $\hat{\beta}$=1.3248 by Hurwicz estimation and method of moments respectively. The K-S test statistic is obtained as 0.06956 and the p-value is 0.1219. So this model is suitable for the data. The ACF plot of residuals and the superimposition of the histogram of the residuals on the pdf of GED are given in Figure 4.5.

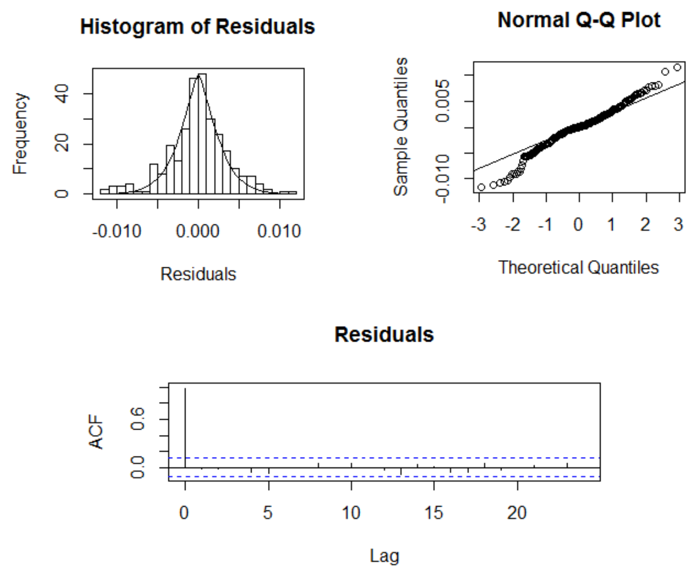

FIGURE 4.5: Histogram of residuals superimposed with GED density and Q-Q plot (top panel) and ACF of the residuals (bottom panel)

Notes

Verification of strong mixing property

Result (Athreya and Pantula (1986)): Let $\{Z_t\}$ be an AR process given by $Z_t = \phi Z_{t-1} + a_t$; $|\phi| < 1$ and $|a_t|$ are iid random variables independent of Z_0. Assume that

1. $E[\{\log|a_1|\}^+]$ is finite and
2. a_1 has a non trivial absolutely continuous component. Then for any initial distribution Λ of Z_0, Z_n is strong mixing.

Theorem 4.1. *Vasudeava and Vasantha Kumari (2013) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed GED-I random variables with parameter ν. Then $\sum_{i=1}^{n} |X_i|^\nu$ is a gamma distributed random variable. (Gamma$(1/(2\lambda^\nu), n/\nu)$)*

Remark: The probability density function of GED-I is given as

$$f_\nu = \frac{\nu \exp\left(-\frac{1}{2}\left|\frac{x}{\nu}\right|^\nu\right)}{\left(\lambda 2^{\frac{1}{\nu}+1}\Gamma\left(\frac{1}{\nu}\right)\right)}, \qquad \nu > 0,\ x \in \mathbb{R},$$

where $\lambda = \left(\frac{2^{-2/\nu}\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}\right)^{\frac{1}{2}}$ and $\Gamma(\cdot)$ denotes the Gamma function.

To verify the above conditions we have to show $E[\{\log|a_1|\}^+]$ is finite, where a_t is defined by (3.2).

Let

$$\log|a_1|^+ = \begin{cases} 0, & \text{if } |a_1| \leq 1, \\ \log|a_1|, & \text{if } |a_1| > 1. \end{cases}$$

Put $y = |a_1|$. Then $E[\log y] = \int \log(y) f(y) dy < \infty$.

Then from Theorem 4.1, if $X \sim GED - I$, then $|X|$ is Gamma$[\frac{1}{2\lambda}, 1]$. Then

$$E[log\, y]^+ = \Gamma(1.414)^{-1} \int_0^\infty log\, y\ exp(-y)\ y^{-1+1.414} dy < \infty$$

Further, a_t is absolutely continuous and hence Z_t is strong mixing.

The results of this Chapter are reported in the article Sri Ranganath (2017a)

Chapter 5

Lindley Stochastic Volatility Model

5.1 Introduction

In this class of models volatility are assumed to be generated by latent stochastic models. One among them is the lognormal stochastic volatility model introduced by Taylor (1986). The presence of unobservable volatility in the model makes the inference problems difficult. Many alternatives for the conditional distribution in the model (2.7) have been suggested by Taylor (1994), Jacquier et al. (2002). Andersson (2001) considered the stochastic volatility model with an inverse Gaussian distributed conditional variance. Later, Abraham et al. (2006) proposed a gamma SV model and investigated its distributional and time series properties. In this chapter, we study the properties of Lindley Stochastic Volatility model.

In Section 5.2, a brief introduction to the Lindley distribution is given. In Section 5.3, the Lindley SV model and its second order properties are described. We discuss the method of moments estimation procedure to estimate the parameters of Lindley SV model in Section 5.4. The asymptotic properties of the estimators are established in Section 5.5. A simulation study is carried out in Section 5.6. As the last Section, the proposed model is illustrated using a financial data.

5.2 Lindley distribution

Lindley distribution was introduced by Lindley (1958, 1970). A random variable is said to follow Lindley distribution, if its probability density function is of the form

$$f(x) = \frac{\theta^2 (x+1) e^{-\theta x}}{\theta + 1} \qquad x > 0,\, \theta > 0 \tag{5.1}$$

and with cumulative distribution function

$$F(x) = 1 - \frac{e^{-(\theta x)}(\theta + \theta x + 1)}{\theta + 1} \qquad x > 0,\, \theta > 0. \tag{5.2}$$

The mean, variance, skewness and kurtosis are given below.

$$\begin{array}{lll}
E(X) & = & \frac{\theta+2}{\theta(\theta+1)}. \\
V(X) & = & \frac{2}{\theta^2} - \frac{1}{(\theta+1)^2}. \\
Skewness & = & \frac{2\left(\theta^3+6\theta^2+6\theta+2\right)}{(\theta^2+4\theta+2)^{3/2}}. \\
K_X & = & \frac{3\left(3\theta^4+24\theta^3+44\theta^2+32\theta+8\right)}{(\theta^2+4\theta+2)^2}. \\
Also, \quad E(X^k) & = & \frac{k!(\theta+k+1)}{\theta^k(\theta+1)}.
\end{array}$$

Lindley distribution can be written as a mixture of an Exponential distribution with shape parameter θ and Gamma distribution, $Gamma(2, \theta)$ with mixing proportion $p = \frac{\theta}{\theta+1}$. Ghitany et al. (2008) has discussed some statistical properties of Lindley distribution.

5.3 Lindley SV model

Let $\{Y_t\}$ be the sequence of returns on certain financial asset at time t, $t = 0, \pm 1, \pm 2, \ldots$. Define the SV model

$$Y_t = \varepsilon_t \sqrt{h_t}, \tag{5.3}$$

$$h_t = \phi h_{t-1} + \eta_t, \qquad t = 1, 2, \ldots. \quad 0 \leq \phi < 1, \tag{5.4}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed standard normal random variables. We assume that the sequence η_1 is independent of h_1 and $\{\varepsilon_t\}$ is independent of h_t for every t. Here, we assume that $\{h_t\}$ is a stationary $AR(1)$ process defined by Bakouch and Popović (2016), with stationary Lindley marginal function given in (5.3). They have established that the distribution of the innovation random variable, η_t is specified as the mixture of the singular and absolute continuous distribution given by

$$f_\eta(x) = \phi\delta(x) + (1 - \phi)g(x), \tag{5.5}$$

where $\delta(x)$ is the Dirac Delta function defined by

$$\delta(x) = \begin{cases} +\infty, \; x = 0, \\ 0, \; x \neq 0. \end{cases}$$

and $g(x)$ is a probability density function having the form

$$g(x) = \frac{\theta^2(1-\phi)^2 + \theta(1-\phi^2) + 2\phi}{(\phi(1-\phi)+1)^2}\theta\exp(-\theta x) + \frac{1-\phi}{(\theta(1-\phi)+1)}\theta^2 x\exp(-\theta x)$$
$$-\frac{\theta+1}{(\theta(1-\phi)+1)^2}\exp(-\frac{\theta+1}{\phi})x \tag{5.6}$$

with $\theta > 0, x \geq 0$. The mean and variance of are

$$E(\eta_t) = \frac{\theta+2}{\theta(\theta+1)}(1-\phi), \qquad V(\eta_t) = \frac{(1-\phi^2)(\theta^2+4\theta+2)}{\theta^2(\theta+1)^2}. \tag{5.7}$$

Then the characteristic function (cf) of Y_t is given by

$$\begin{aligned}\zeta(s) &= E(\exp(isY_t)) \\ &= E(\exp(is\varepsilon_t\sqrt{(h_t)})) \\ &= \int_0^\infty E(\exp(is\sqrt{(x)}\varepsilon_t))f_h(x)dx \qquad [\because h_t = x] \\ &= \frac{\theta}{\theta+1}\frac{2\theta}{2\theta+s^2} + \frac{1}{\theta+1}\left(\frac{2\theta}{2\theta+s^2}\right)^2 \qquad (5.8)\end{aligned}$$

That is,

$$\zeta(s) = \frac{\theta}{\theta+1}L_1 + \frac{1}{\theta+1}L_2 + \frac{1}{\theta+1}L_3 \tag{5.9}$$

Thus $\{Y_t\}$ is a stationary Markov sequence whose marginal distribution has the cf (5.8), which can be expressed as in (5.9), where L_1, L_2, L_3 are independent and identically distributed random variables with Laplace distribution with probability density function

$$f_L(y) = \theta\exp(-2\theta|y|) \qquad -\infty < y < \infty,\ \theta > 0.$$

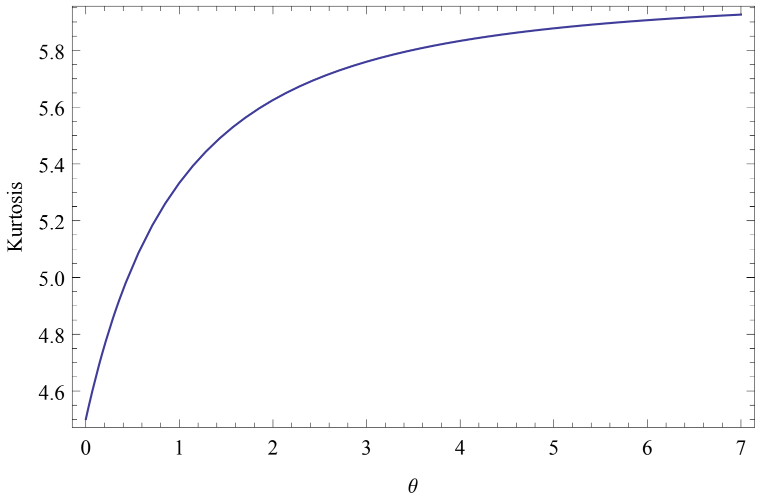

FIGURE 5.1: The plot of kurtosis, K_Y, of Y_t

The odd moments of Y_t are zero and its even moments are given by

$$\begin{aligned} E(Y_t^{2k}) &= E\left((\varepsilon_t\sqrt{h_t})^k\right) \\ &= E(\varepsilon_t^k)E(h_t^{\frac{k}{2}}) \\ &= 1.3.5.\ldots.\left(\frac{k!(\theta+k+1)}{\theta^k(\theta+1)}\right), \quad k=1,2,\ldots. \end{aligned} \tag{5.10}$$

$$E(Y_t^2) = E(\varepsilon_t^2)E(h_t) = \frac{\theta+2}{\theta(\theta+1)} = Var(Y_t) \tag{5.11}$$

$$E(Y_t^4) = E(\varepsilon_t^2)E(h_t^2) = 3\frac{2(\theta+3)}{\theta^2(\theta+1)} \tag{5.12}$$

The kurtosis of Y_t is

$$K_Y = 6\frac{(\theta^2+4\theta+3)}{\theta^2+4\theta+4} < 6.$$

Laplace distribution has a constant kurtosis(=6). But in the present case, $K_Y < 6$, but is a function of θ and hence more flexible. Figure 5.1 shows the flexibility of the kurtosis, K_Y for different θ values. Also, the autocovariance function,

$Cov(h_t, h_{t-k}) = \phi^{|k|}\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}$. So that,

$$
\begin{aligned}
\nu_{Y_t^2}(k) &= E(Y_t^2, Y_{t-k}^2) - E(Y_t^2)E(Y_{t-k}^2) \\
&= E(\varepsilon_t^2 h_t \varepsilon_{t-k}^2 h_{t-k}) - E(\varepsilon_t^2 h_t)E(\varepsilon_{t-k}^2 h_{t-k}) \\
&= E(h_t h_{t-k}) - E(h_t)E(h_{t-k}) \\
&= \phi^{|k|}\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2} - \left(\frac{\theta+2}{\theta(\theta+1)}\right)^2. \qquad (5.13)
\end{aligned}
$$

and the ACF of $\{Y_t^2\}$ is given by

$$\rho_{Y_t^2}(k) = Corr(Y_t^2, Y_{t-k}^2) = \frac{((\theta+4)\theta+2)\phi^k}{(\theta+2)^2} - 1.$$

5.4 Estimation of Lindley-SV model

The likelihood function of a SV model involves multiple integrals. So it is difficult to integrate out the unobservable latent variables in the likelihood function. This possess difficulties in the likelihood based inference of the model. In view of this, the estimation procedure is carried out by Generalized Method of Moments (Hansen (1982)).

Let $(y_1, y_2, \ldots, y_T)$ be a realization of length T from Lindley SV model (5.3) and let $\Theta = (\phi, \theta)'$ be the parameter vector to be estimated. To estimate these parameters, we use the moments $E(Y_t^2) = \frac{\theta+2}{\theta(\theta+1)}$ and $E(Y_t^2 Y_{t-1}^2) = \phi\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}$. Define the function $g(y_t, y_{t-1}, \Theta)$ as

$$g(y_t, y_{t-1}, \Theta) = \begin{pmatrix} y_t^2 - \frac{\theta+2}{\theta(\theta+1)} \\ y_t^2 y_{t-1}^2 - \phi\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2} \end{pmatrix}. \qquad (5.14)$$

Then the moment estimator $\hat{\Theta} = (\hat{\phi}, \hat{\theta})'$ of Θ can be obtained by solving the equation

$$\frac{1}{T}\sum_{t=1}^{T} g(y_t, y_{t-1}, \Theta) = 0.$$

The resulting moment equations for ϕ and θ can be expressed as

$$\bar{Y}^2 = \frac{\theta + 2}{\theta(\theta + 1)}; \quad \hat{\phi} = \bar{Y}_{22}\frac{(\hat{\theta}^2(\hat{\theta} + 1)^2)}{\hat{\theta}^2 + 4\hat{\theta} + 2}, \tag{5.15}$$

where $\bar{Y}^2 = (1/T)\sum_{t=1}^{T} y_t^2$, $\bar{Y}_{22} = (1/T)\sum_{t=1}^{T} y_t^2 y_{t-1}^2$.

5.5 Asymptotic properties of the estimators

The asymptotic properties for generalized method of moment estimators are established using the results of Hansen (1982) as given in Section 3.4.2. According to Hansen (1982) GMM estimators are obtained using a large number of moment equations. Under the assumptions mentioned in Section 3.4.2, Hansen (1982) proved that the generalized moment estimators are consistent and asymptotically normal. The Theorem 3.1 states that $\left\{\sqrt{T}(\hat{\Theta} - \Theta), T \geq 1\right\}$ converges in distribution to a normal random vector with mean 0 and dispersion matrix $[DS^{-1}D']^{-1}$, where D is as given in Section 3.4.2. From Bakouch and Popović (2016), the sequence $\{h_t\}$ is stationary, ergodic and has finite moments. So, these properties follows for the sequence $\{y_t\}$ given in Section 5.3. The partial derivatives are given as

$$\frac{\partial g}{\partial \theta} = \begin{pmatrix} \frac{1}{(\theta+1)^2} - \frac{2}{\theta^2} \\ -\frac{2(\theta(\theta(\theta+6)+6)+2)\phi}{\theta^3(\theta+1)^3} \end{pmatrix}$$

and

$$\frac{\partial g}{\partial \phi} = \begin{pmatrix} 0 \\ \frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2} \end{pmatrix}.$$

Thus, we can confirm the existence of the partial derivatives, $\partial g/\partial\Theta$ and also they are continuous for all Θ. Similarly, $E(\partial g/\partial\Theta)$ and $E(\omega_0\omega_0')$ exists and are finite. Hence, the regularity conditions stated in Section 3.4.2 hold good for Lindley-SV model. To compute the asymptotic standard errors of the estimators, we need the dispersion matrix D. The computations for obtaining the D matrix are discussed as follows.

Let $\quad \Gamma_{(k)} = \begin{pmatrix} \gamma_{11}^{(k)} & \gamma_{12}^{(k)} \\ \gamma_{21}^{(k)} & \gamma_{22}^{(k)} \end{pmatrix}, \quad k = 0, \pm 1, \pm 2, \ldots$

and $\Gamma_{(k)}$=$\Gamma_{(-k)}$, $k = 1, 2, \ldots$. Then the 2×2 matrix S is given by $S = \Gamma_{(0)} + 2\sum_{k=1}^{\infty}\Gamma_{(k)}$.

When $k = 0$, the elements of $\Gamma_{(0)}$= $E\left(\omega_t\omega_t'\right)$ are obtained as

$$\gamma_{11}^{(0)} = \frac{6(\theta+3)}{\theta^2(\theta+1)} - \frac{(\theta+2)^2}{\theta^2(\theta+1)^2};$$

$$\gamma_{12}^{(0)} = \gamma_{21}^{(0)} = \frac{90(\theta+4)}{\theta^3(\theta+1)} - \frac{(\theta+2)(\phi(\theta^2+4\theta+2))}{\theta(\theta+1)(\theta^2)(\theta+1)^2};$$

$$\gamma_{22}^{(0)} = 9\left(\frac{24(\theta+5)\phi^2}{\theta^4(\theta+1)} + \frac{(2(\theta+3))}{\theta^4(\theta+1)}\left(\frac{6(1-\phi)}{\theta^2} + \frac{2(1-\phi)}{\theta^2} + \frac{2(1-\phi)\phi^2}{(\theta+1)^2}\right)\right) - \left(\frac{\left(\theta^2+4\theta+2\right)\phi}{\theta^2(\theta+1)^2}\right)^2.$$

The elements of $\Gamma_{(k)}\, fork = 1, 2, \ldots$ are

$$\gamma_{11}^{(k)} = \phi^{|k|}\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2} - \frac{(\theta+2)^2}{\theta^2(\theta+1)^2};$$

$$\gamma_{12}^{(k)} = \phi^{k+2}h_3 + 2\phi^{k+1}e_1h_2 + \phi^k e_2h_2 + h_1h_{12}\sum_{j=0}^{k-1}\phi^j - c_2^k h_1 - c_1h_{12} + c_1c_2^k;$$

$$\gamma_{21}^{(k)} = \phi\left(\phi^{2(k-1)}h_3 + 2\phi^{k-1}e_1h_2\sum_{j=1}^{k}\phi^{k-j}\right) + e_1\left(\phi^{k-1}e_2 + e_1h_1\sum_{j=1}^{k}\phi^{k-j}\right)$$
$$- c_1h_{12} - c_2h_1 + c_1c_2;$$

$$\begin{aligned}\gamma_{22}^{(k)} &= \phi^{2(k+1)+1}\left(\phi^3h_4 + 3\phi^2e_1h_3 + 3\phi e_2h_2 + e_3h_1\right) + 2\phi^k\sum_{j=1}^{k}\phi^{k-j}e_1\left(\phi^2h_3\right.\\ &\left.+2\phi e_1h_2 + e_2h_1\right) + \phi e_1h_{12}\sum_{j=1}^{k}\phi^{k-j} + e_1\phi^k\left(\phi^2h_3 + 2\phi e_1h_2 + e_2h_1\right)\\ &+ e_1\sum_{j=1}^{k}\phi^{k-j}e_1h_{12} - c_2^kh_{12} - c_2h12 + c_2c_2^k,\end{aligned}$$

where $c_1 = e_1 = h_1 = (\theta+2)/(\theta(\theta+1))$, $e_2 = 2(\theta^2+4\theta+3-\phi)/(\theta^2(\theta+1)^2)$, $h_2 = 6(\theta+3)/(\theta^2(\theta+1))$, $h_3 = 3(\theta+2)/(\theta^3(\theta+1))$ $h_{12} = c_2 = \phi(\theta^2+4\theta+2)/(\theta^2(\theta+1)^2)$, $c_2^k = \phi^{|k|}(\theta^2+4\theta+2)/(\theta^2(\theta+1)^2)$.

The 2×2 matrix D is evaluated using the form $D = E(df(y_t, \Theta)/d\Theta)$ and its elements are :

$$D_{11} = \frac{2}{\theta^2} - \frac{1}{(\theta+1)^2};$$

$$D_{12} = \frac{2\phi(\theta^3 + 6\theta^2 + 6\theta + 2)}{\theta^3(\theta+1)^3};$$

$$D_{21} = 0;$$

$$D_{22} = \frac{\theta^2 + 4\theta + 2}{\theta^2(\theta+1)^2}.$$

Hence the asymptotic dispersion matrix becomes $\frac{1}{T}(\Sigma)$, where $\Sigma = [DS^{-1}D']^{-1}$. The asymptotic standard errors of the estimators are computed by utilising the diagonal elements of Σ.

5.6 Simulation Study

A simulation study is carried out to check the performance of the estimators. We consider the sample sizes 2000 and 3000. The distribution and density function of innovation random variable for a Lindley Markov sequence, $\{h_t\}$, is given in (5.5) and (5.6). First, we generate $\{h_t\}$. Based on this, we simulate the sequence y_t using (5.3). Then we obtain the estimates by solving the equations in (5.15).

The experiment was carried out 1000 times for specified values of the parameters. The average estimates and the corresponding mean square error values based on the simulation are tabulated in Tables 5.1 and 5.2.

Table 5.1: The average estimates and the corresponding mean square error of moment estimates based on sample of size, n=2000 for different θ and ϕ values.

	True Values		Average estimate values			
n	θ	ϕ	$\hat{\theta}$	std dev	$\hat{\phi}$	std dev
	0.60	0.10	0.56196	0.02928	0.06212	0.15700
2000	0.50	0.30	0.53980	0.02922	0.24063	0.19880
	0.40	0.50	0.45211	0.03045	0.46554	0.25560
	0.30	0.75	0.35474	0.03369	0.79559	0.34780
	0.50	0.60	0.46179	0.03276	0.55609	0.28289

Table 5.2: The average estimates and the corresponding mean square error of moment estimates based on sample of sizes, n=3000 for different θ and ϕ values.

	True Values		Average estimate values			
n	θ	ϕ	$\hat{\theta}$	std dev	$\hat{\phi}$	std dev
	0.60	0.10	0.58050	0.01982	0.07426	0.11393
3000	0.50	0.30	0.51864	0.02082	0.26982	0.16568
	0.40	0.50	0.43616	0.02102	0.48105	0.18839
	0.30	0.75	0.33427	0.02304	0.77842	0.27844
	0.50	0.60	0.48061	0.02387	0.57134	0.19232

5.7 Data Analysis

We analyse the daily stock price index returns using the Lindley SV model. The closing index data of Nikkei 225 for the period 1^{st} January 2012 to 31^{st} December 2015 of Tokyo Stock Exchange (TSE) is considered. The time series plot and the plot of the log-return series are given in Figure 5.2 and Figure 5.3 respectively. The summary statistics of the return series are given in Table 5.3, where $Q(20)$ and $Q^2(20)$ are the Ljung-Box statistic for the return and squared return series with lag 20, respectively. The corresponding χ^2 table value at 5% significance level is 10.117, which suggest that the return series is uncorrelated and the squared return series shows significant correlation. The acf of the returns and the squared

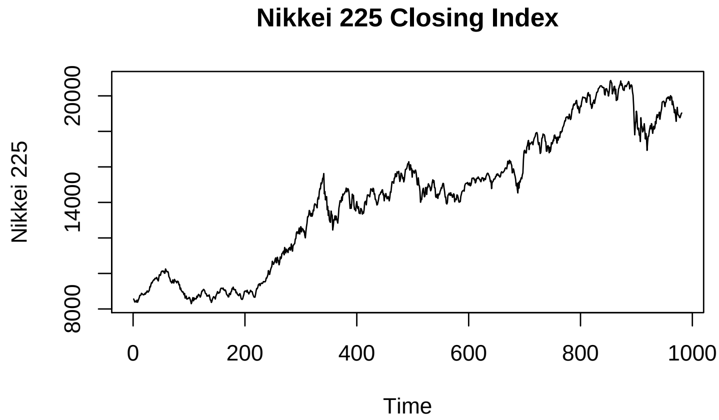

FIGURE 5.2: The time series plot of Nikkei 225 closing index

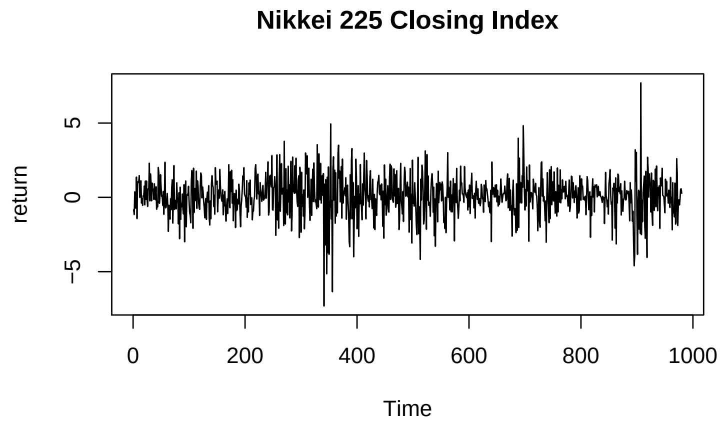

FIGURE 5.3: The time series plot of Nikkei 225 return series

TABLE 5.3: Summary statistics of Nikkei 225 return series

Statistics	Nikkei Closing Index
Sample size	980
Std.Dev	1.35447
Minimum	-7.3159
Maximum	7.70886
$Q(20)$	3.7131
$Q^2(20)$	11.5735

returns are plotted in Figure 5.4. It can be observed from Figure 5.4 that the serial

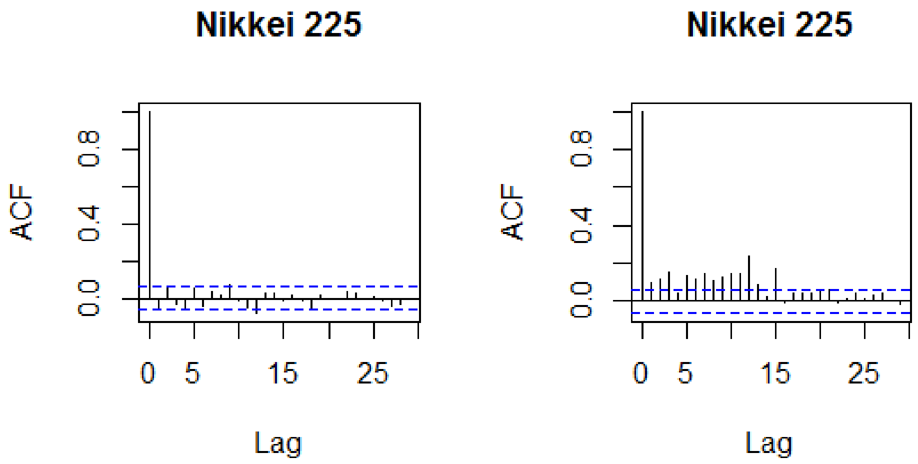

FIGURE 5.4: ACF of returns(left panel) and the squared returns(right panel)

correlation in the return series are insignificant where as the ACF of the squared returns is significant. The parameters are estimated and obtained as $\hat{\theta} = 0.8386$ and $\hat{\phi} = 0.6230$. To validate the model, we need to perform the model diagnostic check based on the residuals. That is, we need to check whether the assumptions on the model (5.3) are satisfied with respect to the data we have analysed. Here, the model is defined in terms of the unobservable volatilities, h_t. One can estimate these volatilities by utilising the Kalman-Filter method. It is a recursive algorithm that computes estimates for the unobserved components at time t, based on the information available at the same time. To employ this method, the model should be represented in state space form with the random variables involved having a normal distribution. This is because, the Kalman method assumes that the distribution of underlying rvs is normal. Jacquier et al. (2002) and Tsay (2005) discuss about this method in detail. Here we approximate the distribution of η_t by a normal distribution and then adopt Kalman filter method for estimating the volatilities.

The state space representation of the SV model given in (5.3) can be written as

$$log\, Y_t^2 = -1.27 + log\, h_t + \vartheta_t, \qquad E(\vartheta_t) = 0, V(\vartheta_t) = \frac{\pi^2}{2}; \tag{5.16}$$

and $h_t = \phi h_{t-1} + \eta_t$, where η_t is assumed to be normally distributed with mean $E(\eta_t) = (\theta+2)(1-\phi)/\theta(\theta+1)$ and variance $V(\eta_t) = (1-\phi^2)(\theta^2+4\theta+2)/(\theta^2(\theta+1)^2$ which are given in (5.7). If the distribution of ϑ_t is approximated by a normal distribution then the preceding system (5.16) becomes a standard dynamic linear model, to which Kalman filter can be applied. Let $\overline{h}_{t|t-1}$ be the prediction of h_t based on the information available at time $t-1$ and $\Omega_{t|t-1}$ be the variance of the predictor. Here we are making a presumption that the update that uses the information at time t as $\overline{h}_{t|t}$ and the variance of the update as $\Omega_{t|t}$. The equations that recursively compute the predictions and updating are given by

$$\begin{aligned}
\overline{h}_{t|t-1} =& \phi\overline{h}_{t-1|t-1} + (\theta+2)(1-\phi)/\theta(\theta+1) \\
\Omega_{t|t-1} =& \phi^2\Omega_{t-1|t-1} + (1-\phi^2)(\theta^2+4\theta+2)/(\theta^2(\theta+1)^2
\end{aligned}$$

and

$$\begin{aligned}
\overline{h}_{t|t} =& \overline{h}_{t|t-1} + \frac{\Omega_{t|t-1}}{f_t}[log\, Y_t^2 + 1.27 - \overline{h}_{t|t-1}] \\
\Omega_{t|t} =& \Omega_{t|t-1}(1 - \frac{\Omega_{t|t-1}}{f_t}),
\end{aligned}$$

where $f_t = \Omega_{t|t-1} + \frac{\pi^2}{2}$. Then we can compute the residuals by the equation $\hat{\varepsilon} = y_t/\sqrt{\overline{h}_t}$ and use this sequence for the residual analysis. The system is initialised at the unconditional values, $\Omega_0 = (\theta^2+4\theta+2)/(\theta^2(\theta+1)^2$ and $h_0 = (\theta+2)/\theta(\theta+1)$. The parameters θ and ϕ in the above system are replaced by their estimates $\hat{\theta}$ and $\hat{\phi}$ respectively. The ACF plot of the residuals is given in Figure 5.5. Further the

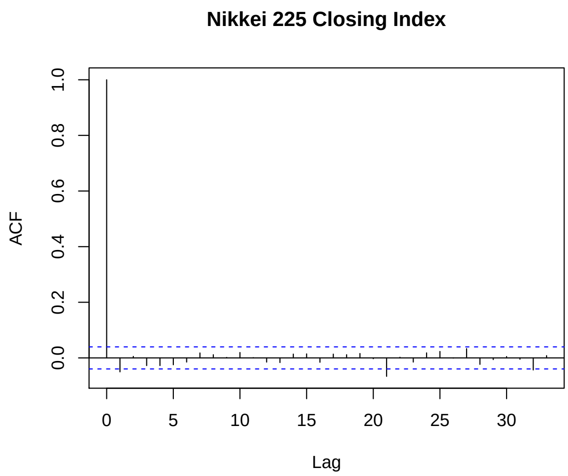

FIGURE 5.5: ACF of residuals

significance of ACF in the residuals is computed using the Ljung-Box statistic for the series $\{\hat{\varepsilon}\}$ and $\{\hat{\varepsilon^2}\}$ and the values obtained are 3.687 and 6.135 respectively. All these values are less than the 5% Chi-square critical value (10.117) at degrees of freedom 20. Hence we conclude that there is no significant serial dependence among the residuals and squared residuals. In Figure 5.6, we superimpose the histogram

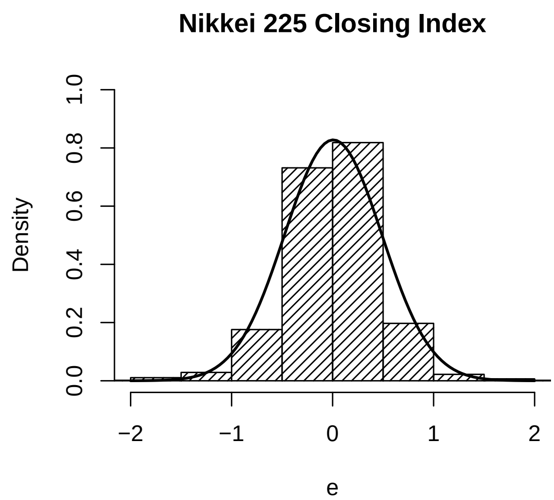

FIGURE 5.6: Histogram of residuals with superimposed standard normal density

of the residuals with standard normal density to check whether the series follows standard normal distribution. From the figure, the standard normal density is a good fit for the residuals.

Hence our proposed model is capable of capturing the stylized facts of the above data set.

The results of this Chapter are reported in Sri Ranganath and Balakrishna (2018).

Chapter 6

Lindley Autoregressive Conditional Duration model

6.1 Introduction

The modelling of time interval between trades have gathered momentum in recent years. The study of this irregularly space trade durations contains useful information about intraday market activities. Engle (1982) proposed the autoregressive conditional duration model to analyse the dynamic behaviour of financial duration data. They assume the Exponential and Weibull distribution as the conditional distribution of an interval given the past information. Lunde (1999) considered the generalized Gamma distribution and Grammig and Maurer (2000) assumed the Burr distribution as ACD specifications. To ensure the non-negativity durations without any parameter restrictions Bauwens and Giot (2000) proposed a logarithmic ACD model. Hautsch (2001) suggested an additive ACD model based on power transformation which can be written in terms of Box-Cox transformation. Pacurar (2008) gives a survey of the theoretical and empirical literature on the ACD models. The Exponential, Weibull and Gamma distributions are usually considered for modelling the conditional durations. In this Chapter, we propose an ACD model with Lindley distribution as the conditional distribution. Ghitany et al. (2008) has discussed some statistical properties of Lindley distribution. They showed that, due to the flexible nature of Lindley distribution it performs better than exponential distribution in some applications. So, Lindley distribution can be considered as an alternative to exponential distribution in modelling duration models. Bakouch and Popović (2016) proposed an autoregressive process of first order based on Lindley marginal distribution. Raqab et al. (2017) consider the problem of the model selection or discrimination among three different positively skewed lifetime distributions, the Lindley, Weibull, and Gamma distributions. They use the likelihood ratio test and minimum Kolmogorov distance tools to analyse the distributions and concluded that the Lindley distribution is closer to Weibull distribution in the sense of likelihood ratio and Kolmogorov criteria.

In Section 6.2, the Lindley-ACD(1,1) model is defined. The second order properties are described in Section 6.3. We discuss the estimation procedure by conditional maximum likelihood method in Section 6.4. A simulation study is carried out in Section 6.5. In Section 6.6, we present the data analysis using the proposed duration model.

6.2 Lindley ACD model

Let X follow Lindley distribution with mean $\frac{\theta+2}{\theta(\theta+1)}$. Then $\varepsilon = X/\frac{\theta+2}{\delta(\theta+1)}$ follows a unit mean Lindley distribution with pdf (6.1.)

$$f_{\varepsilon_i}(\varepsilon) = \frac{\theta^2}{\theta+1}\left(1+\frac{\theta+2}{\theta(\theta+1)}\varepsilon\right)\exp^{-(\frac{\theta+2}{\theta+1})\varepsilon} \qquad \varepsilon > 0. \tag{6.1}$$

Now we define the Lindley ACD(1,1) model as

$$\begin{aligned} X_i &= \psi_i \varepsilon_i \\ \psi_i &= \omega + \alpha X_{i-1} + \beta \psi_{i-1}\,, \qquad i = 1, 2, \ldots, n \end{aligned} \tag{6.2}$$

The conditional pdf of X_i given ψ_i is

$$f_{(X_i|\psi_i)}(x_i) = \frac{1}{\psi_i} f_{\varepsilon_i}\left(\frac{x_i}{\psi_i}\right) = \frac{\theta(\theta+2)\left(1 + \frac{(\theta+2)x_i}{\theta(\theta+1)\psi_i}\right)\exp\left(-\frac{(\theta+2)\frac{x_i}{\psi_i}}{\theta+1}\right)}{(\theta+1)^2}\frac{1}{\psi_i}, \quad \theta > 0. \tag{6.3}$$

6.3 Properties of Lindley ACD model

Consider the model given in (6.2). The conditional mean of the ACD model is given by $E[X_i|F_{t_i-1}] = \psi_i$. Taking expectation the unconditional mean is,

$$\begin{aligned} E(X_i) &= E[E(\psi_i \varepsilon_i | F_{i-1})] = E(\psi_i), \\ E(\psi_i) &= \omega + \alpha E(X_{i-1}) + \beta E(\psi_{i-1}). \end{aligned} \tag{6.4}$$

Under weak stationarity, $E(\psi_i) = E(\psi_{i-1})$ so that, (6.4) gives

$$\mu_x \equiv E(X_i) = E(\psi_i) = \frac{\omega}{1 - \alpha - \beta}. \tag{6.5}$$

Consequently, for weak stationarity of $\{X_i\}$ we need the condition $0 \leq \alpha + \beta < 1$. Also, $E({X_i}^2) = E[E({\psi_i}^2 {\varepsilon_i}^2 | F_{i-1})]$. From (6.4), we have $E(\varepsilon_i^2) = \frac{2(1+\theta)(3+\theta)}{(2+\theta^2)}$. So, taking the square of ψ_i in (6.4) and the expectation and using weak stationarity of

ψ_i and X_i, we have,

$$
\begin{aligned}
E(\psi_i^2) &= E(\omega^2 + \alpha^2 X_{i-1}^2 + \beta^2 \psi_{i-1}^2 + 2\alpha\omega X_{i-1} + 2\beta\omega\psi_{i-1} + 2\alpha\beta X_{i-1}\psi_{i-1}) \\
&= \omega^2 + \alpha^2 E(X_{i-1}^2) + \beta^2 E(\psi_{i-1}^2) + 2\omega\alpha\mu_x + 2\omega\beta\mu_x + 2\alpha\beta\mu_x^2 \\
&= \omega^2 + \alpha^2 E(\psi_i^2)\frac{2(1+\theta)(3+\theta)}{(2+\theta)^2} + \beta^2 E(\psi_i^2) + 2\omega\alpha\mu_x + 2\omega\beta\mu_x + 2\alpha\beta\mu_x^2 \\
&= \frac{\omega^2 + 2\omega\alpha\mu_x + 2\omega\beta\mu_x + 2\alpha\beta\mu_x^2}{1 - \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2 + \beta^2 - 2\alpha\beta} \\
&= \frac{\mu_x^2(1-(\alpha+\beta)^2)}{1 - \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2 + \beta^2 - 2\alpha\beta}
\end{aligned}
$$

so that

$$
E({X_i}^2) = \frac{\mu_x^2(1-(\alpha+\beta)^2)}{1 - \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2 - \beta^2 - 2\alpha\beta}\frac{2(1+\theta)(3+\theta)}{(2+\theta^2)}. \tag{6.6}
$$

Using $Var(X_i) = E(X_i^2) - (E(X_i))^2$ and (6.6) , we have the unconditional variance as

$$
\begin{aligned}
V(X_i) &= \frac{\mu_x^2(1-(\alpha+\beta)^2)}{1 - \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2 + \beta^2 - 2\alpha\beta}\frac{2(1+\theta)(3+\theta)}{(2+\theta^2)} - \mu_x^2 \\
&= \frac{(\frac{2(1+\theta)(3+\theta)}{(2+\theta)^2} - 1)\mu_x^2(1-\beta^2-2\alpha\beta)}{1 - \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2 + \beta^2 - 2\alpha\beta} \\
&= \frac{(\delta-1)\mu_x^2(1-\beta^2-2\alpha\beta)}{1-\delta\alpha^2+\beta^2-2\alpha\beta},
\end{aligned} \tag{6.7}
$$

where $\delta = \frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}$.

Autocorrelation function

The k^{th} order auto-covariance function of $\{X_i\}$ is defined as

$$\begin{aligned}\nu_X(k) &= Cov(X_i, X_{i-k}) = Cov(\psi_i, X_{i-k})\\ &= Cov(\omega + \alpha X_{i-1} + \beta\psi_{i-1}, X_{i-k})\\ &= \alpha Cov(X_{i-1}, X_{i-k}) + \beta Cov(\psi_{i-1}, X_{i-k})\\ &= (\alpha+\beta)\nu_{k-1}. \end{aligned} \tag{6.8}$$

The first order autocovariance function of X_i is

$$\begin{aligned}\nu_1 &= Cov(X_i, X_{i-1}) = Cov(\psi_i, X_{i-1})\\ &= Cov(\omega + \alpha X_{i-1} + \beta\psi_{i-1}, X_{i-1})\\ &= \alpha\nu_0 + \beta Var(\psi_{i-1})\\ &= \alpha\frac{(\delta-1)\mu_x^2(1-\beta^2-2\alpha\beta)}{1-\delta\alpha^2+\beta^2-2\alpha\beta} + \beta Var(\psi_{i-1}),\end{aligned}$$

$$\text{where} \qquad \begin{aligned}Var(\psi_i) &= E(\psi_i^2) - E(\psi_i)^2\\ &= \frac{\mu_x^2(1-(\alpha+\beta)^2)}{1-\frac{2(1+\theta)(3+\theta)}{(2+\theta)^2}\alpha^2+\beta^2-2\alpha\beta} - \mu_x^2\\ &= \frac{(\delta-1)\alpha^2\mu_x^2}{1-\delta\alpha^2+\beta^2-2\alpha\beta}.\end{aligned}$$

ACF of lag k for $k > 1$ is

$$\rho_X(k) = (\alpha+\beta)\rho_{k-1}, \tag{6.9}$$

$$\text{with} \quad \rho_1 = \frac{\alpha(1-\beta^2)-\alpha\beta}{1-\beta^2-2\alpha\beta}$$

Hazard function of Lindley distribution

The transaction duration in finance is inversely related to trading intensity, which in turn depends on the arrival of new information. So the choice of the distribution of the error term in (6.2) impacts the conditional intensity or hazard function of the ACD model. As the hazard function of durations is not constant over time, the exponential specification is quite restrictive. The hazard function of Lindley distribution shows flexibility over the exponential distribution. For a random variable X, the survival function is defined as

$$S(x) \equiv P(X > x) = 1 - P(X \leq x) \; x > 0, \tag{6.10}$$

which gives the probability that a subject, which follows the distribution of X, survives upto the time x. The hazard function (or the intensity function) of X is then defined by

$$h(x) = \frac{f(x)}{S(x)}, \tag{6.11}$$

where $f(.)$ and $S(.)$ are the pdf and survival function of X, respectively. The hazard function of Lindley distribution is $h(x) = (\theta^2(1+x))/(\theta+1+\theta x) \quad x \geq 0$.

Hazard function of Lindley-ACD(1,1) model

The conditional hazard function of Lindley-ACD(1,1) is given by

$$h(x_i) = \frac{\theta(\theta+2)\left(\frac{(\theta+2)x_i}{\theta(\theta+1)\psi_i} + 1\right)}{(\theta+1)^2\psi_i + (\theta+2)x_i} \tag{6.12}$$

For an exponential distribution, the hazard function is constant. The hazard function curves for different θ values are shown in Figure 6.1. Here $h(x)$ is an increasing function in x and θ. This shows the flexibility of Lindley distribution

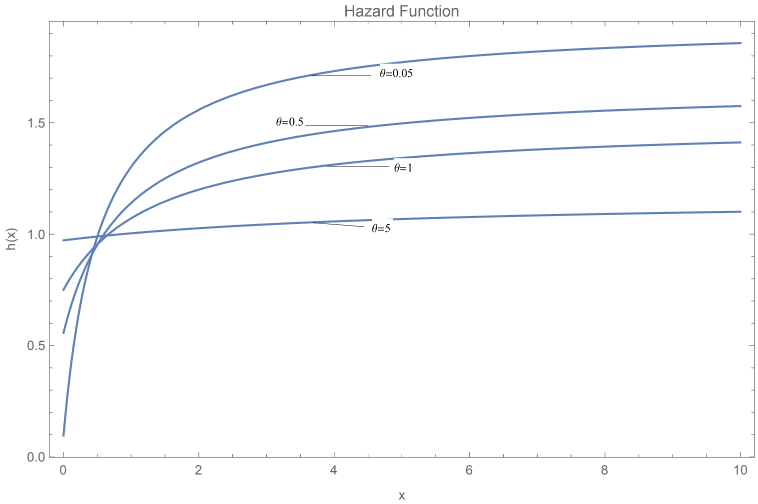

FIGURE 6.1: Plot of Hazard rate

over the exponential distribution.

6.4 Estimation of Lindley-ACD(1,1) model

Here we discuss the problem of estimation of Lindley-ACD(1,1) model. Lu et al. (2016) proposed the moment closed form estimator (MCFE) for the autoregressive conditional duration model. The ACD model is represented as Autoregressive Moving Average using identical transformation. Then using the sample mean and sample autocorrelation function the estimators of the parameters ω, α and β can be obtained. Once these estimators are known, the estimate for θ is computed.

Maximum likelihood method

The parameter vector of the Lindley-ACD(1,1) model is $\Theta = (\theta, \omega, \alpha, \beta)'$. From (6.3), we have the conditional pdf of X_i given ψ_i. The likelihood function is defined

as

$$L(X|\Theta) = f(X_1|\Theta)\prod_{i=2}^{n} f_{X_i|\psi_i}(X_i|F_{i-1};\Theta), \tag{6.13}$$

where $f(X_1|\Theta)$ is the density function of the initial random variable where $F_i = \sigma(X_1, X_2, \ldots, X_i)$. We consider the conditional log-likelihood function by ignoring $f(X_1;\Theta)$ as this is of negligible effect on the overall likelihood function as the sample size becomes larger. The conditional log-likelihood is given by

$$\begin{aligned}
\log L &= \sum_{i=2}^{n} log\left(f_{X_i|\psi_i}(X_i|F_{i-1};\Theta)\right)\\
&= \sum_{i=2}^{n} log\left[\frac{\theta(\theta+2)}{(\theta+1)^2}\left(\frac{\theta(\theta+1)\psi_i+(\theta+2)X_i}{\theta(\theta+1)\psi_i}\right)\exp\left(-\frac{\theta+2}{\theta+1}\frac{X_i}{\psi_i}\right)\frac{1}{\psi_i}\right]\\
&= \text{n } \log\theta(\theta+2) - 2\text{n } \log(\theta+1) + \sum_{i=2}^{n} log\left(\theta(\theta+1)(\omega+\alpha X_{i-1}+\beta\psi_{i-1})+(\theta+2)X_i\right)\\
&\quad \sum_{i=2}^{n} log\left(\theta(\theta+1)(\omega+\alpha X_{i-1}+\beta\psi_{i-1})\right) - \frac{\theta+2}{\theta+1}\sum_{i=2}^{n}\frac{X_i}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})}\\
&\quad \sum_{i=2}^{n} log(\omega+\alpha X_{i-1}+\beta\psi_{i-1}).
\end{aligned} \tag{6.14}$$

The ML estimator of $\hat{\Theta} = (\hat{\theta}, \hat{\omega}, \hat{\alpha}, \hat{\beta})'$ of $\Theta = (\theta, \omega, \alpha, \beta)'$ can be obtained by solving the following likelihood equations.

$$\begin{aligned}
\frac{\partial \log L}{\partial\theta} = 0 = &\frac{(\theta+2)}{(\theta+1)^2}\sum_{i=2}^{n}\frac{X_i}{\beta\psi_{i-1}+\alpha X_{i-1}+\omega} - \frac{2n}{\theta+1} - \frac{(2\theta+1)n}{\theta(\theta+1)}\\
&+ \sum_{i=2}^{n}\frac{\theta\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+(\theta+1)\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+X_{i-1}}{\theta(\theta+1)\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+(\theta+2)X_{i-1}}\\
&- \frac{1}{\theta+1}\sum_{i=2}^{n}\frac{X_i}{\beta\psi_{i-1}+\alpha X_{i-1}+\omega}.
\end{aligned} \tag{6.15}$$

$$\frac{\partial}{\partial \omega}\log L = 0 = \sum_{i=2}^{n}\frac{\theta(\theta+1)}{\theta(\theta+1)\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+(\theta+2)X_{i-1}} \\ -\frac{(\theta+2)}{\theta+1}\sum_{i=2}^{n}-\frac{X_i}{\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)^2}-2\sum_{i=2}^{n}\frac{1}{\beta\psi_{i-1}+\alpha X_{i-1}+\omega}. \tag{6.16}$$

$$\frac{\partial}{\partial \alpha}\log L = 0 = \sum_{i=2}^{n}\frac{\theta(\theta+1)X_{i-1}}{\theta(\theta+1)\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+(\theta+2)X_{i-1}}- \\ \frac{(\theta+2)}{\theta+1}\sum_{i=2}^{n}-\frac{X_{i-1}X_i}{\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)^2}-2\sum_{i=2}^{n}\frac{X_{i-1}}{\beta\psi_{i-1}+\alpha X_{i-1}+\omega}. \tag{6.17}$$

$$\frac{\partial}{\partial \beta}\log L = 0 = \sum_{i=2}^{n}\frac{\theta(\theta+1)\psi_{i-1}}{\theta(\theta+1)\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)+(\theta+2)X_{i-1}} \\ -\frac{(\theta+2)}{\theta+1}\sum_{i=2}^{n}-\frac{X_i\psi_{i-1}}{\left(\beta\psi_{i-1}+\alpha X_{i-1}+\omega\right)^2}-2\sum_{i=2}^{n}\frac{\psi_{i-1}}{\beta\psi_{i-1}+\alpha X_{i-1}+\omega}. \tag{6.18}$$

As the likelihood equations have no explicit solutions, numerical methods are used to obtain the estimates. We use the optim() function in R. This is a General-purpose optimization based on NelderMead, quasi-Newton and conjugate-gradient algorithms. We use the moment closed form estimator for the ACD model (Lu et al. (2016)) as the initial value.

6.5 Simulation Study

In this Section, a simulation study is carried out to assess the performance of the estimators. We consider different values for the parameter vector Θ, with sample size 1500, 2500, 4000 and 5000 and took 500 replications. As θ value increases the estimated value approaches closer to true value, for large sample size. In Table 6.2 we take $\theta = 2.5$ for different sample sizes. The true values, estimated values and the corresponding mean square error are given in the Table 6.1 and Table 6.2 with mean square error given in parenthesis. In Lindley distribution the estimator of $\hat{\theta}$ is positively biased. The simulation results show a similar behaviour to $\hat{\theta}$ for Lindley-ACD(1,1) model.

6.6 Data Analysis

In this Section, the intraday transactions of exchange rates data of Canadian dollar vs Swiss Franc downloaded from the Website of Dukascopy, Swiss Forex bank and Marketplace is illustrated. We consider the traded data on 1^{st} November, 2017. The transactions occurred during the normal trading hours from $9:30AM$ to $4:00PM$ Eastern time are taken. It consists of 65378 transactions and 65377 duration data. The downloaded tick data consists of Local time with date,Ask,Bid, Ask Volume and Bid Volume. We took the trade entered time (HH:MM:SS) and compute the durations between each consecutive transactions in seconds. After removing the zero durations we have 9547 intraday durations and the plot is given in Figure 6.2. As the data exhibit diurnal pattern, we consider the adjusted time

TABLE 6.1: The average ML estimators and the corresponding mean square error for Lindley-ACD(1,1) model

n	True Values $(\theta, \omega, \alpha, \beta)$	$\hat{\theta}$ (mse)	$\hat{\omega}$ (mse)	$\hat{\alpha}$ (mse)	$\hat{\beta}$ (mse)
1500	(0.50,1.50, 0.20, 0.60)	0.51885	1.67318	0.20305	0.5725
		(0.00424)	(0.01723)	(0.0012)	(0.00288)
	(1, 1, 0.30, 0.5)	1.01211	1.05086	0.30351	0.48687
		(0.00938)	(0.00786)	(0.00144)	(0.00231)
	(1.50, 1.50, 0.60, 0.20)	1.58073	1.52419	0.59654	0.19661
		(0.02069)	(0.00683)	(0.00171)	(0.00162)
	(2, 1, 0.70, 0.10)	2.22813	1.01408	0.69883	0.09557
		(0.07296)	(0.00427)	(0.00181)	(0.00143)
2500	(0.50,1.50, 0.20, 0.60)	0.51197	1.60878	0.20466	0.58093
		(0.00347)	(0.01232)	(0.00094)	(0.00215)
	(1, 1, 0.30, 0.5)	1.02328	1.02392	0.29992	0.49494
		(0.00757)	(0.00648)	(0.00104)	(0.00185)
	(1.50,1.50, 0.60, 0.20)	1.52906	1.52406	0.59706	0.1968
		(0.01363)	(0.00571)	(0.00137)	(0.00132)
	(2, 1, 0.70, 0.10)	2.10475	1.00993	0.6983	0.09743
		(0.02694)	(0.00312)	(0.00149)	(0.00111)
4000	(0.50,1.50, 0.20, 0.60)	0.50662	1.56516	0.20267	0.58856
		(0.00251)	(0.0096)	(0.00075)	(0.00166)
	(1, 1, 0.30, 0.5)	1.01429	1.0121	0.30084	0.49728
		(0.0057)	(0.00478)	(0.00087)	(0.00141)
	(1.50,1.50, 0.60, 0.20)	1.51474	1.50228	0.60096	0.19957
		(0.01084)	(0.00424)	(0.00097)	(0.00094)
	(2, 1, 0.70, 0.10)	2.07841	1.00161	0.69992	0.09959
		(0.01874)	(0.00241)	(0.00114)	(0.00082)
5000	(0.50,1.50, 0.20, 0.60)	0.50544	1.55843	0.20319	0.58922
		(0.00228)	(0.0081)	(0.00066)	(0.00145)
	(1, 1, 0.30, 0.5)	1.00619	1.01446	0.30131	0.49613
		(0.00509)	(0.00409)	(0.00075)	(0.00123)
	(1.50,1.50, 0.60, 0.20)	1.52172	1.50128	0.60008	0.19919
		(0.00832)	(0.00389)	(0.00093)	(0.00092)
	(2, 1, 0.70, 0.10)	2.00911	1.00276	0.70014	0.10016
		(0.015)	(0.00225)	(0.00096)	(0.00073)

duration $X_i^* = X_i/f(t_i)$, where X_i is the difference of the observed duration i^{th} and $i-1^{th}$ transactions. We follow the method described in (Tsay, 2014),(pp 298-300) to remove this deterministic component. Let $f(t_i)$ be a deterministic function consisting of the cyclical component of X_i. $f(t_i)$ depends on the underlying asset and the systematic behaviour of the market. We assume

$$f(t_i) = exp[d(t_i)], \qquad d(t_i) = \beta_0 + \beta_1 f_1(t_i) + \beta_2 f_2(t_i), \tag{6.19}$$

TABLE 6.2: The average ML estimators and the corresponding mean square error for Lindley-ACD(1,1) model for $(\theta, \omega, \alpha, \beta)$=(2.5,1,0.30,0.50)

Sample size	True Values $(\theta, \omega, \alpha, \beta)$	$\hat{\theta}$ (mse)	$\hat{\omega}$ (mse)	$\hat{\alpha}$ (mse)	$\hat{\beta}$ (mse)
2500	(2.5,1,0.30,0.50)	3.73919	1.02765	0.30091	0.493564
		0.65433	0.00422	0.00081	0.00130
4000		2.66466	1.01882	0.30014	0.49592
		0.02557	0.00340	0.00062	0.00102
5000		2.64926	1.01558	0.30063	0.49587
		0.02503	0.00303	0.00056	0.00093
10000		2.55664	1.00371	0.30030	0.49889
		0.01275	0.00209	0.00039	0.00064

where

$$f_1(t_i) = \frac{t_i - 43200}{23400} \text{and} \quad f_2(t_i) = f_1^2(t_i),$$

where 43200 denotes the 12 : 00 noon and 23400 is number of trading hours measured in seconds. The coefficients β_j in (6.19) are obtained by the least squares method of the linear regression

$$log(X_i) = \beta_0 + \beta_1 f_1(t_1) + \beta_2 f_2(t_2) + \varepsilon_i.$$

The residual is then given by

$$\hat{\varepsilon}_i = log(X_i) - \hat{\beta}_0 - \hat{\beta}_1 f_1(t_1) - \hat{\beta}_2 - \hat{\beta}_2 f_2(t_2).$$

Then $f(t_i) = exp\{\hat{e}_i\}$. Using $f(t_i)$, we obtain the adjusted duration X_i^*. Figure 6.3 shows the time plot of adjusted durations. From the plot, the diurnal pattern is largely removed.

The summary of the duration data and the adjusted duration data are given in Table 6.3

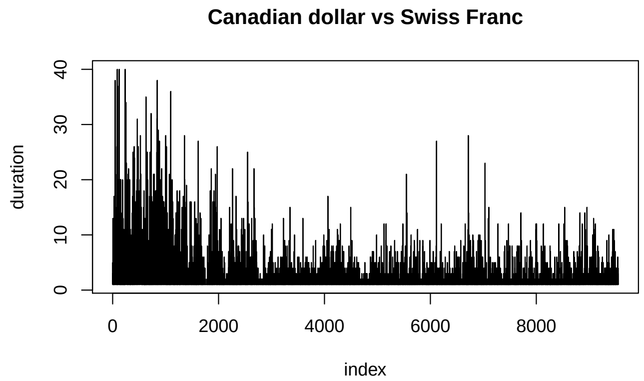

FIGURE 6.2: Time plot of nonzero durations of Canadian dollar vs Swiss Franc .

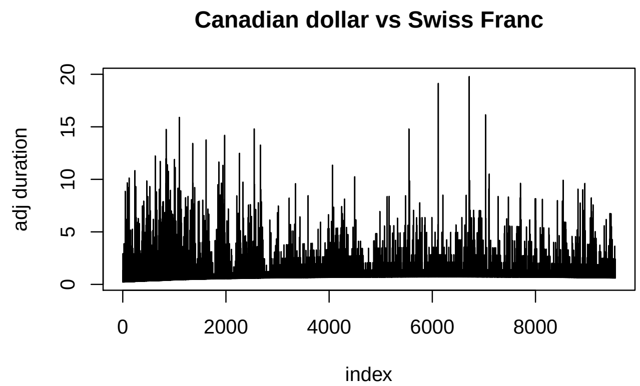

FIGURE 6.3: Time plot of adjusted nonzero durations of Canadian dollar vs Swiss Franc .

TABLE 6.3: Descriptive statistics for nonzero durations and adjusted duration data of Canadian dollar vs Swiss Franc

Statistic	duration data	adjusted duration data
Sample size	9547	9547
Minimum	1	0.2255
Maximum	40	19.7920
Mean	2.451	1.3591
Median	1.023	0.7038
Q(10)	19	0.7038

The estimates of the parameters are carried out using the conditional ML method. Then the fitted model is

$$X_i = \psi_i \varepsilon_i, \qquad \psi_i = 0.02892 + 0.04422 X_{i-1} + 0.93463 \psi_{i-1}$$

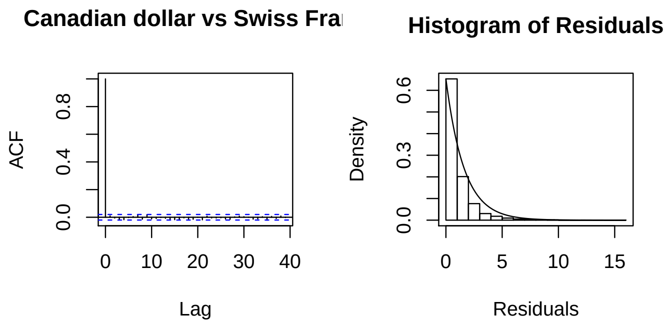

FIGURE 6.4: ACF plot of residual series of Lindley-ACD(1,1) model and histogram of the residuals superimposed by Lindley density

where ε_i follows a standardized Lindley distribution with parameter $\theta = 0.005$. The residual series is computed by $\hat{\varepsilon}_i = X_i/\hat{\psi}_i$. Here $\hat{\varepsilon}_i$ should be uncorrelated random variables. The ACF plot of innovations and the histogram of the innovations superimposed by the unit mean Lindley distribution is given in Figure 6.4. Also the Ljung-Box statistics value with corresponding p-value is given in Table 6.4.

TABLE 6.4: Lung-Box statistics results for residuals and squared residuals with lags 10 and 20

Ljung-Box	$\hat{\varepsilon}_i$	$\hat{\varepsilon_i^2}$
$Q(10)$	17.083.	15.217
(p-value)	0.07254	0.1244
$Q(20)$	30.173	17.893
(p-value)	0.0671	0.5944

From the above results we can confirm that the residuals of Lindley-ACD(1,1) model have no significant serial correlation and conditional heteroscedasticity at 5% level of significance.

The results of this Chapter are reported in Sri Ranganath and Balakrishna (2018).

Chapter 7

Bayesian Analysis of Inverse Gaussian Stochastic Conditional Duration Model

7.1 Introduction

The empirical analysis of durations between the occurrences of certain financial events is important in understanding the market behaviour. To describe the evolution of such durations, Bauwens and Veredas (2004) proposed a class of models called the stochastic conditional duration models. One can refer Engle and Russell (1998), Pacurar (2008) for details on such duration models. The SCD model assumes that the conditional mean of durations between events is generated by a latent stochastic process. The likelihood based inference for such models requires evaluation of multiple integral with respect to latent variables. In view of these difficulties Bauwens and Veredas (2004) estimated the parameters by quasi-maximum likelihood by expressing the model into a linear state space form and then applying Kalman filter method. Feng et al. (2004) proposed an extension to Bauwens and Veredas (2004) SCD model in order to capture the asymmetric behaviour or leverage effect in the duration process. They adopted the Markov Chain Maximum Likelihood(MCML) approach proposed by Durbin and Koopman (1997). Strickland et al. (2006) proposed a Bayesian Markov Chain Monte Carlo method in which the sampling scheme employed is a hybrid of the Gibbs and Metropolis-Hastings algorithm, with the latent vector sampled in blocks. Knight and Ning (2008) discussed the empirical characteristic function(ECF) and the generalized method of moments estimation. Bauwens and Galli (2009) developed a Maximum Likelihood estimation based on the efficient importance sampling (EIS) method to estimate the parameters. Xu et al. (2010) extend the SCD model proposed by Bauwens and Veredas (2004) by imposing mixtures of bivariate normal distributions on the innovations of the observation and latent equations of the duration process. The estimation was carried out by extending the ECF approach of Knight and Ning (2008). Men et al. (2015) introduced a correlation between the error process and the innovations of the duration process and adopted Monte Carlo methods for estimation. Ramanathan et al. (2016) introduced a new procedure for estimation, filtering and smoothing in SCD models, based on estimating functions. Majority of the literature on SCD models assume that the errors follow either a Weibull, Gamma or exponential distribution. Balakrishna and Rahul (2014) proposed a stochastic conditional duration model having inverse Gaussian(IG) distribution for innovations. In this Chapter, we consider the Bayesian analysis of SCD model with inverse Gaussian error random variables. We propose Bayesian MCMC methods to estimate the parameters of the model. Here we follow the algorithm mentioned in Men et al. (2016), Edwards and Sokal (1988) and Neal (2003). Since it is difficult to obtain the analytical conditional densities of observed data, the auxiliary particle filter in Pitt and Shephard (1999) is employed to evaluate the likelihood function.

A brief review of inverse Gaussian SCD model is given in Section 7.2. The Bayesian estimation methodology and MCMC algorithm are discussed in Section 7.3. Simulation studies are carried out in Section 7.4. Section 7.5 presents the application of proposed method to real life data sets.

7.2 The Stochastic Conditional Duration Model

Let τ_i be the time of occurrence of an event (or transaction) of interest with $\tau_0 = 0$ and $X_i = \tau_i - \tau_{i-1}$, $i = 1, 2, ..., n$ be the i^{th} trade duration, which is defined as the waiting time between two consecutive transactions of an underlying asset from time i to $i + 1$.

The SCD model is defined by

$$\begin{aligned} X_i &= e^{\psi_i}\varepsilon_i \\ \psi_i &= \phi\psi_{i-1} + \eta_i\,, \qquad i = 1, 2, \ldots, n \\ \psi_0 &\sim N\left(0, \frac{\sigma^2}{1-\phi^2}\right)\,, \end{aligned} \tag{7.1}$$

where $|\phi| < 1$ to ensure the stationarity of the process and η_i follows independent and identically distributed (iid) $N(0, \sigma^2)$ so that $\{\psi_i\}$ follows a Gaussian AR(1) sequence and $\{\varepsilon_i\}$ is an iid sequence on the positive support with common pdf $f(\varepsilon_i)$ and η_j is independent of ε_i $\forall\, i, j$. Note that the model depends on the unobservable ψ_i, called the latent variable. Most of the SCD models available in the literature assume that the innovations follow iid exponential, gamma or Weibull distributions. One can refer Bauwens and Veredas (2004), Strickland et al. (2006), Knight and Ning (2008), Durbin and Koopman (1997) and so on for the details on such models. Xu et al. (2010) proposed a SCD model by assuming a bivariate mixture of normal distribution for (ε_i, η_i). Balakrishna and Rahul (2014) assumed an inverse Gaussian distribution for ε_i and estimated the model parameters by efficient importance sampling method. Here, we propose a Bayesian method for this SCD model with inverse Gaussian innovations. We consider the unit mean inverse Gaussian distribution for ε_i in the model (7.1).

A random variable Y is said to have an inverse Gaussian distribution with parameters μ and λ, and is denoted by $IG(\mu, \lambda)$ if its probability density function (pdf) is of the form

$$f(y; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi y^3}} \exp\left(-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right), \quad y > 0, \ \ \mu > 0, \ \ \lambda > 0. \tag{7.2}$$

If we restrict ε_i to follow a unit mean inverse Gaussian distribution then its pdf is of the form

$$f_\varepsilon(\varepsilon_i) = \sqrt{\frac{\lambda}{2\pi\varepsilon_i^3}} \exp\left(-\frac{\lambda(\varepsilon_i - 1)^2}{2\varepsilon_i}\right), \qquad \varepsilon_i > 0. \tag{7.3}$$

In order to develop an estimation procedure for the model let $\boldsymbol{X} = (x_1, x_2, ..., x_{n-1}, x_n)'$ be a vector of observations from the model (7.1) and $\boldsymbol{\psi} = (\psi_1, \psi_2, ..., \psi_n)'$ be a vector of associated latent variable, where $'$ denotes the transpose of a vector. If we denote the joint density function of $(\boldsymbol{X}, \boldsymbol{\psi})$ by $f(\boldsymbol{X}, \boldsymbol{\psi}|\boldsymbol{\theta})$, then the likelihood function of the parameter $\boldsymbol{\theta} = (\phi, \sigma, \lambda)'$ based on the observations is

$$L(\boldsymbol{\theta}; \boldsymbol{X}) = \int f(\boldsymbol{X}, \boldsymbol{\psi}|\boldsymbol{\theta}) d\boldsymbol{\psi}, \tag{7.4}$$

which is an n-fold integral with respect to $\boldsymbol{\psi}$. We need to evaluate this integral to obtain the likelihood function. Towards that end, some numerical methods such as MCMC are to be used. In the next Section, we propose a Bayesian method for estimating the parameters.

7.3 Bayesian Estimation

7.3.1 Algorithm

We consider the problem of estimation for the SCD model when the innovations of the durations follow an inverse Gaussian distribution in a Bayesian framework. By Bayes theorem, the joint conditional distribution of $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$ given the observations is

$$f(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{X}) \propto f(\boldsymbol{X}|\boldsymbol{\psi}, \boldsymbol{\theta})f(\boldsymbol{\psi}|\boldsymbol{\theta})f(\boldsymbol{\theta}), \tag{7.5}$$

where $f(\boldsymbol{X}|\boldsymbol{\psi}, \boldsymbol{\theta})$ is the density of $\boldsymbol{X}$ given $(\boldsymbol{\theta}, \boldsymbol{\psi})$, $f(\boldsymbol{\psi}|\boldsymbol{\theta})$ is the density of $\boldsymbol{\psi}$ and $f(\boldsymbol{\theta})$ is the prior density of $\boldsymbol{\theta}$. Bayesian inference on $(\boldsymbol{\theta}, \boldsymbol{\psi})$ is based upon the posterior distribution $f(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{X})$. From this joint posterior, the marginal $f(\boldsymbol{\theta}|\boldsymbol{X})$ can be used to make inferences about the parameters of SCD model with inverse Gaussian innovations and the marginal $f(\boldsymbol{\psi}|\boldsymbol{X})$ provides the inference about the latent variables. We assume that the prior distribution of ϕ, σ, λ are mutually independent. We take prior distribution of ϕ as a normal distribution truncated in the interval $(-1, 1)$ which results in flat density over the supporting region. For σ^2 we take an inverse gamma prior distribution as in Pitt and Shephard (1999) which is a conjugate prior. For the parameter λ, we use a truncated Cauchy prior distribution with density

$$f(\lambda) \propto \frac{1}{1+\lambda^2}, \qquad \lambda > 0 \tag{7.6}$$

which is considered in Bauwens and Lubrano (1998).

To construct an appropriate Markov Chain sampler, the joint posterior should be expressed as proportional to various conditional distributions. So in (7.5), the joint posterior, $f(\boldsymbol{\psi}|\boldsymbol{\theta})$ is further decomposed into a set of conditionals $f(\psi_i|\psi_{-i}, \boldsymbol{\theta}, \boldsymbol{X})$ where ψ_{-i} means all terms of $\boldsymbol{\psi}$ except ψ_i. Here, since $\boldsymbol{\psi}$ is a vector of latent variables, it is difficult to sample from these univariate conditional distribution. So we follow the Markov Chain Monte Carlo algorithm proposed by Men et al. (2016) for estimation.

While developing MCMC algorithm, the latent variables are augmented with the vector of parameters and then the estimation is carried out. We start drawing samples from the conditional posterior distribution of the latent variable $\boldsymbol{\psi}$ and the parameters $\boldsymbol{\theta}$. The sampling algorithm required to draw the samples from each conditional posterior and the required Steps are detailed below.

The parameters $\boldsymbol{\theta} = (\phi, \sigma, \lambda)'$ are initialized with values $(0.5, 0.5, 1)'$. The initial value of $\boldsymbol{\psi}$ is generated from latent AR(1) model by using the above value for $\boldsymbol{\theta}$.

Step 1. Sample $\boldsymbol{\psi} = (\psi_1, \psi_2, ..., \psi_n)$ by drawing ψ_i from $f(\psi_i|\psi_{-i}, \boldsymbol{X}, \boldsymbol{\theta})$ for $i = 1, 2, 3, \ldots, n$. By employing the Markovian structure of the model in (7.1), $f(\psi_i|\psi_{-i}, \boldsymbol{X}, \boldsymbol{\theta})$ can be expressed as $f(\psi_i|\psi_{i-1}, \psi_{i+1}, \boldsymbol{X}, \boldsymbol{\theta})$. A single-move M-H algorithm is used to sample ψ_i. The conditional distribution of latent random variables ψ_i, given the other parameters in the model have been previously sampled is given by

$$f(\psi_1|\boldsymbol{X}, \psi_2, \boldsymbol{\theta}) \propto f(x_1|\psi_1) f(\psi_1|\boldsymbol{\theta}) f(\psi_2|\psi_1, x_1, \boldsymbol{\theta})$$

$$f(\psi_i|\boldsymbol{X}, \psi_{i-1}, \psi_{i+1}, \boldsymbol{\theta}) \propto f(\boldsymbol{X}|\psi_i) f(\psi_i|\psi_{i-1}, x_{i-1}, \boldsymbol{\theta}) f(\psi_{i+1}|\psi_i, x_i, \boldsymbol{\theta}) \;\; i = 2, \ldots, n-1$$

$$f(\psi_n|\boldsymbol{X}, \psi_{n-1}, \boldsymbol{\theta}) \propto f(x_n|\psi_n) f(\psi_n|\psi_{n-1}, x_{n-1}, \boldsymbol{\theta}),$$

where $f(x_i|\psi_i)$, $i = 1, 2, \ldots, n$ are the conditional density functions of the durations, $f(\psi_1|\boldsymbol{\theta})$ is the density of the latent state ψ_1 , $f(\psi_i|\psi_{i-1}, x_{i-1}, \boldsymbol{\theta})$ is the conditional density of ψ_i given ψ_{i-1} and $f(\psi_{i+1}|\psi_i, x_i, \boldsymbol{\theta})$ is the conditional density of ψ_{i+1} given ψ_i. Since x_n is the last observation, the posterior distribution of ψ_n depends only on x_n, x_{n-1} and ψ_{n-1}. The conditional distribution of ψ_1 and ψ_n are given respectively. The conditional distribution of parameter ψ_1 is

$$\begin{aligned} f(\psi_1|\boldsymbol{X}, \psi_2, \boldsymbol{\theta}) &\propto f(x_1|\psi_1) f(\psi_1|\boldsymbol{\theta}) f(\psi_2|\psi_1, x_1, \theta) \\ &= f(x_1|\psi_1) \exp\left(-\frac{(1-\phi^2)\psi_1^2}{2\sigma^2}\right) \exp\left(-\frac{(\psi_2 - \phi\psi_1)^2}{2\sigma^2}\right). \end{aligned}$$

The conditional distribution of parameter ψ_n is

$$\begin{aligned} f(\psi_n|\boldsymbol{X}, \psi_{n-1}, \boldsymbol{\theta}) &\propto f(x_n|\psi_n) f(\psi_n|\psi_{n-1}, x_{n-1}, \theta) \\ &\propto f(x_n|\psi_n) \exp\left(\frac{(\psi_n - \phi\psi_{n-1})^2}{2\sigma^2}\right). \end{aligned}$$

If the conditional distribution does not possess a simple form as in the present case then it is not possible to draw the samples directly. In such cases one of the obvious choice is to consider an accept/reject method. Following Men et al. (2016) we use a single move Metropolis Hastings algorithm to sample the latent states, where the proposal distribution is simulated by slice sampler method. The slice sampler method proposed by Edwards and Sokal (1988) and Neal (2003) is a method to sample random variables which do not have simple probability density functions or their probability density functions are known up to a normalizing constant. As the slice sampler adapts to the analytical structure of the underlying density, it is more efficient. Also it ensures faster convergence to the underlying distribution. So, to generate random variates from the conditional distribution we employ the method of slice sampler. In the case of analysing a data, to obtain the estimate of these unobservable latent variables $\boldsymbol{\psi}$, we use Auxiliary Particle Filter proposed by Pitt and Shephard (1999). In the following discussion, we obtain specific forms of the required samplers.

For $i = 2, 3, \ldots, n-1$,

$$f(\psi_i|\psi_{i-1}) \sim N(\phi\psi_{i-1}, \sigma^2) \text{ and } f(\psi_{i+1}|\psi_i) \sim N(\phi\psi_i, \sigma^2).$$

Substituting for the corresponding normal densities and on simplifying the square terms we get

$$f(\psi_i|\psi_{-i}, \boldsymbol{\theta}) \propto \exp\left(-\frac{(1+\phi^2)}{2\sigma^2}\left(\psi_i^2 - 2\psi_i\left(\frac{\phi}{\phi^2+1}(\psi_{i-1}+\psi_{i+1})\right)\right)\right)$$

$$\propto N(\omega_1, \sigma_1^2)$$

where $\omega_1 = \frac{\phi}{1+\phi^2}\left(\psi_{i-1}+\psi_{i+1}\right)$ and $\sigma_1^2 = \frac{\sigma^2}{1+\phi^2}$.

Hence the conditional posterior distribution of ψ_i can be represented as

$$
\begin{aligned}
f(\psi_i|\boldsymbol{X}, \psi_{i-1}, \psi_{i+1}, \boldsymbol{\theta}) \propto & f(x_i|\psi_i) f(\psi_i|\psi_{i-1}, \boldsymbol{\theta}) f(\psi_{i+1}|\psi_i, \boldsymbol{\theta}) \qquad (7.7)\\
\propto & \sqrt{\frac{\lambda e^{\psi_i}}{x_i^3}} \exp\left(-\frac{\lambda(x_i - e^{\psi_i})^2}{2e^{\psi_i} x_i}\right) \exp\left(-\frac{(\psi_i - \phi\psi_{i-1})^2}{2\sigma^2}\right)\\
& \exp\left(-\frac{(\psi_{i+1} - \phi\psi_i)^2}{2\sigma^2}\right)\\
\propto & (\lambda \exp(\psi))^{\frac{1}{2}} \exp\left(-\frac{\lambda(x_i - e^{\psi_i})^2}{2e^{\psi_i} x_i}\right) \exp\left(-\frac{(\psi_i - a_i)^2}{2b}\right)
\end{aligned}
\tag{7.8}
$$

where $a_i = \frac{\phi(\psi_i + \psi_{i+1})}{1+\phi^2}$ and $b = \frac{\sigma^2}{1+\phi^2}$.

In (7.8), the posterior distribution is proportional to a product of three positive functions that cannot be simulated directly. So we propose a method of slice sampler and summarize its algorithm below.

Algorithm for slice sampler for ψ_i. Let us rewrite the conditional distribution in (7.8) as

$$
g(\psi_i) \propto \exp\left(-\frac{\lambda(x_i - e^{\psi_i})^2}{2e^{\psi_i} x_i}\right) \exp\left(-\frac{(\psi_i - a_{1i})^2}{2b}\right), \tag{7.9}
$$

where $a_{1i} = a_i + \frac{b}{2}$ and follow the algorithm given below.

To start the slice sampling procedure, the sampled value of ψ_i from the last MCMC step is set as the initial value.

i Initialize $\psi_i^{(0)}$. Set t=0.

ii Draw a random observation u_1 uniformly from the interval

$$
\left(0, \exp\left(-\frac{\lambda(x_i - e^{\psi_i^{(t)}})^2}{2e^{\psi_i^{(t)}} x_i}\right)\right).
$$

Then we define an interval for ψ_i through the inequality $u_1 \leq \exp\left(-\frac{\lambda(x_i - e^{\psi_i})^2}{2e^{\psi_i}x_i}\right)$, which is equivalent to

$$\psi_i \geq log\left(\frac{x_i(1+\frac{1}{\varepsilon_i^2})}{2\left(1-\frac{(log(u_1))}{\lambda}\right)}\right). \tag{7.10}$$

iii Similarly draw a random observation u_2 uniformly from the interval $\left(0, \exp\left(-\frac{(\psi_i - a_{1i}^{(t)})^2}{2b}\right)\right)$, where $a_{1i}^{(t)}$ is calculated from (7.9). We define an interval for ψ_i through the inequality $u_2 \leq \exp\left(-\frac{(\psi_i - a_{1i})^2}{2b}\right)$, which is equivalent to

$$a_{1i}^{(t)} - \sqrt{-2\,b\,log(u_2)} \leq \psi_i \leq a_{1i}^{(t)} + \sqrt{-2\,b\,log(u_2)}. \tag{7.11}$$

iv Draw $\psi_i^{(t+1)}$ uniformly from the interval $\left(max\left\{log\left(\frac{x_i(1+\frac{1}{\varepsilon_i^2})}{2\left(1-\frac{(log(u_1))}{\lambda}\right)}\right), a_{1i}^{(t)} - \sqrt{-2\,b\,log(u_2)}\right\}, a_{1i}^{(t)} + \sqrt{-2\,b\,log(u_2)}\right)$ determined by the inequalities (7.10) and (7.11).

v Stop, if a stopping criterion is met; otherwise, set $t = t + 1$ and repeat from ii.

Step 2. Sample ϕ.

The prior distribution of ϕ is assumed to follow a univariate normal distribution truncated in the interval (-1,1). Given a truncated normal prior distribution, $\phi \sim N(\alpha_\phi, \beta_\phi^2)$ and the other parameters in the model have been previously sampled, the conditional distribution of ϕ is The conditional distribution of parameter ϕ is

$$
\begin{aligned}
f(\phi|\boldsymbol{X},\sigma,\lambda) &\propto f(\boldsymbol{\psi}|\boldsymbol{\theta},\boldsymbol{X})f(\phi) \qquad (7.12)\\
&= f(\psi_1|\boldsymbol{\theta},\boldsymbol{X})\prod_{i=2}^{n} f(\psi_i|\psi_{i-1},\boldsymbol{\theta},\boldsymbol{X})\exp\left(-\frac{(\phi-\alpha_\phi)^2}{2\beta_\phi^2}\right)\\
&\propto \exp\left(-\frac{(1-\phi^2)\psi_1^2}{2\sigma^2}\right)\exp\left(-\frac{\sum_{i=1}^{n}(\psi_i-\phi\psi_{i-1})^2}{2\sigma^2}\right)\\
&\qquad \exp\left(-\frac{(\phi-\alpha_\phi)^2}{2\beta_\phi^2}\right)(1-\phi^2)^{\frac{1}{2}} \qquad (7.13)\\
&\propto \exp\left(-\frac{1}{2}\left(\phi^2\left\{\frac{\sum_{i=1}^{n}\psi_1^2}{\sigma^2}+\frac{1}{\beta_\phi^2}\right\}-2\phi\left\{\frac{\sum_{i=1}^{n}\psi_i\psi_{i-1}}{\sigma^2}+\frac{\alpha_\phi}{\beta_\phi^2}\right\}\right)\right)\\
&\qquad (1-\phi^2)^{\frac{1}{2}} \qquad (7.14)\\
&\propto N\left(\frac{d}{c},\frac{1}{c}\right)\sqrt{1-\phi^2},
\end{aligned}
$$

where $c=\frac{\sum_{i=1}^{n}\psi_1^2}{\sigma^2}+\frac{1}{\beta_\phi^2}$, $\quad d=\frac{\sum_{i=2}^{n}\psi_i\psi_{i-1}}{\sigma^2}+\frac{\alpha_\phi}{\beta_\phi^2}$. It is proportional to the product of a normal distribution and a positive function. Hence we can use the slice sampling method to sample the parameter ϕ.

Step 3. Sample σ^2.

We sample σ^2 by taking an inverse gamma prior distribution i.e., $\sigma^2 \propto$ inverse Gamma$(\alpha_{\sigma^2},\beta_{\sigma^2})$, where α_{σ^2} and β_{σ^2} are hyperparameters. As the prior for σ^2 is a conjugate prior the sampling is carried out by simulating from the inverse Gamma distribution with the corresponding parameters obtained. The conditional distribution of parameter σ^2 , is given by

$$
\begin{aligned}
f(\sigma^2|\boldsymbol{X},\phi,\lambda) &= f(\boldsymbol{\psi}|\boldsymbol{\theta},\boldsymbol{X})f(\alpha_{\sigma^2},\beta_{\sigma^2}) \\
&= f(\psi_1|\theta,\boldsymbol{X})\prod_{i=2}^{n} f(\psi_i|\psi_{i-1},\boldsymbol{\theta},\boldsymbol{X})f(\alpha_{\sigma^2},\beta_{\sigma^2}) \\
&\propto \exp\left(-\frac{\psi_i^2(1-\phi^2)}{2\sigma^2}\right)\times\exp\left(\frac{\sum_{i=2}^{N}(\psi_i-\phi\psi_{i-1})^2}{2\sigma^2}\right) \\
&\qquad \times(\sigma^2)^{-\frac{n}{2}}\times\frac{(\beta_{\sigma^2})^{\alpha_{\sigma^2}}}{\Gamma(\alpha_{\sigma^2})}(\sigma^2)^{-\alpha_{\sigma^2}-1}\exp(\frac{-\beta_{\sigma^2}}{\sigma^2}). \\
&\propto (\sigma^2)^{-\frac{(2\,\alpha_{\sigma^2})-n+2}{2}}\exp\left(-\frac{\psi_i^2(1-\phi^2)-\sum_{i=2}^{N}(\psi_i-\phi\psi_{i-1})^2+2\,\beta_{\sigma^2}}{2\,\sigma^2}\right) \\
&\propto inverse\,Gamma(a,b),
\end{aligned}
$$

where $a=\alpha_{\sigma^2}+\frac{n}{2}$ and $b=\frac{(1-\phi^2)(\psi_i)^2}{2}+\frac{1}{2}\sum_{i=2}^{n}(\psi_i-\phi\psi_{i-1})^2+\beta_{\sigma^2}$.

Step 4. Sample λ.

The conditional distribution of λ is

$$
\begin{aligned}
f(\lambda|\boldsymbol{X},\boldsymbol{\psi},\phi,\sigma^2) &= f(\boldsymbol{X}|\boldsymbol{\psi},\lambda)f(\lambda) \\
&= f(\lambda)\prod_{i=1}^{n}\sqrt{\frac{\lambda e^{\psi_i}}{x_i^3}}\exp\left(-\frac{\lambda(x_i-e^{\psi_i})^2}{2e^{\psi_i}x_i}\right),
\end{aligned}
\tag{7.15}
$$

where $f(\lambda)$ is a prior density of λ, given by (7.6). Here the samples cannot be simulated directly from (7.15). So we use a random walk MH algorithm with standard normal distribution as the proposal distribution to sample λ . The acceptance probability is computed using equation (7.15).

In summary, the sampling procedure for $(\boldsymbol{\theta}',\boldsymbol{\psi}')'$ is as follows :

- Sample $\boldsymbol{\psi}'$ using the single-move Metropolis Hastings algorithm with proposal distribution simulated by the method of slice sampling which is briefly explained in Step 1.
- Sample ϕ from (7.12) following the explanations in Step 2.
- Sample σ^2 directly from the inverse Gamma density using Step 3.
- Sample λ using a random walk MH algorithm with the acceptance probability computed through (7.15) given in Step 4.

In the next section we demonstrate the applications of the above methods through a simulation study.

7.4 Simulation study

A simulation study is carried out to assess the performance of the Bayes estimators, described in the previous Section. From (7.1), we generate 5000 observations. Then the MCMC algorithm discussed in Section 7.3.1 is run and first 25000 iterations were discarded as burn-in from 100000 iterations. The parameters are estimated and the simulation results are tabulated in Table 7.1. The plots of histograms of posterior samples of ϕ ,σ, λ are shown in Figure 7.1 (a), 7.1 (b) and 7.1 (c) respectively. The trace plots are shown in Figure 7.2 (a), 7.2 (b) and 7.2 (c) respectively.

To illustrate the application of Bayesian estimation method we analyse two sets of data. We perform model diagnosis based on the residuals to explain model adequacy. The residuals of SCD model with inverse Gaussian innovations are defined

TABLE 7.1: True and estimated parameters of the IG-SCD model

	ϕ	σ	λ
True.	0.65	1.5	1
Est.	0.651699	1.49213	1.09243
mse.	0.0000155	0.0000291	0.000047
HPD CI(95%)	(0.64943, 0.66751)	(1.47612, 1.54811)	(0.96213, 1.1321)
	ϕ	σ	λ
True.	0.75	1.5	1
Est.	0.763302	1.529561	0.999191
mse.	0.000039	0.000094	0.0000039
HPD CI(95%)	(0.74633, 0.78840)	(1.47808, 1.57561)	(0.9557, 1.1477)
True.	0.85	1.5	1
Est.	0.84699	1.55057	0.999242
mse.	0.0000322	0.000095	0.0000041
HPD CI(95%)	(0.82633, 0.86840)	(1.49808, 1.58561)	(0.9657, 1.1413)
True.	0.90	1.5	1
Est.	0.895114	1.569688	1.052775
mse.	0.000105	0.0000372	0.000028
HPD CI(95%)	(0.88438, 0.91799)	(1.481613, 1.59031)	(0.99861, 1.06868)

*Highest Probability Density Confidence Interval (HPD CI)

as $\hat{\varepsilon}_i = x_i/e^{\hat{\psi}_i}$, where $\hat{\psi}_i$ is the estimator of ψ_i. The estimators of the parameters ϕ, σ and λ can be obtained by the MCMC algorithm discussed in Section 7.3. To obtain the estimates of the unobservable component ψ_i, we employ an auxiliary particle filter proposed by Pitt and Shephard (1999). This method is described in the following subsection.

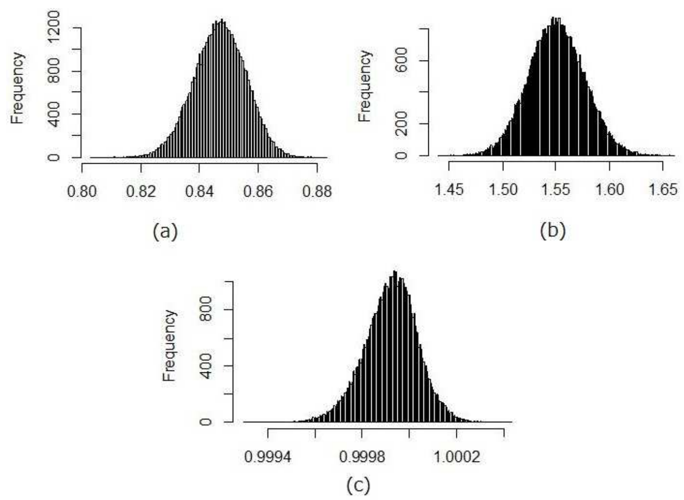

FIGURE 7.1: Histogram of the posterior samples using simulated data.

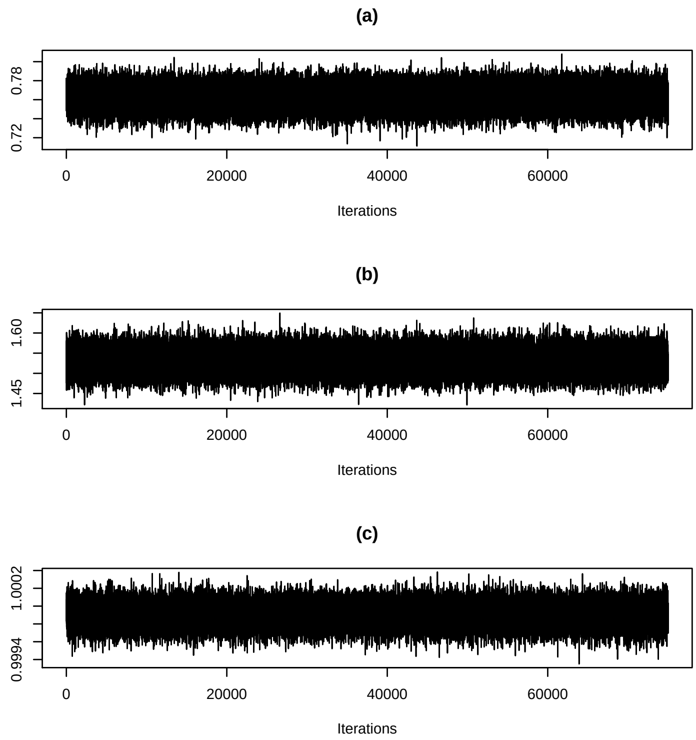

FIGURE 7.2: Trace plots of the posterior samples using simulated data.

Particle Filter

Particle filters are a class of simulation-based filters that recursively approximate the filtering distribution using a collection of particles with some probability masses. The particles are samples of unknown states from the state space, and the particle weights are probability mass computed by Bayes theory. The basic idea is the recursive computation of relevant probability distribution and approximation of probability distribution.

By successive conditional decomposition, the likelihood of the IG-SCD model is

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = f(x_1|\boldsymbol{\theta}) \prod_{i=2}^{n} f(x_i|\mathcal{F}_{i-1}, \boldsymbol{\theta}), \tag{7.16}$$

where $\mathcal{F}_i = \sigma(x_1, x_2, \ldots, x_i)$, the sigma field generated by $(x_1, x_2, \ldots, x_i)$ is the information known at time i. The conditional density of x_{i+1} given $\boldsymbol{\theta}$ and $\mathcal{F}_i$ is given by

$$\begin{aligned} f(x_{i+1}|\mathcal{F}_i, \boldsymbol{\theta}) &= \int f(x_{i+1}|\psi_{i+1}, \boldsymbol{\theta}) dF(\psi_{i+1}|\mathcal{F}_i, \boldsymbol{\theta}) \\ &= \int f(x_{i+1}|\psi_{i+1}, \boldsymbol{\theta}) f(\psi_{i+1}|\psi_i, \boldsymbol{\theta}) dF(\psi_i|\mathcal{F}_i, \boldsymbol{\theta}). \end{aligned} \tag{7.17}$$

The difficulty in obtaining an analytical form of the above integral leads to the utilization of APF. Suppose that we have a particle sample $\{\psi_i^{(t)}, t = 1, 2, \ldots, N\}$ of ψ_i from the filtered distribution $(\psi_i|\mathcal{F}_i, \theta)$ with weights $\{\pi_t, t = 1, 2, \ldots, N\}$ such that $\sum_{t=1}^N \pi_t = 1$. From this sample, the one-step ahead predictive density of ψ_{i+1} is

$$f(\psi_{i+1}|\mathcal{F}_i, \boldsymbol{\theta}) \approx \sum_{t=1}^{N} \pi_t f(\psi_{i+1}|\psi_i^{(t)}, \boldsymbol{\theta}) \tag{7.18}$$

The one step ahead prediction distribution of ψ_{i+1} can then be sampled and the conditional density (7.17) can be evaluated numerically by

$$f(x_{i+1}|\mathcal{F}_i, \boldsymbol{\theta}) \approx \sum_{t=1}^{N} \pi_t f(x_{i+1}|\psi_{i+1}^{(t)}, \boldsymbol{\theta}), \tag{7.19}$$

where $\psi_{i+1}^{(t)}$ are particles from the prediction distribution of $(\psi_{i+1}|\mathcal{F}_i, \boldsymbol{\theta})$. The predictive density of ψ_{i+1} should be known for the approximation (7.18) to be feasible. From latent AR(1) process, we have ψ_{i+1} has a conditional normal distribution $\psi_{i+1} \sim N(\phi\psi_i, \sigma^2)$.

Given the particle sample from a filtered distribution $(\psi_i|\mathcal{F}_i, \boldsymbol{\theta})$ we need to sample $(\psi_{i+1}|\mathcal{F}_{i+1}, \boldsymbol{\theta})$. For that, we follow the procedure of Chib et al. (2006) and Men et al. (2016), which is summarized below.

Algorithm for APF

1. (a) Given a sample $\{\psi_i^{(t)}, t = 1, 2, \ldots, N\}$ from $(\psi_i|\mathcal{F}_i, \boldsymbol{\theta})$, calculate the expectation $\hat{\psi}_{i+1}^{*(t)} = E(\psi_{i+1}|\psi_i^{(t)})$ and

$$\pi_t = f(x_{i+1}|\hat{\psi}_{i+1}^{*(t)}, \boldsymbol{\theta}), \qquad t = 1, 2, \ldots, N$$

and sample N times with replacement the integers 1,2,. . . ,N with probability $\hat{\pi}_t = \frac{\pi_t}{\sum_{t=1}^{N} \pi_t}$. Let the sampled indexes be $k_1, k_2, \ldots, k_N$ and associate these with particles $\{\hat{\psi}_i^{*(k_1)}, \hat{\psi}_i^{*(k_2)}, \ldots, \hat{\psi}_i^{*(k_N)}\}$.

2. For each value of k_t from 1(a), sample the values $\{\psi_{i+1}^{*(1)}, \ldots, \psi_{i+1}^{*(N)}\}$ from

$$\psi_{i+1}^{*(t)} = \phi\psi_i^{(k_t)} + \eta_{i+1}, \qquad t = 1, \ldots, N$$

where $\eta_{i+1} \sim N(0, 1)$.

3. Calculate the weights of the values $\{\psi_{i+1}^{*(1)}, \ldots, \psi_{i+1}^{*(N)}\}$ as

$$\pi_t^* = \frac{f(x_{i+1}|\psi_{i+1}^{*(t)}, \boldsymbol{\theta})}{f(x_{i+1}|\hat{\psi}_{i+1}^{*(k_t)}, \boldsymbol{\theta})}, \qquad t = 1, 2, \ldots, N$$

and using these weights resample the values $\{\psi_{i+1}^{*(1)}, \ldots, \psi_{i+1}^{*(N)}\}$ N times with replacement to obtain a sample $\{\psi_{i+1}^{(1)}, \ldots, \psi_{i+1}^{(N)}\}$ from the filtered distribution $(\psi_{i+1}|\mathcal{F}_{i+1}, \boldsymbol{\theta})$. Here we use N=2000.

7.5 Data Analysis

We demonstrate the applications of the model, by analysing two sets of data. Data sets are on intraday trades of IBM OHLC bar data downloaded from Algoseek Website and intraday trades of US Brent Crude Oil downloaded from the Website of a Swiss Forex bank. Only trades between 9:30:00 am and 4:00:00 pm are recorded as this is the normal trading hour.

7.5.1 IBM trades data

The model is applied to intraday IBM trades data as on 16^{th} June 2015. Consider the 1 second Trade OHLC Bar data with a sample size of 6708 observations. The data sets are obtained from Website of Algoseek. The trade durations are defined as the time intervals between consecutive trades, measured in seconds. We ignore the zero trade duration and the time plot of the nonzero intra-day IBM trade duration series is shown in Figure 7.3. Now, removing the effect of the diurnal pattern we take the adjusted time duration (Tsay (2014), pp 298-300) to model the intraday pattern. In Table 7.2 the summary statistics of IBM trades data are given.

Table 7.2: Descriptive statistics for IBM Trades data

Statistic	IBM Trades data
Sample size	6708
Minimum	1
Maximum	37
Mean	3.48
Median	2

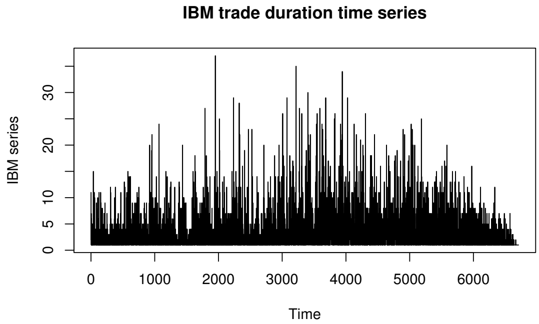

FIGURE 7.3: IBM trade duration time series.

The estimation of the parameters is carried out by the MCMC algorithm described in Section 7.3.1 and the estimates are provided in Table 7.3. For model diagnosis we compute residuals, $\hat{\varepsilon}_i = x_i/e^{\hat{\psi}_i}$, where $\hat{\psi}_i$ is the estimator of ψ_i which is obtained by APF. If the fitted model is adequate then the acf of $\{\hat{\varepsilon}_i\}$ will be negligible. The residual acf plot given in Figure 7.4 indicates that they are uncorrelated. Bauwens and Veredas (2004) considered the time series version of Spearman's ρ correlation coefficient and the p-value plots instead of the Ljung-Box statistic. Here we follow a rank portmanteau statistic given in Dufour and Roy (1986) to check the lack of autocorrelation of the obtained residuals. The p-value obtained is 0.7. Also the run test confirms the independence of the residuals $\{\hat{\varepsilon}_i\}$. In Figure 7.5 the histogram of the residuals is superimposed by the Inverse Gaussian density curve for the IG-SCD model. So it can be concluded that the fitted model is adequate for explaining the dynamics, which generated the data.

TABLE 7.3: Estimated parameters of the IG-SCD model based on IBM Trades duration data.

	ϕ	σ	λ
Est.	0.69803	1.2366	1.1566
std error.	0.00143	0.00906	0.00658

FIGURE 7.4: ACF plot of residuals

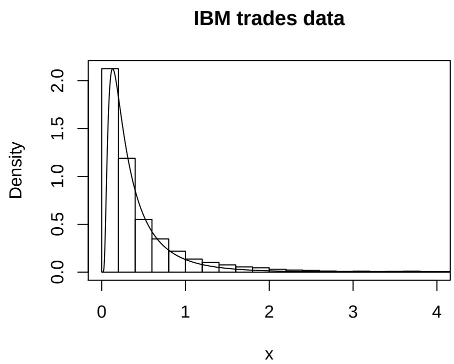

FIGURE 7.5: Histogram of residuals superimposed by inverse Gaussian density for IG-SCD model.

7.5.2 US Brent Crude Oil

The second set of data is the intraday trades data of US Brent Crude Oil downloaded from the Website of a Swiss Forex bank and Marketplace. The intraday trade of the Brent Crude Oil on 20 February 2017 is considered. Taking the normal trading hours and ignoring the zero durations the sample size obtained is 1625 trade durations. The time plot of the nonzero intra-day durations and the time plot of the adjusted duration series is shown in Figure 7.6 and 7.7. The summary of data is given in Table 7.4.

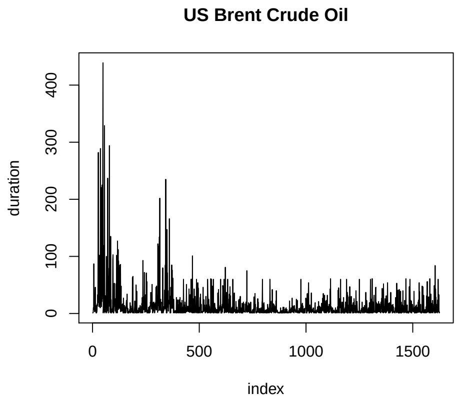

FIGURE 7.6: Duration plot of US Brent Crude Oil.

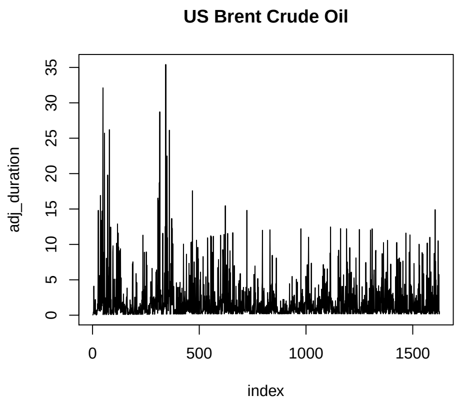

FIGURE 7.7: Adjusted Duration plot of US Brent Crude Oil.

TABLE 7.4: Descriptive statistics for US Brent Crude Oil Trades data

Statistic	IBM Trades data
Sample size	1625
Minimum	1
Maximum	439
Mean	14.4
Median	6

The parameters are estimated by the algorithm mentioned in Section 3 and is tabulated in Table 7.5. From the residual plot given in Figure 7.8 and the p-value(=0.8)

TABLE 7.5: Estimated parameters of the IG-SCD model based on US Brent Crude Oil trades duration data.

	ϕ	σ	λ
Est.	0.77263	1.34885	0.945069
std error.	0.00567	0.00975	0.01018

obtained from rank portmanteau statistic we conclude that the residuals are independent. In Figure 7.9 the histogram of the residuals is superimposed by the inverse Gaussian density curve and is of good fit.

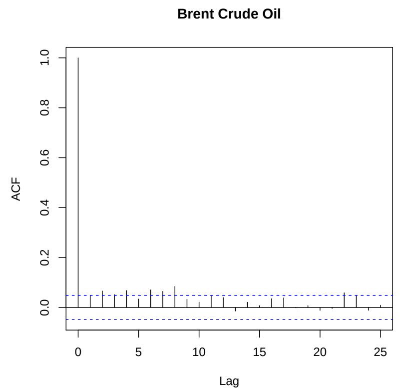

FIGURE 7.8: Residual plot of US Brent Crude Oil.

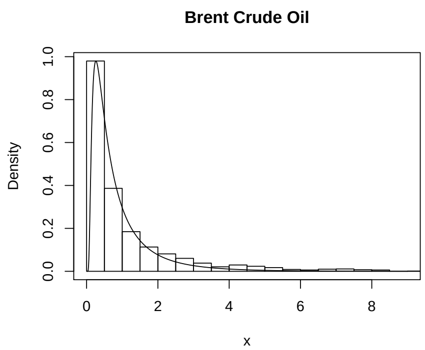

FIGURE 7.9: Histogram of residuals superimposed by inverse Gaussian density for IG-SCD model.

The results of this Chapter are reported in Sri Ranganath and Balakrishna(2017b).

Chapter 8

Conclusions and Future Works

The classical method of time series deals with the linear models with Gaussian errors. Many real life situations, in particular the financial time series cannot be explained by the Box-Jenkins methodology. So, to study and understand the behaviour of financial data, many non-Gaussian non-linear models have been introduced in the literature. In this thesis, we propose non-Gaussian models and studied their suitability to model non-Gaussian time series data. The study of financial data also reveal the absence of correlation among returns but the absolute or squared returns shows significant correlation. This type of behaviour cannot be modelled with a linear model. So, to model the conditional variance we propose stochastic volatility model and conditional duration model and examined their applicability to model financial data. Also we propose MCMC Bayesian estimation procedure to inverse gaussian conditional duration model.

We have proposed an ARMA model with Generalized Error Distribution innovations. The properties of the proposed model are studied. The detailed analysis of ARMA(1,1), MA(1) and AR(1) models are considered. Estimation of AR(1) model is carried out by Generalized Method of Moments and the method of Maximum Likelihood. Simulation studies are carried out to assess the performance of the estimators. Asymptotic properties of the estimators are established using GMM. Two sets of data are analysed to illustrate the application of the proposed model. We found that the proposed model is able to explain the characteristics of the data. The validity of the model is mainly performed with statistical plots. More theories are to be developed to diagnose the model under non-normal errors. Also, due to the non differentiability of the likelihood function, the asymptotic properties of Maximum Likelihood estimates are established for $\beta = 2$ and $\beta > 3$. We hope to consider some other efficient estimation methods and will be taken up in the future work.

The problem of estimation is an important stage in stochastic modelling. In autoregressive model of first order, the least square estimator is chosen frequently to estimate the autoregressive parameters. Due to the estimation problem associated with the least square estimator, many authors study the problem of obtaining a robust estimate to estimate the autoregressive parameter. In Chapter 4, we propose Hurwicz estimator to estimate the autoregressive coefficient in an AR(1) model with GED innovations. A Monte Carlo simulation is carried out to study the nature of LS and Hurwicz estimator when the errors follow GED. A comparison study of the two estimators is done using the statistical measures such as bias, mean square error and mean absolute deviation. The performance of LS and Hurwicz estimator are analysed with respect to Pitman-Closeness Criterion. Also, a simulation study on the coverage level of bootstrap prediction interval and length of intervals is done. From the above studies we found that Hurwicz estimator performs better in heavy tail distribution than LS estimator. The asymptotic distribution of the proposed estimator is analysed through parametric bootstrap methods. As a future work, we plan to study and establish its theoretical asymptotic properties.

A number of models have been introduced to describe the evolution of conditional variances in modelling stochastic volatilities. We propose a Lindley distributed errors to define an SV model. The parameters of the model are estimated using method of moments and the asymptotic properties are studied. Also the applicability of the model is illustrated through real data set. Some other efficient methods like MCMC Bayesian approach or Efficient Importance Sampling will be handier here.

The analysis and modelling of conditional durations is of great importance in financial data to study the market behaviour. We propose Lindley ACD model to analyse the transaction duration. The increasing nature of the hazard function makes it as an alternative to exponential distribution. The properties of the proposed model are studied and the parameters are estimated by conditional maximum likelihood estimation method. To illustrate the model, we have analysed a real data. Other estimation methods can be accommodated to check the performance of the estimators. The increasing nature of the hazard function makes it as an alternative to exponential distribution. As an additional choice to analyse transaction durations in financial point process, a Stochastic Conditional Duration model based on the Lindley distribution can be studies, which will be taken up in the future work.

One of the problems in dealing with the stochastic conditional duration models is its estimation. The likelihood based inference for such models needs the evaluation of multiple integral with respect to latent variable. We propose Bayesian MCMC estimation methods to estimate the parameters of Inverse Gaussian SCD model. Simulation studies are conducted to check the performance of the estimators. Two sets of data are analysed and proposed method performs well. The diagnosis tools employed in this model needs to be developed. More rigorous statistical tests are to be developed to perform model diagnosis in the presence of latent variables, which will be of great interest and we hope to consider this for our future work.

Appendix A

R code for estimation of parameters of GED-AR model

$n1 = 2100$

$ph = 0.3;$

$mu = 0;$

$si = 2;$

$p = 1.3;$

$s = si/(p^{(1/p)});$

$p1 = c()$

s1=c()

phi1=c()

r1=c()

for(j in 1:100)

{ e=rnormp(n1,mu,s,p)

x=c()

x[1]=1;

$for(o\, in\, 2 : n1)\{$

$x[o] = ph * x[o-1] + e[o]\ \}$

$x1 = x[1001 : n1]$

$n = length(x1)$ $r1[j] = sum((x1[2:n]) * (x1[1:n-1]))/sum((x1[1:n])^2);$

$p3 = seq(.1, 5, .001);$

$fnp3 = (gamma(3/p3)^2)/((gamma(1/p3)^2) * ((((6 * r1[j]^2) * (gamma(3/p3)^2))/$

$((1 - r1[j]^2) * (gamma(1/p3)^2))) + (gamma(5/p3)/gamma(1/p3))))$

$con = ((mean(x1^2))^2 * (1 - r1[j]^2))/(mean(x1^4) * (1 + r1[j]^2));$

fn=abs(fnp3-con);

val=which.min(fn);

p1[j]=p3[val]

$s1[j] = sqrt((mean(x1^2) * (1 - r1[j]^2) * (gamma(1/p1[j]))))/gamma(3/p1[j]))$

}

$mean(r1);$

$sd(r1);$

$mean(s1);$

$sd(s1);$

$mean(p1);$

$sd(p1);$

Appendix B

R code for estimation of parameters of GED-AR(1) Model taking Hurwicz estimator

---

library(normalp)

$n1 = 450;$

$ph = 0.3;$

$mu = 0;$

$si = 1.5$

$p = 2;$

$s = si/(p^{(1/p)});$

$phi2 = c();$

$s2 = c();$

$p2 = c();$

$for(t\, in\, 1 : 500)$

{

$e = rnormp(n1, mu, s, p)$

$x = c()$

$x[1] = 1;$

$for(o\, in\, 2 : n1)$

$\{$

$x[o] = ph * x[o - 1] + e[o]$

$\}$

$x1 = x[401 : n1]$

$n = length(x1)$

$phi1 = c()$

$s1 = c()$

$p1 = c()$

$r1 = sum((x1[2 : n] - mean(x1)) * (x1[1 : n - 1]-$

$mean(x1)))/sum((x1[1 : n] - mean(x1))^2)$

$phi1[1] = r1;$

$s1[1] = sum((abs(x - mu))^2)/n * (gamma(1/p)/gamma(3/p));$

$p1[1] = p;$

$for(k\, in\, 1 : 20)$

$\{$

$phi = c();$

$phih = r1;$

$d = ((abs(x1[2 : n] - (phih * x1[1 : n - 1])))/s1[k])$

$p0 = c()$

$p0 = p$

$for(j\, in\, 1 : 10)$

$\{$

$fp = (n/p0[j]) + ((n/(p0[j] * p0[j]))*$

$digamma(1/p0[j])) - sum((d^{p0[j]}) * (log(d)))$

$fdp = -(n/(p0[j]^2)) - (((2*n)/(p0[j]*p0[j]*p0[j]))*digamma(1/p0[j]))$

$-((n/(p0[j]^4))*trigamma(1/p0[j])) - sum((d^{p0[j]}) * ((log(d))^2))$

$p0[j+1] = p0[j] - (fp/fdp)$

$if((p0[j+1] - p0[j]) < .00001)\, pp = p0[j+1]$

}

$sh = ((pp/n) * sum((abs(x1[2:n] - (phih * x1[1:n-1])))^{pp}))^{(1/pp)}$

$phi1[k+1] = phih$

$s1[k+1] = sh$

$p1[k+1] = pp$

$if((phi1[k+1] - phi1[k]) < 0.001)\, phih1 = phi1[k+1]$

$if((p1[k+1] - p1[k]) < 0.001)\, pp1 = p1[k+1]$

$if((s1[k+1] - s1[k]) < 0.001)\, sh1 = s1[k+1]$

}

$s2[t] = sh1$

$p2[t] = pp1$

$phi2[t] = phih1$

}

$mean(phi2)$

$var(phi2)$

$mean(s2)$

$var(s2)$

$mean(p2)$

$var(p2)$

$mean(abs(ph - phi2))$

$mean(abs(si - s2))$

$mean(abs(p - p2))$

Appendix C

R code for estimation of parameters of Lindley SV model

Estimation of Lindley-SV

$--$

$rm(list = ls(all = TRUE))$

#Quadratic equation function to solve theta

$--$

$result < -function(a, b, c)$

{

$if(delta(a, b, c) > 0)${#first case D >0

$x_1 = (-b + sqrt(delta(a, b, c)))/(2 * a)$

$x_2 = (-b - sqrt(delta(a, b, c)))/(2 * a)$

$result = c(x_1, x_2)$

} $elseif(delta(a, b, c) == 0)$

{#second case D=0

$x = -b/(2 * a)$

}

$else${"There are no real roots."}#third case D<0

} #Constructing delta

$delta <- function(a, b, c)$

{

$b^2 - 4 * a * c$

} #MAIN program

$n1 = 2000$

$y = rep();$

$h1 = rep();$

$h = rep();$

$eta = c();$

$phihat = rep();$

$thetahat = rep();$

$y1 = rep();$

$biasphi = rep();$

$e = rep()$

$m1 = rep();$

$v1 = rep();$

$m2 = rep();$

$v2 = rep();$

$c1 = rep();$

$eta1 = rep();$

$u = rep();$

$theta = 0.1;$

$phi = 0.1;$

$h[1] = 1;$ initial valuues

$w1 = ((theta * theta) * ((1 - phi) * (1 - phi))) + (theta * (1 - (phi * phi))) + (2 * phi)$

$w2 = (1 - phi); w3 = phi;$

$w4 = (theta * (1 - phi) + 1);$

$w2 = (1 - phi);$

$w3 = phi * (theta + 1);$

$w4 = (theta * (1 - phi) + 1);$

$(w1/(w4 * w4)) + (w2/w4) - (w3/((w4 * w4))) \quad sumofweights = 1$

$for(kin1 : 1000)$

{ #k loop begins for replications

$for(i\, in\, 1 : n1)\{$

$u[i] = runif(1)$

$if(u[i] < phi)\{$

$eta1[i] = 0$

}

$else\{\ eta1[i] = rgamma(1, 2, (1/theta)) +$

$rexp(1, theta) + rexp(1, ((theta + 1)/phi))$

}

}

$y1[1] = e[1] * h[1]$

$for(j\, in\, 2 : n1)$

{

$h[j] = phi * h[j - 1] + eta1[j]$

}

$g = (h^{(0.5)})$

$e = rnorm(n1, 0, 1);$

$y1 = g * e;$

$y = y1[1001 : n1]$

$n = length(y)$

$mu2 = mean(y[2 : n] * y[2 : n])$

$a = mu2$ quadratic equation coefficients

$b = (a - 1)$ quadratic equation coefficients

$c = (-2)$ quadratic equation coefficients

$res < -result(a, b, c);\ res$

$thetahat[k] = res[1];$

$for(i\, in\, 2 : n)$

$\{$

$c1[i - 1] = (1/n) * ((y[i]^2 * y[i - 1]^2))$

$\}$

$c1 = sum(c1)$

$c2 = ((thetahat[k] + 2)/(thetahat[k] * (thetahat[k] + 1)))^2$

$c3 = (thetahat[k]^2 + 4 * thetahat[k] + 2)/((thetahat[k]^2 * (thetahat[k] + 1)^2))$

$phihat[k] = (c1 - c2)/c3$

} k loop ending

$msephi = var(phihat) + (mean(phihat) - phi)^2;$

$mseth = var(thetahat) + ((mean(thetahat) - theta))^2;$

$mean(thetahat)$

$mean(phihat)$

$---$

Appendix D

R code for estimation of parameters of Lindley ACD(1,1) model

MLE of Lindley-ACD(1,1)

−−

$rm(list = ls(all = TRUE))$

$n1 = 5500;$

$n = n1 - 500;$

$y = c();$

$y1 = c();$

$psi = c();$

$theta = 1;$

$ome = 1;\ alp = 0.3;$

$bet = 0.5;$

$lam = (theta + 2)/(theta * (theta + 1));$

$est1 = rep();$

$est2 = rep();$

$est3 = rep();$

$est4 = rep();$

$L = function(par)$

$\{$

$theta = par[1];$

$w = par[2];$

$alpha = par[3];$

$beta = par[4];$

$a = n * log(theta * (theta + 2)) - 2 * n * log(theta + 1)$

$b = rep();$

$c = rep();$

$d = rep();$

$e = rep()$

$psi = rep();$

$psi[1] = 0.5$

$for(i \quad in \quad 2 : n)$

$\{$

$b[i-1] = ((theta * (theta + 1) * (w + alpha * y[i-1] + beta * psi[i-1]))$

$+(theta + 2) * y[i])$

$c[i-1] = ((theta * (theta + 1) * (w + alpha * y[i-1] + beta * psi[i-1])))$

$d[i-1] = ((theta + 2)/(theta + 1)) * (y[i]/(w + alpha * y[i-1] + beta * psi[i-1]))$

$e[i-1] = ((w + alpha * y[i-1] + beta * psi[i-1]))$

$psi[i] = w + alpha * y[i-1] + beta * psi[i-1]$

$\}$

$d1 = which(b < 0);$

$if(length(d1) == 0)\ b = b;$

$if(length(d1) > 0)\ b = b[-d1];$

$b = log(b)$

$d2 = which(c < 0);$

$if(length(d2) == 0)\ c = c;$

$if(length(d2) > 0)\ c = c[-d2];$

$c = log(c)$

$d3 = which(e < 0);$

$if(length(d3) == 0)\ e = e;$

$if(length(d3) > 0)\ e = e[-d3];$

$e = log(e)$

$b = sum(b);$

$c = sum(c);$

$d = sum(d);$

$e = sum(e);$

$l = a + b - c - d - e$

$return(-l)$

$\}$

$w = ome;$

$alpha = alp;$

$beta = bet;$

$Init = c(theta, w, alpha, beta)$

$for(j \quad in \quad 1 : 500)\{$

$eps1 = rexplindley(n1, theta, 1)$

$eps = eps1/lam;$

$psi[1] = 0.5;$

$y1[1] = 1;$

$for(i \quad in \quad 2 : n1)\{$

$psi[i] = ome + alp * y1[i-1] + bet * psi[i-1];$

$y1[i] = psi[i] * eps[i];$

$\}\ y = y1[501 : n1];$

$RES = optim(Init, L, method = "L{-}BFGS{-}B", lower = c(0.01, 0.01, 0.01, 0.01),$

$upper = c(Inf, Inf, 1, 1))$

$est1[j] = RES\mathrm{par}[1]$

est2[j]=RESpar[2]

$est3[j] = RES\mathrm{par}[3]$

est4[j]=RESpar[4]

$\}\ mean(est1);$

$seest1 < -sd(est1)/sqrt(length(est1));$

$mean(est2);$

$seest2 < -sd(est2)/sqrt(length(est2));$

$mean(est3);$

$seest3 < -sd(est3)/sqrt(length(est3));$

$mean(est4);$

$seest4 < -sd(est4)/sqrt(length(est4));$

$---$

Appendix E

R code for Bayesian estimation of parameters of IG-SCD model

$---$

$rm(list = ls(all = TRUE))$

$met4 = function(iters)$

$\{$

$xvec = numeric(iters)$

$lamcand = 2$

$for(i \quad in \quad 1 : iters)$

$\{$

$repeat\{$

$xs = lamcand + rnorm(1)$

$if(xs > 0)$

$break$

$\}$

$A = (log(1/(1 + xs^2)) - sum((log(p1[1 : n] * xs^(1/2))) - xs * p2[1 : n]))/$

$(log(1/(1 + lamcand^2)) - sum((log(p1[1 : n] * lamcand^{(1/2)}))-$

$lamcand * p2[1 : n]))$

$if(runif(1) < A)$

$lamcand = xs$

$xvec[i] = lamcand$

}

$return(xvec)$

}

$main\, program$

n1=4000

x=c(); psi=c(); phi=0.5;sigma=1;mu=1;lambda=1;u1=c();u2=c();

a=c();mua=c();lamb=c();postlamb=c();lambcand=c();la=c();

slphi=c();psi11=c(); psi1=phi;a[1]=0.5; iters=n1;

pe=c();up1=c();up2=c();phiest=c();

betphi=10;alpphi=0.5;phi1=c();phi1[1]=phi;

alpsig=1.5;betsig=0.2;sig2=c();lambout=c();

sig2[1]=sigma;

$for(s \quad in \quad 1:100)\{$

eta=rnorm(n1,0,sigma)

e=rinvgauss(n1, mu,lambda)

x[1]=phi;psi[1]=.5;

$for(i \quad in \quad 2:n1)\{$

psi[i]=phi*psi[i-1]+eta[i]

$x[i] = exp(psi[i]) * e[i]; \}$

n1=length(x)

psi1=psi[501:n1];

x1=x[501:n1];

sampling psi

n=length(x1)

for(l in 2:(n-1))

{

$a[l] = (phi/(1 + phi^2)) * (psi1[l] + psi1[l + 1])$

}

$b = sigma/(1 + phi^2);$

$for(l1 \quad in \quad 2 : (n - 1))$

{

mua[l1]=a[l1]+(b/2)

}

$psi1[1] = sqrt(lambda*exp(psi1[1])/x[1]^3)*exp(-lambda*(((x1[1]-exp(psi1[1]))^2))/$

$(2 * exp(psi1[1] * x1[1]))) * exp((-(1 - phi^2) * psi1[1])/2 * sigma)*$

$exp(-(psi1[2] - phi * psi1[1])^2/2 * sigma)$

t=2;

while(t<(n-1))

{

$l1 = exp(-lambda * (((x1[t] - exp(psi1[t]))^2))/(2 * exp(psi1[t] * x1[t])))$

u1=runif(1,0,l1)

$q1 = log(x[t]/(1+(1/sqrt(lambda))*sqrt(-2*x[t]*log(u1/sqrt(lambda))/lambda)))$

$l2 = exp(-(((psi1[t] - mua[t]))^2)/(2 * b))$

u2=runif(1,0,l2)

q2=mua[t]-(sqrt(-2*b*log(u2)))

q3=mua[t]+(sqrt(-2*b*log(u2)))

if(max(q1,q2) <q3)

$\{psi1[t] = runif(1, max(q1, q2), q3)\}$

else $\{psi1[t] = psi1[t-1]\}$

$t = t+1; \}$

psi1[n]=phi;

t1=1; slphi[t1]=phi;

$c1 = (sum(psi1[2:n]^2)/sigma) + (1/betphi^2);$

$d1 = ((sum(psi1[1:(n-1)] * psi1[2:n]))/sigma) + (alpphi/betphi^2);$

$mup = d1/c1 \quad np = n$

$while(t1 < np)\{$

$l11 = as.numeric(exp(-as.brob(((slphi[t1]) - (d1/c1))^2/(2))))$

$u21 = runif(1, 0, l11) \quad l12 = sqrt((1 - slphi[t1]^2))$

$u22 = runif(1, 0, l12)$

$pq1 = ((d1/c1) - sqrt((-2/c1) * log(u21)))$

$pq2 = ((d1/c1) + sqrt((-2/c1) * log(u21)))$

$pq3 = sqrt(1 - u22^2)$

$if(pq1 < min(pq2, pq3)) \quad slphi[t1+1] = runif(1, pq1, min(pq2, pq3))$

$elseslphi[t1+1] = slphi[t1]$

$t1 = t1 + 1;\ \}\ si1 = (alpsig + (n/2));$

$si2 = (((1 - (phi^2)) * (psi1[1])^2)/2+$

$(sum((psi1[2:n] - (phi * psi1[1:(n-1)]))^2))/2 + (betsig))$

$sig2[s] = rinvgamma(1, si1, si2)$

$p1 = (exp(psi1[1:n])/(x1[1:n]^3))^(1/2);$

$p2 = ((x1[1:n] - exp(psi1[1:n]))^2)/(2 * exp(psi1[1:n]) * x1[1:n]);$

$lambou = met4(iters)$

$lambout[s] = mean(lambou)\}$

List of Published/Communicated Papers

1. Balakrishna, N., and Sri Ranganath, C. G. (2015). Arma Models with generalized error distributed innovations. Journal of Indian Statistical Association, 53, 11-34.

2. Sri Ranganath, C. G. (2017a). Hurwicz estimator for Autoregressive model with Generalized Error Distributed Innovations. Journal of the Indian Society for Probability and Statistics. (Revised and resubmitted)

3. Sri Ranganath, C. G., and Balakrishna, N. (2017b). Bayesian Analysis of Inverse Gaussian Stochastic Conditional Duration Model(Submitted).

4. Sri Ranganath, C. G., and Balakrishna, N. (2018). Stochastic volatility and conditional durations generated by Lindley distributions. (Submitted).

References

Abraham, B., & Balakrishna, N. (1999). Inverse gaussian autoregressive models. *Journal of Time Series Analysis*, *20*(6), 605–618.

Abraham, B., Balakrishna, N., & Sivakumar, R. (2006). Gamma stochastic volatility models. *Journal of Forecasting*, *25*(3), 153–171.

Agro, G. (1995). Maximum likelihood estimation for the exponential power function parameters. *Communications in Statistics-Simulation and Computation*, *24*(2), 523–536.

Akaike, H. (1973). Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255–265.

Andel, J. (1988). On AR(1) processes with exponential white noise. *Communications in Statistics-Theory and Methods*, *17*(5), 1481–1495.

Anderson, D. N., & Arnold, B. C. (1993). Linnik distributions and processes. *Journal of Applied Probability*, *30*(2), 330–340.

Andersson, J. (2001). On the normal inverse gaussian stochastic volatility model. *Journal of Business & Economic Statistics*, *19*(1), 44–54.

Andrews, D. W. (1983). *First order autoregressive processes and strong mixing* (Tech. Rep.). Cowles Foundation for Research in Economics, Yale University.

Andrews, D. W. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica: Journal of the Econometric Society*, 139–165.

Athreya, K. B., & Pantula, S. G. (1986). A note on strong mixing of arma processes. *Statistics & probability letters*, *4*(4), 187–190.

Bakouch, H. S., & Popović, B. V. (2016). Lindley first-order autoregressive model with applications. *Communications in Statistics-Theory and Methods*, *45*(17), 4988–5006.

Balakrishna, N., & Nampoothiri, K. (2003). A cauchy autoregressive model and its applications. *Journal of the Indian Statistical Association*, *41*(2), 143–156.

Balakrishna, N., & Rahul, T. (2014). Inverse gaussian distribution for modeling conditional durations in finance. *Communications in Statistics-Simulation and Computation*, *43*(3), 476–486.

Balakrishna, N., & Shiji, K. (2014). Extreme value autoregressive model and its applications. *Journal of Statistical Theory and Practice*, *8*(3), 460–481.

Balakrishna, N., & Sri Ranganath, C. G. (2015). ARMA models with generalized error distributed innovations. *Journal of the Indian Statistical Association*, *53*(1-2), 11–34.

Bauwens, L., & Galli, F. (2009). Efficient importance sampling for ml estimation of scd models. *Computational Statistics & Data Analysis*, *53*(6), 1974–1992.

Bauwens, L., & Giot, P. (2000). The logarithmic acd model: an application to the bid-ask quote process of three nyse stocks. *Annales d'Economie et de Statistique*, *60*, 117–149.

Bauwens, L., & Lubrano, M. (1998). Bayesian inference on garch models using the gibbs sampler. *The Econometrics Journal*, *1*(1), 23–46.

Bauwens, L., & Veredas, D. (2004). The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of econometrics*, *119*(2), 381–412.

Bell, C., & Smith, E. (1986). Infrence for non-negative autoregressive schemes. *Communications in Statistics-Theory and Methods*, *15*(8), 2267–2293.

Berkoun, Y., & Douki, H. (2011). Testing independence in linear process with non-normal innovations. *Communications in Statistics-Theory and Methods*, *40*(13), 2377–2384.

Berkoun, Y., & Fellag, H. (2011). Hurwicz's estimator of the autoregressive model with non-normal innovations. *Applicationes Mathematicae*, *38*, 211–218.

Berkoun, Y., Fellag, H., & Zieliński, R. (2003). Robust testing serial correlation in AR(1) processes in the presence of a single additive outlier. *Communications in Statistics-Theory and Methods*, *32*(8), 1527–1540.

Billingsley, P. (1961). *Statistical inference for markov processes* (Vol. 2). University of Chicago Press.

Birnbaum, A. (1964). Median-unbiased estimators. *Bulletin of Mathematical Statistics*, *11*(1), 25–34.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1-2), 5–59.

Bondon, P. (2009). Estimation of autoregressive models with epsilon-skew-normal innovations. *Journal of Multivariate Analysis*, *100*(8), 1761–1776.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time series analysis: forecasting and control. *Englewood Cliffs, NJ: Prentice Hall.*.

Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, *65*(332), 1509–1526.

Box, G. E. P., & Tiao, G. C. (1962). A further look at robustness via bayes's theorem. *Biometrika*, *49*(3/4), 419–432.

Breton, A., & Pham, D. T. (1989). On the bias of the least squares estimator for the first order autoregressive process. *Annals of the Institute of Statistical Mathematics*, *41*(3), 555–563.

Campbell, J. Y., Lo, A. W.-C., MacKinlay, A. C., et al. (1997). *The econometrics of financial markets* (Vol. 2). Princeton University Press, Princeton, NJ.

Chib, S., Nardari, F., & Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, *134*(2), 341–371.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.

Denby, L., & Martin, R. D. (1979). Robust estimation of the first-order autoregressive parameter. *Journal of the American Statistical Association*, *74*(365), 140–146.

Dewald, L., & Lewis, P. (1985). A new laplace second-order autoregressive time-series model–nlar (2). *IEEE Transactions on Information theory*, *31*(5), 645–651.

Diananda, P. H. (1949). Note on some properties of maximum likelihood estimates. *Mathematical Proceedings of the Cambridge Philosophical Society*, *45*(4), 536–544.

Dufour, J.-M., & Roy, R. (1986). Generalized portmanteau statistics and tests of randomness. *Communications in Statistics-Theory and Methods*, *15*(10), 2953–2972.

Durbin, J., & Koopman, S. J. (1997). Monte carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika*, *84*(3), 669–684.

Edwards, R. G., & Sokal, A. D. (1988). Generalization of the fortuin-kasteleyn-swendsen-wang representation and monte carlo algorithm. *Physical Review D*, *38*(6), 2009.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, R. F., & Kraft, D. (1983). Multiperiod forecast error variances of inflation estimated from arch models. *Applied Time Series Analysis of Economic Data*, 293–302.

Engle, R. F., & Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, 1127–1162.

Fama, E. F. (1965). The behaviour of stock-market prices. *Journal of Business*, *38*, 34–105.

Fellag, H. (2010). Some aspects of stability in time series small sample case. *Afrika Statistika*, *5*(1), 1537–1551.

Fellag, H., & Zieliński, R. (1996). Bias of the lse estimator of the first order autoregressive model under tukey contamination. *Communications in Statistics-Theory and Methods*, *25*(7), 1537–1551.

Feng, D., Jiang, G. J., & Song, P. X.-K. (2004). Stochastic conditional duration models with leverage effect for financial transaction data. *Journal of Financial Econometrics*, *2*(3), 390–421.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(3), 350–363.

Gaver, D., & Lewis, P. (1980). First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability*, *12*(3), 727–745.

Ghitany, M., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, *78*(4), 493–506.

Gorodestkii, V. V. (1977). On the strong mixing property for linear sequences. *Teoriya Veroyatnostei i ee Primeneniya*, *22*(2), 421–423.

Grammig, J., & Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*, *3*(1), 16–38.

Guo, J.-H. (2000). Robust estimation for the coefficient of a first order autoregressive process. *Communications in Statistics-Theory and Methods*, *29*(1), 55–66.

Haddad, J. N. (2000). On robust estimation in the first order autoregressive processes. *Communications in Statistics-Theory and Methods*, *29*(1), 45–54.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hautsch, N. (2001). Modelling intraday trading activity using box-cox acd models.

Hughes, G. L., Rao, S. o., & Rao, T. S. (2007). Statistical analysis and time-series models for minimum/maximum temperatures in the antarctic peninsula. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *463*, 241–259.

Hurlimann, W. (2012). On non-gaussian AR(1) inflation modeling. *Journal of Statistical and Econometric Methods*, *1*(1), 93–101.

Hurwicz, L. (1950). Least squares bias in time series. *Statistical Inference in Dynamic Economic Models*, *10*, 365–383.

Jacquier, E., Polson, N. G., & Rossi, P. E. (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, *20*(1), 69–87.

Jayakumar, K., & Pillai, R. (1993). The first-order autoregressive mittag–leffler process. *Journal of Applied Probability*, *30*(2), 462–466.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions, vol. 2 of wiley series in probability and mathematical statistics: applied probability and statistics.* Wiley, New York,.

Kendall, M. G. (1954). Note on bias in the estimation of autocorrelation. *Biometrika*, *41*(3-4), 403–404.

Knight, J., & Ning, C. Q. (2008). Estimation of the stochastic conditional duration model via alternative methods. *The Econometrics Journal*, *11*(3), 593–616.

Lawrance, A. (1991). Directionality and reversibility in time. *International Statistical Review*, *59*(1), 67–79.

Lawrance, A., & Lewis, P. (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 165–202.

Li, W., & McLeod, A. (1988). Arma modelling with non-gaussian innovations. *Journal of Time Series Analysis*, *9*(2), 155–168.

Lindley, D. V. (1958). Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 102–107.

Lindley, D. V. (1970). *Introduction to probability and statistics from a bayesian viewpoint* (Tech. Rep.).

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303.

Lu, W., Ke, R., & Liang, J. (2016). A moment closed form estimator for the autoregressive conditional duration model. *Statistical Papers*, *57*(2), 329–344.

Luger, R. (2006). Median-unbiased estimation and exact inference methods for first-order autoregressive models with conditional heteroscedasticity of unknown form. *Journal of Time Series Analysis*, *27*(1), 119–128.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. *Unpublished paper*.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The journal of business*, *36*(4), 394–419.

Marriott, F., & Pope, J. (1954). Bias in the estimation of autocorrelations. *Biometrika*, *41*(3/4), 390–402.

McLeod, A. (1978). On the distribution of residual autocorrelations in box-jenkins models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 296–302.

Men, Z., Kolkiewicz, A. W., & Wirjanto, T. S. (2015). Bayesian analysis of asymmetric stochastic conditional duration model. *Journal of Forecasting*, *34*(1), 36–56.

Men, Z., Kolkiewicz, A. W., & Wirjanto, T. S. (2016). Bayesian inference of asymmetric stochastic conditional duration models. *Journal of Statistical Computation and Simulation*, *86*(7), 1295–1319.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, *32*(7), 685–694.

Neal, R. M. (2003). Slice sampling. *Annals of statistics*, 705–741.

Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Nelson, H. L., & Granger, C. W. J. (1979). Experience with using the box-cox transformation when forecasting economic time series. *Journal of Econometrics*, *10(1)*, 57–69.

Niehsen, W. (1999). Generalized gaussian modeling of correlated signal sources. *IEEE Transactions on Signal Processing*, *47*(1), 217–219.

Pacurar, M. (2008). Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys*, *22*(4), 711–751.

Pan, L., & Politis, D. N. (2016). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference*, *177*, 1–27.

Pitt, M. K., & Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, *94*(446), 590–599.

Pogány, T. K., & Nadarajah, S. (2010). On the characteristic function of the generalized normal distribution. *Comptes Rendus Mathématique*, *348*(3-4), 203–206.

Provost, S. B., & Sanjel, D. (2005). Inference about the first-order autoregressive coefficient. *Communications in Statistics -Theory and Methods*, *34*(5), 1183–1201.

Ramanathan, T., Mishra, A., & Abraham, B. (2016). Estimation, filtering and smoothing in the stochastic conditional duration model: an estimating function approach. *Stat*, *5*(1), 11–21.

Raqab, M. Z., Al-Jarallah, R. A., & Al-Mutairi, D. K. (2017). Closeness of lindley distribution to weibull and gamma distributions. *Communications for Statistical Applications and Methods*, *24*(2), 129–142.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Sen, P. K. (1968). Estimates of the regression coefficient based on kendall's tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389.

Shephard, N. (1996). Statistical aspects of arch and stochastic volatility. *Monographs on Statistics and Applied Probability*, *65*, 1–68.

Strickland, C. M., Forbes, C. S., & Martin, G. M. (2006). Bayesian analysis of the stochastic conditional duration model. *Computational Statistics & Data Analysis*, *50*(9), 2247–2267.

Subbotin, M. T. (1923). On the law of frequency of error. *Mathematics Collection*, *31*(2), 296–301.

Taylor, S. J. (1986). *Modelling financial time series.* John Wiley & Sons.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical finance*, *4*(2), 183–204.

Tiku, M. L., Wong, W.-K., Vaughan, D. C., & Bian, G. (2000). Time series models in non-normal situations: Symmetric innovations. *Journal of Time Series Analysis*, *21*(5), 571–596.

Tjøstheim, D., & Paulsen, J. (1983). Bias of some commonly-used time series estimates. *Biometrika*, *70*(2), 389–399.

Tong, H., Chan, K., Cox, D., Cutler, C. D., Guégan, D., Jensen, J. L., ... others (1995). A personal overview of non-linear time series analysis from a chaos perspective [with discussion and rejoinder]. *Scandinavian Journal of Statistics*, 399–445.

Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543). John Wiley & Sons, New Jersey.

Tsay, R. S. (2014). *An introduction to analysis of financial data with r.* John Wiley & Sons.

Varanasi, M. K., & Aazhang, B. (1989). Parametric generalized gaussian density estimation. *The Journal of the Acoustical Society of America*, *86*(4), 1404–1415.

Vasudeava, R., & Vasantha Kumari, J. (2013). On general error distributions. *ProbStat Forum*, *06*, 89–95.

Weiss, G. (1977). Shot noise models for the generation of synthetic streamflow data. *Water Resources Research*, *13*(1), 101–108.

Wenzel, T. (2002). Pitman-closeness as a measure to evaluate the quality of forecasts. *Communications in Statistics-Theory and Methods*, *31*(4), 535–550.

Withers, C. (1981). Conditions for linear processes to be strong-mixing. *Probability Theory and Related Fields*, *57*(4), 477–480.

Xu, D., Knight, J., & Wirjanto, T. S. (2010). Asymmetric stochastic conditional duration modela mixture-of-normal approach. *Journal of Financial Econometrics*, *9*(3), 469–488.

Yakowitz, S. J. (1973). A stochastic model for daily river flows in an arid region. *Water Resources Research*, *9*(5), 1271–1285.

Zielinski, R. (1999). A median-unbiased estimator of the ar (1) coefficient. *Journal of Time Series Analysis*, *20*(4), 477–481.