APPLICATIONS OF SOME NON-GAUSSIAN TIME SERIES IN MODELLING STOCHASTIC VOLATILITY AND CONDITIONAL DURATIONS

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

RAHUL T.

Department of Statistics
Cochin University of Science and Technology
Cochin-682022
March 2017

CERTIFICATE

Certified that the thesis entitled **"Applications of Some non-Gaussian Time Series in Modelling Stochastic Volatility and Conditional Durations "** is a bonafide record of work done by Mr. Rahul T. 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

March 2017

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

March 2017

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 **RAHUL T.**

March 2017

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 supervising guide, Dr. N. Balakrishna, Professor and Formerly Head, 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 and Prof. V. K. Ramachandran Nair former Heads, Department of Statistics, CUSAT for their extensive support. I wish to express my sincere thanks to Prof. K. C. James, Prof. Asha Gopalakrishnan, Prof. P. G. Sankaran, Head, Department of Statistics, CUSAT and Prof. S. M. Sunoj for their wholehearted support.

My most heartfelt thanks to Dr. N. Balakrishnan, Professor, McMaster University, Canada for his valuable suggestions and guidance during my research period.

I thank Prof. C. M. Latha, Prof. Uma T. P., Prof. P. K. Venugopalan of Department of Statistics, Sree Kerala Varma College, Thrissur and Sri. Ajith Kumar Raja and Smt. Renuka Raja of Sakthan Thampuran College of Mathematics and Arts, Thrissur for their love and support throughout in my life.

I remember with deep gratefulness all my former teachers.

I take this opportunity to record my sincere respect and heartiest gratitude to all the Senior Officers and my colleagues of Department of Statistics and Information Management and Internal Debt Management Department of Reserve Bank of India, Mumbai for their support. Without their cooperation, it would have been almost impossible for me to continue research work with the job.

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

I also would like to thank the Department of Science and Technology for the financial support provided during the initial phase of my research at CUSAT, under a major research project on "Modelling and Analysis of Financial Time Series".

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 am deeply indebted to my beloved father, mother and sister for their encouragement, prayers and blessings given to me. I am also indebted to my wife, Shiji K., for her constant help and encouragement throughout my research.

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

RAHUL T.

To My Loving Teachers

Contents

List of Tables **xvii**

List of Figures **xxi**

1 Introduction **1**

1.1 Motivation 1

1.2 Overview of non-Gaussian Time Series 3

1.3 Examples of Time Series 7

1.4 Some Basic Concepts 9

1.4.1 Stochastic Process 9

1.4.2 Stationary Processes 10

1.4.3 Autocorrelation and Partial Autocorrelation Function 10

1.5 Linear Time Series Models 12

1.5.1 Autoregressive Models 12

1.5.2 Moving Average Models 14

1.5.3 Autoregressive Moving Average Models 15

1.6 Box-Jenkins Modelling Techniques 16

1.6.1 Model Identification 17

1.6.2 Parameter Estimation 18

1.6.3 Diagnosis Methods 18

1.6.4 Forecasting 20

1.7 Examples for Box-Jenkins Methodology 20

1.8 Outline of the Thesis 26

2 Models for Financial Time Series 31

2.1 Introduction 31

2.2 Stylized facts of Financial Time Series 33

2.3 Models for Volatility 37

2.3.1 Autoregressive Conditional Heteroscedastic Models 37

2.3.2 Generalized ARCH (GARCH) Models 40

2.3.3 Stochastic Volatility Models 43

2.4 Models for Durations 45

2.4.1 Autoregressive Conditional Duration (ACD) Models 46

2.4.2 Stochastic Conditional Duration (SCD) Models 47

3 Birnbaum-Saunders Autoregressive Moving Average Processes 51

3.1 Introduction 51

3.2 Birnbaum-Saunders distribution 53

3.3 BS-AR(1) Model 55

3.4 BS-MA(1) Model 61

3.5 BS-ARMA(1,1) Model 64

3.6 Estimation of Parameters 68

3.6.1 BS-AR(1) Model 68

3.6.2 BS-MA(1) Model 70

3.6.3 BS-ARMA(1,1) Model 72

3.7 Asymptotic Properties of the estimators for BS-AR(1) Model 74

3.8 Simulation Study 78

3.8.1 Performance Analysis of the Models 83

3.9 Data Analysis 85

4 Modelling Stochastic Volatility using Birnbaum-Saunders Markov Sequences 91

4.1 Introduction 91

4.2 BS-SV Model and Properties 93

4.3 Estimation of Parameters 96

4.3.1 Parameter Estimation by Method of Moments 97

4.3.2 Parameter Estimation by Efficient Importance Sampling . . 98

4.4 Simulation Study 102

4.5 Data Analysis 105

5 Asymmetric Laplace Stochastic Volatility Model 111

5.1 Introduction 111

5.2 Asymmetric Laplace distribution 113

5.3 First order Asymmetric Laplace Autoregressive Process 115

5.4 Asymmetric Laplace SV Model 118

5.5 Parameter Estimation 121

5.6 Asymptotic Properties of Estimators 123

5.7 Simulation Study 129

5.8 Data Analysis 131

6 Inverse Gaussian distribution for Modelling Conditional Durations in Finance 137

6.1 Introduction 137

6.2 Review of ACD Models 139

6.3 Inverse Gaussian ACD Model 146
6.3.1 Properties of IG-ACD Model 147
6.4 Estimation of IG-ACD Model 149
6.5 Extended Generalized Inverse Gaussian ACD Model 151
6.5.1 Special Cases 153
6.6 Inverse Gaussian SCD Model and Properties 154
6.7 Estimation of IG-SCD Model 157
6.8 Simulation Study 163
6.8.1 IG-ACD Model 163
6.8.2 IG-SCD Model 163
6.9 Data Analysis 167

7 Conclusions and Future Research 175

Appendix A Matlab code for estimation of parameters of BS-ARMA Model 181

Appendix B R code for computation of I_1 191

Appendix C R code for estimation of parameters of BS-SV Model 193

Appendix D R code for estimation of parameters of AL-SV Model 211

Appendix E R code for estimation of parameters of IG duration models 217

Bibliography 239

List of Tables

1.1 AIC and BIC of fitted models for crude oil price data 22

1.2 AIC and BIC for fitted models for rice production data 26

2.1 Summary statistics for BSE log-returns 35

3.1 The average estimates and the corresponding mean square error for the MLEs of BS-AR(1), when $\alpha = 2$, $\beta = 1$ and for different ρ's . . 79

3.2 The average estimates and the corresponding mean square error for the MLEs of BS-MA(1), when $\alpha = 2$, $\beta = 1$ and for different θ's . . 80

3.3 The average estimates and the corresponding mean square error for the MLEs of BS-ARMA(1,1), when $\alpha = 2$, $\beta = 1$ and for different (ρ, θ) values. 81

3.4 Estimated coverage probabilities $(CP(.))$ for selected values of $\alpha = 2, \beta = 1$ and different values of ρ 82

3.5 Forecast evaluation statistics for simulated data 84

3.6 Summary statistics for transformed data 85

3.7 Parameter Estimates using ML methods 87

3.8 Model evaluation statistics 90

4.1 The average estimates and the corresponding mean square error for the MMEs, when α=2, β=1 and for different ρ's. 103

4.2 The average estimates and the corresponding mean square error for the EIS-MLEs, when α=2, β=1 and for different ρ's. 104

4.3 Descriptive statistics of the return series 106

4.4 Estimates of parameters and Ljung-Box statistic for residuals 108

5.1 The average estimates and the corresponding mean square error of moment estimates based on sample of size n=1000, when κ=2, θ=1 and for different values of ρ and σ 130

5.2 The average estimates and the corresponding mean square error of moment estimates based on sample of size n=3000, when κ=2, θ=1 and for different values of ρ and σ 131

5.3 Summary statistics of return series 133

5.4 Parameter estimates using method of moments 133

5.5 Ljung-Box Statistic for the residuals and squared residuals 135

6.1 The average ML estimates and the corresponding mean square error for IG-ACD model 164

6.2 The average estimates and the corresponding mean square error for the EIS ML estimates 165

6.3 Descriptive statistics for the data 170

6.4 Parameter estimates for the data 170

6.5 Ljung-Box Statistics for standardized residual series and its squared process with lags 10 and 20. 173

List of Figures

1.1 Monthly crude oil price from April 2000 to March 2015 7

1.2 Annual rice production in India from 1950-51 to 2014-15 8

1.3 Stationary series of monthly crude oil price from April 2000 to March 2015 21

1.4 ACF and PACF plot for monthly crude oil price 22

1.5 Plot of residuals and ACF of residuals for monthly crude oil price . 23

1.6 Stationary series of annual rice production in India from 1950-51 to 2014-15 24

1.7 ACF and PACF plot for annual rice production 24

1.8 Plot of residuals and ACF of residuals for annual rice production . . 25

2.1 Time series plot of BSE index and returns 34

2.2 ACF of returns and squared returns of BSE index 36

2.3 Histogram of BSE index return and normal approximation 36

3.1 ACF of the Gaussian AR(1) and the corresponding BS Markov sequence

3.2 ACF and PACF of the Gaussian MA(1) and the corresponding BS-MA sequence. . . . 64

3.3 ACF and PACF of the Gaussian ARMA(1,1) and the corresponding BS-ARMA sequence 67

3.4 Time series plot of the actual series and adjusted series 86

3.5 ACF and PACF of seasonally adjusted and de-trended series 87

3.6 Histograms of the data with superimposed BS density 88

3.7 Time series and ACF plots of the residual series from fitted BS and Gaussian models 89

4.1 The plot of kurtosis of return and the ACF of squared return 95

4.2 Time series plot the data and the return 106

4.3 ACF of the returns and the squared returns 107

4.4 ACF of the residuals 109

4.5 Histogram of residuals with superimposed standard normal density 110

5.1 The plot of kurtosis of r_t 119

5.2 The ACF of squared returns for different combinations of parameters 121

5.3 Time series plot of the original data and the returns 132

5.4 ACF of the returns and the squared returns 134

5.5 ACF of the residuals 135

6.1 Hazard rate of inverse Gaussian distribution when $\mu = 1$. 149

6.2 Unconditional hazard functions of IG-SCD model for simulated data and its empirical hazard rate. 167

6.3 Time series plot of adjusted durations. 169

6.4 Time plot of standardized innovation series of IG-ACD(1,1) and IG-SCD models. 171

6.5 ACF of the standardized residual series of IG-ACD(1,1) and IG-SCD models. 172

6.6 Histogram of standardized residuals superimposed by inverse Gaussian density for IG-ACD model and IG-SCD model. 173

6.7 Empirical and estimated hazard functios of adjusted durations of exchange rates data. 174

Chapter 1

Introduction

1.1 Motivation

In an effort to understand the changing world around us, observations of one kind or another are frequently made sequentially over time. For example, the daily maximum temperature is increasing every year, the price of gold fluctuates day by day, the index of Bombay Stock Exchange fluctuates every now and then, etc. A record of such observations made sequentially in time is referred to as *time series*. Systematic studies of such time series help us to uncover the dynamical law governing its generation. However, a complete uncovering of the law may not be possible in practice as only partial observations are available in most of the cases. The major objectives of time series analysis are: (1) Understanding of the dynamic structure of data generating mechanism. (2) Construction of empirical time series models incorporating as much available background theory as possible, (3) Check if the model captures the important features of the observed data, (4) Predicting the future behaviour of the series.

The data are generated either by controlled experiments or by the nature. In either case the observations are subject to random errors, and they may be fluctuating around a constant or time-varying level. One can view such data as a realization of a more general stochastic process. That is, an observed time series can be viewed as a realization of a discrete parameter stochastic process. To understand the data generating mechanism, one has to use appropriate stochastic models, which link the observations at different time points.

The analysis of time series in the classical set up, assumes that the series is a realization of some Gaussian process and the value at a time point t is a linear function of past observations. Linearity is the basic assumption in the theory and methods of classical time series analysis developed by Box and Jenkins (1970). This is widely preferred, since most parameter estimation techniques can lead to analytically tractable solutions under this assumption. Moreover this Gaussian assumption has been based on the central limit theorem and is valid for processes having finite variance. Therefore processes having infinite variances cannot be modelled as Gaussian. The studies on financial and econometric time series have established these facts. The study of non-Gaussian time series is motivated mainly by two aspects. First is that it gets stationary sequences having non-normal marginal random variables; second is to study the point processes generated by sequences of non-negative dependent random variables. This includes the counting processes generated when the sequence of times is Markovian, such as first order Autoregressive (AR(1)) sequence (cf, Gaver and Lewis (1980)). In view of this, a large number of non-linear and non-Gaussian time series models are introduces in the literature, see Tong (1995). Our objective in this thesis is to study various aspects of financial time series and develop some non-Gaussian time series to model non-negative variables like volatility, durations, price etc. in finance.

1.2 Overview of non-Gaussian Time Series

The theory and methods for analysing time series in classical set-up is based on the assumption that such series are realizations of linear Gaussian processes, cf. Box and Jenkins (1976). However, we come across practical situations in which the observed series are generated by non-Gaussian processes. In modelling such non-Gaussian time series, the usual practice is to make suitable transformations to remove skewness in the data and then fit a Gaussian model. But there are cases where the assumption that the transformed data follows Gaussian distribution is unlikely to be true (cf. Lawrance (1991)). For this reason, a number of non-Gaussian time series models have been introduced in the literature during the last four decades. For example, Lawrance and Kottegoda (1977) explain the need for using time series models having non-Gaussian marginal distributions for modelling river flow and other hydrological time series data. In economic studies, Nelson and Granger (1979) considered a set of 21 time series data of which only six were found to be Gaussian. Time series models with Weibull marginal distribution for wind velocity (Brown et al. (1984)), Laplace marginal distribution for image source modelling (Gibson (1983)), Linnik marginal distribution for stock price return (Anderson and Arnold (1993)) are some other examples.

In the development of non-Gaussian time series models, it is observed that the method of analysis depends on the type of marginal distribution. When we insist a specific stationary marginal distribution for a model, its innovation distribution takes a different form unlike in the Gaussian case. In particular, if we restrict the variables to be non-negative, then for most of the standard distributions, the innovation random variable does not have a closed-form expression for its density, which poses difficulties in the associated likelihood inference. For example, Gaver and Lewis (1980) introduced an autoregressive model of order one (AR(1)) with gamma marginal distribution to study the properties of point processes generated by such sequences. Lawrance and Lewis, in a series of papers, discussed several autoregressive moving average (ARMA) sequences with exponential and gamma marginals; see Lawrance and Lewis (1985) and the references contained therein. Properties of other Markov sequences with non-Gaussian marginals such as gamma (Sim (1990), Adke and Balakrishna (1992)), inverse Gaussian (Abraham and Balakrishna (1999)), Cauchy (Balakrishna and Nampoothiri (2003)), normal-Laplace (Jose et al. (2008)), approximated beta distribution (Popović (2010), Popović et al. (2010)), extreme value (Balakrishna and Shiji (2014a)) have also been discussed in the literature.

The modelling of non-negative random variables play a major role in the study of financial time series, where one has to model the evolution of conditional variances known as stochastic volatility (see Tsay (2005)). During the last two decades, there has been an increasing interest in modelling the dynamic evolution of the volatility of high-frequency series of financial returns. The stochastic volatility(SV) models have been widely used to model a changing variance of financial time series data.

These models usually assume Gaussian distribution (Jacquier et al. (1994); Kim et al. (1998)) for asset returns conditional on the latent volatility. However, empirical studies show that the volatility of asset returns are not constant and the returns are more peaked around the mean and have fatter tails than those implied by the normal distributions. These empirical observations have led to the models in which the volatility of returns follows non-Gaussian distributions. To account for heavy tails observed in returns series, Harvey et al. (1994), Liesenfeld and Jung (2000), Chib et al. (2002), Berg et al. (2004), Jacquier et al. (2004), Omori et al. (2007), Asai (2008), Choy et al. (2008), Nakajima and Omori (2009), Asai and McAleer (2011), Wang et al. (2011), Nakajima and Omori (2012) and Delatola and Griffin (2013) assume that the conditional distribution of returns follow Student's t-distribution. The other studies used the Normal Inverse Gaussian distribution (see Barndorff-Nielsen (1997) and Andersson (2001)), the Generalized Error Distribution (see Liesenfeld and Jung (2000)), and the Generalized-t distribution (see Wang (2012) and Wang et al. (2013)) for incorporating the leptokurtic nature of conditional distribution of returns. Bauwens et al. (2012) give a detailed discussion on SV models with various distributional assumptions to account for non-normality of data and time varying volatility simultaneously.

A more straight forward way is to use an AR(1) model for non-negative random variables to generate the volatility sequence. The standard SV model in the literature assumes a Gaussian AR(1) model for generating the log-volatility sequence. As an alternative to this normal-lognormal SV models, Abraham et al. (2006) proposed a SV model in which the volatility sequence is generated by a gamma AR(1) sequence of Gaver and Lewis (1980) and Balakrishna and Shiji (2014b) developed a SV model generated by first order Gumbel extreme value autoregressive process.

A relatively new area of research in the context of financial time series is the modelling and analysis of time duration between consecutive events. The duration between transactions in finance is important, for it may signal the arrival of new information concerning the underlying asset. Engle and Russell (1998) use an idea similar to that of the generalized autoregressive conditional heteroscedastic models to propose an autoregressive conditional duration (ACD) model and show that the model can successfully describe the evolution of time durations for (heavily traded) stocks. A feature of Engle and Russell's linear ACD specification with exponential or Weibull errors is that the implied conditional hazard functions are restricted to be either constant or increasing/decreasing. Zhang et al. (2001), Hamilton and Jorda (2002) and Bauwens and Veredas (2004) questioned whether this assumption is an adequate one. As an alternative to the Weibull distribution used in the original ACD model, Lunde (1999) employs a formulation based on the generalized Gamma distribution, while Grammig and Maurer (2000) and Hautsch (2001) utilize the Burr and generalized F distributions respectively. Recently, Bauwens and Veredas (2004) proposed the stochastic conditional duration model (SCD), in which the evolution of the durations is assumed to be driven by a latent factor. The motivation for the use of the latent variable is that it captures general unobservable information on the market. A recent review of the literature on the ACD models and their applications in finance can be found in Pacurar (2008).

The contents of this thesis are on various aspects of modelling and analysis of non-Gaussian and non-negative time series in view of their applications in finance to model stochastic volatility and conditional durations.

1.3 Examples of Time Series

A time series is an ordered sequence of observations. Time series analysis deals with statistical methods for analysing and modelling such ordered sequence of observations. Although the ordering is usually through time, particularly in terms of some equally spaced time intervals, the ordering may also be taken through other dimensions, such as space. Time series occur in a variety of fields such as agriculture, business, finance, economics, engineering, medical studies etc. In this section, we describe some examples of time series.

Before going into more formal analysis, it is useful to examine some real time series data by plotting them against time. The first example is the monthly crude oil price in dollars per barrel in Indian market, one of the widely discussed time series. The data consists of 180 observations from April 2000 to March 2015. The time series plot of the data is shown in Figure 1.1. It is obvious from the figure that the series

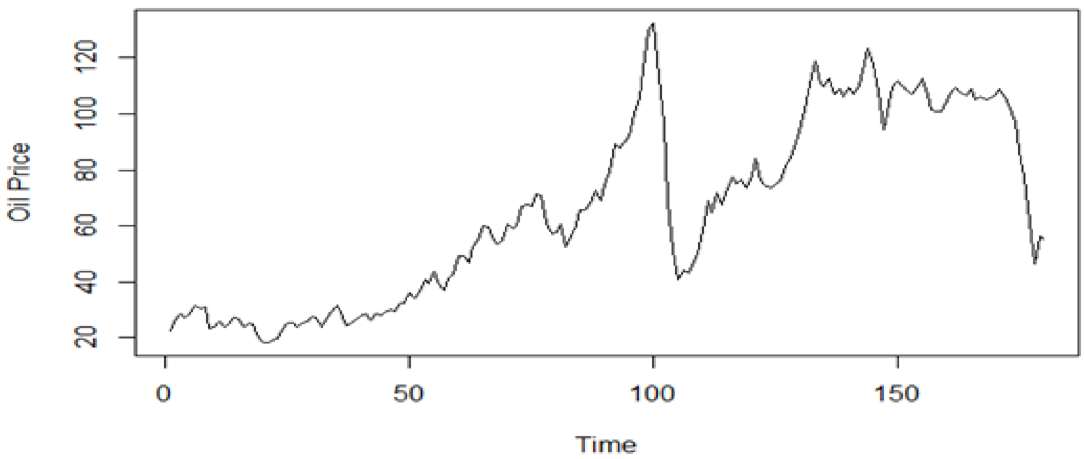

FIGURE 1.1: Monthly crude oil price from April 2000 to March 2015

is non-stationary because its mean is not constant through time. This is a typical economic series where time series analysis could be used to formulate a model for forecasting future values of the oil price.

Next, we consider annual rice production (in Million Tonnes) in India from 1950-51 to 2014-15. The data on rice production were obtained from Ministry of Agriculture,

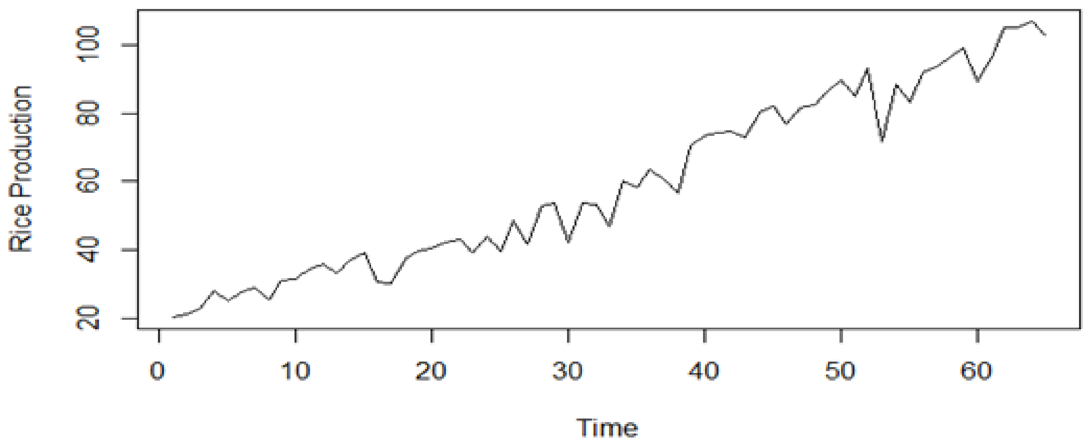

FIGURE 1.2: Annual rice production in India from 1950-51 to 2014-15

Government of India. From the Figure 1.2 it is apparent that the data exhibit a clear positive trend. A proper trend analysis and forecast of production of such an important crop is significant to stabilize the price and ensure profits for the farmers.

Other examples include (1) Monthly index of industrial production, (2) the maximum temperature at a particular location on successive days, (3) electricity consumption in a particular area for successive one-hour periods, (4) daily exchange rate of a domestic currency with foreign currency, (5) weekly interest rates, and (6) monthly price indices, etc.

In the upcoming sections, we list some of the basic concepts which facilitate the systematic development of the thesis.

1.4 Some Basic Concepts

We begin with basic definition of stochastic processes, stationary process, the autocorrelation and partial autocorrelation functions etc. that are necessary for proper understanding of time series models. We also give a simple introduction to linear time series models and Box-Jenkins modelling techniques, which play a fundamental role in time series analysis.

1.4.1 Stochastic Process

A stochastic process is a family of time indexed random variables $X(\omega, t)$, where ω belongs to a sample space and t belongs to an index set. For a given ω, $X(\omega, t)$, as a function of t, is called a sample function or realization. The population that consists of all possible realizations is called the ensemble in stochastic processes. Thus, a time series is a realization or a sample function from a certain stochastic process. With proper understanding that a stochastic process, $X(\omega, t)$, is a set of time indexed random variables defined on a sample space, we usually suppress the variable ω and simply write $X(\omega, t)$ as $X(t)$ or X_t. The mean function and variance function of the process are defined as $\mu_t = E(X_t)$ and $\sigma_t^2 = Var(X_t) = E(X_t - \mu_t)^2$.

1.4.2 Stationary Processes

A time series $\{X_t\}$ is said to be strictly stationary if the joint distribution of $(X_{t_1}, X_{t_2}, ..., X_{t_n})$ is identical to that of $(X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k})$ for all t and k, where n is an arbitrary positive integer and $(t_1, t_2, ..., t_n)$ is a collection of n integers. In other words, strict stationarity requires that the joint distribution of $(X_{t_1}, X_{t_2}, ..., X_{t_n})$ is invariant under time shift. This is very strong condition that is hard to verify empirically. A weaker version of stationarity is often assumed.

A time series $\{X_t\}$ is said to be weakly stationary if

(i) $E(X_t) = \mu$, a constant,

(ii) $Var(X_t) < \infty$,

(iii) $Cov(X_t, X_s)$ is a function of $|t - s|$ only.

From the above definitions, it is clear that, if $\{X_t\}$ is strictly stationary and its first two moments are finite, then $\{X_t\}$ is also weakly stationary. The converse is not true in general. However, a Gaussian process is weakly stationary if and only if it is strictly stationary.

1.4.3 Autocorrelation and Partial Autocorrelation Function

Let $\{X_t : t = 0, \pm 1, \pm 2, ...\}$ be a stochastic process, the covariance between X_t and X_{t-k} is known as the autocovariance function at lag k and is defined by

$$\text{Cov}\,(X_t,\, X_{t-k}) = E(X_t - E(X_t))(X_{t-k} - E(X_{t-k})).$$

Hence, the correlation coefficient between X_t and X_{t-k}, is called Autocorrelation function (ACF) at lag k, and is given by

$$\rho_X(k) = \text{Corr}\,(X_t,\, X_{t-k}) = \frac{\text{Cov}(X_t,\, X_{t-k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t-k})}}, \tag{1.1}$$

where $Var(.)$ is the variance function of the process.

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

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

$$\phi_{1,\,1} = \text{Corr}(X_1,\, X_0) = \rho_1$$

and

$$\phi_{k,\,k} = \text{Corr}(X_k - \hat{X}_k,\, X_0 - \hat{X}_0), \quad k \geq 2,$$

where $\hat{X}_k = l_1\, X_{k-1} + l_2\, X_{k-2} + \cdots + l_{k-1}\, X_1$ is the linear predictor. Both $(X_k, \hat{X}_k)$ and $(X_0, \hat{X}_0)$ are correlated with $\{X_1,\, X_2, ..., X_{k-1}\}$. By stationarity, the PACF, is the correlation between X_t and X_{t-k} obtained by fixing the effect of $X_{t-1},\, X_{t-2}, ..., X_{t-(k-1)}$.

1.5 Linear Time Series Models

The most popular class of linear time series models are autoregressive moving average (ARMA) models, including purely autoregressive (AR) and purely moving average (MA) models as special cases. ARMA models are frequently used to model linear dynamic structures, to depict linear relationships among lagged variables, and to serve as vehicles for linear forecasting. This section gives a brief overview of linear time series models.

1.5.1 Autoregressive Models

A stochastic model that can be extremely useful in the representation of certain practically occurring series is the autoregressive model. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock η_t. Let us denote the values of a process at equally spaced time $t,\ t-1,\ t-2,\ ...$ by $X_t,\ X_{t-1},\ X_{t-2}, ...$, then X_t can be described by the following expression:

$$X_t = \rho_1 X_{t-1} + \rho_2 X_{t-2} + ... + \rho_p X_{t-p} + \eta_t. \tag{1.2}$$

Or equivalently $\varphi(B)X_t = \eta_t$ with $\varphi(B) = 1 - \rho_1 B - \rho_2 B^2 - \cdots - \rho_p B^p$, where B is the back shift operator, defined by $BX_t = X_{t-1}$, $\{\eta_t\}$ is a sequence of uncorrelated random variables with mean zero and constant variance, termed as innovations and $\varphi(B)$ is referred to as the characteristic polynomial associated with an AR(p) process. As X_t is a linear function of its own past p values, the process $\{X_t\}$ is referred to as an Autoregressive process of order p (AR(p)). This is rather like a multiple regression model, but X_t is regressed not on independent variables but on past values of X_t; hence the prefix ‘auto’. The resulting AR(p) process is weakly stationary if all the roots of the associated characteristic polynomial equation $\varphi(B) = 0$ lie outside the unit circle.

For a stationary AR(p) processes, the autocorrelation function, $\rho_X(k)$, can be found by solving a set of difference equations called the Yule-Walker equations given by

$$(1 - \rho_1 B - \rho_2 B^2 - \cdots - \rho_p B^p)\rho_X(k) = 0\,, \quad k > 0.$$

The plot of ACF of a stationary AR(p) model would then show a mixture of damping sine and cosine patterns and exponential decays depending on the nature of its characteristic roots.

The autoregressive model of order 1 (AR(1)) is important as it has several useful features. It is defined by

$$X_t = \rho X_{t-1} + \eta_t, \tag{1.3}$$

where $\{\eta_t\}$ is a white noise with mean 0 and variance σ^2. The sequence $\{X_t\}$ is weakly stationary AR(1) process when $|\rho| < 1$. Under stationarity, we have $E(X_t) = 0$, $Var(X_t) = \sigma^2/(1 - \rho^2)$ and the autocorrelation function is given by

$$\rho_X(k) = \rho^k, \;\; k = 0, 1, 2, ...$$

This result says that the ACF of a weakly stationary AR(1) series decays exponentially in k. If we assume that the innovation sequence $\{\eta_t\}$ is independent and identically distributed then the AR(1) sequence is Markovian.

1.5.2 Moving Average Models

Another type of model of great practical importance in the representation of observed time series is the finite moving average process. In this model, the observation X_t at time t is expressed as a linear function of the present and past shocks. A moving average model of order q (MA(q)) is defined by

$$X_t = \eta_t - \theta_1 \, \eta_{t-1} - \theta_2 \, \eta_{t-2} - \ldots - \theta_q \, \eta_{t-q}. \tag{1.4}$$

Or, $X_t = \Theta(B)\eta_t$, 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 θ_i's are constants, $\{\eta_t\}$ is a white noise sequence.

The definition implies that

$$E(X_t) = 0; Var(X_t) = \sigma^2 \sum_{i=1}^{q} \theta_i^2$$

and the ACF is,

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

Hence, for an MA(q) model, its ACF vanishes after lag q.

In particular an MA(1) model for $\{X_t\}$ is defined by

$$X_t = \eta_t - \theta \, \eta_{t-1}.$$

So, X_t is a linear function of the present and immediately preceding shocks. The MA(q) process will always be stationary as it is a finite linear combination of shocks, but it is invertible if $|\theta| < 1$. The unconditional variance is given by $Var(X_t) = (1 + \theta^2)\,\sigma^2$.

The ACF of the MA(1) process is

$$\rho_X(k) = \begin{cases} -\theta/(1+\theta^2), & k = 1 \\ 0, & k = 2, 3, ... \end{cases} .$$

1.5.3 Autoregressive Moving Average Models

A natural extension of the pure autoregressive and pure moving average processes is the mixed autoregressive moving average process. An ARMA model with p AR terms and q MA terms is called an ARMA (p, q) model. The advantage of ARMA process relative to AR and MA processes is that it gives rise to a more parsimonious model with relatively few unknown parameters.

A mixed process of considerable practical importance is the first order autoregressive moving average (ARMA(1, 1)) model.

$$X_t - \rho X_{t-1} = \eta_t - \theta\, \eta_{t-1}. \tag{1.6}$$

The process is stationary if $|\rho| < 1$ and invertible if $|\theta| < 1$. The mean, variance and the autocorrelation function of the ARMA(1, 1) model are respectively given by

$$E(X_t) = 0, \;\; Var(X_t) = E(X_t^2)$$

and the ACF is

$$\rho_X(k) = \begin{cases} \frac{\rho\theta^2 - \theta\,\rho^2 + \rho - \theta}{1+\theta^2 - 2\theta\,\rho}, & if \; k = 1 \\ \rho \;.\rho_{k-1}, & if \quad k = 2, 3, ... \end{cases} . \tag{1.7}$$

Thus the autocorrelation function decays exponentially from the starting value ρ_1, which depends on θ as well as on ρ.

A more general model that encompasses AR(p) and MA(q) model is the autoregressive moving average, or ARMA(p, q), model

$$X_t - \rho_1 X_{t-1} - \rho_2 X_{t-2} - ... - \rho_p X_{t-p} = \eta_t - \theta_1\,\eta_{t-1} - \theta_2\,\eta_{t-2} - ... - \theta_q\,\eta_{t-q}. \tag{1.8}$$

The model is stationary if AR(p) component is stationary and invertible if MA(q) component is so. One may refer Box et al. (1994) for detailed analysis of linear time series models.

1.6 Box-Jenkins Modelling Techniques

This section examines the Box-Jenkins methodology for model building and discusses its possible contribution to post-sample forecasting accuracy. A three step procedure is used to build a model. First a tentative model is identified through analysis of historical data. Second, the unknown parameters of the model are estimated. Third, through residual analysis, diagnostic checks are performed to determine the adequacy of the model. We shall now discuss each of these steps in more detail.

1.6.1 Model Identification

At this stage of time series modelling, the analysis intends to suggest a tentative model to a time series by examining the time plot and the graphical representation of each of the autocorrelation function and partial autocorrelation function. Such plots could reveal certain properties of a time series like non-stationarity and outlier. The sample correlogram and partial correlogram help us to determine the order of the model. Autocorrelation function of an autoregressive process of order p tail off and its partial autocorrelation function has a cut off after lag p. On the other hand, the autocorrelation function of moving average process cuts off after lag q, while its partial autocorrelation tails off after lag q. If both autocorrelation and partial autocorrelation tail off, a mixed process is suggested. Furthermore, the autocorrelation function for a mixed process, contains a p-th order AR component and q-th order moving average component, and is a mixture of exponential and damped sine waves after the first $q-p$ lags. The PACF for a mixed process is dominated by a mixture of exponential and damped sine waves after the first $q-p$ lags.

1.6.2 Parameter Estimation

Estimating the model parameters is an important aspect of time series analysis. There are several methods available in the literature for estimating the parameters, (see Box et al. (1994)). All of them should produce very similar estimates, but may be more or less efficient for any given model. The main approaches to fitting Box–Jenkins models are non-linear least squares and maximum likelihood estimation. The least squares estimator (LSE) of the parameter is obtained by minimizing the sum of the squared residuals. For pure AR models, the LSE leads to the linear Ordinary Least Squares (OLS) estimator. If moving average components are present, the LSE becomes non-linear and has to be solved by numerical methods. The maximum likelihood (ML) estimator maximizes the (exact or approximate) log-likelihood function associated with the specified model. To do so, explicit distributional assumption for the innovations has to be made. Other methods for estimating model parameters are the method of moments (MM) and the generalized method of moments (GMM), which are easy to compute but not very efficient.

1.6.3 Diagnosis Methods

After estimating the parameters one has to test the model adequacy by checking the validity of the assumptions imposed on the errors. This is the stage of diagnosis check. Model diagnostic checking involves techniques like over fitting, residual plots, and more importantly, checking that the residuals are approximately uncorrelated. This makes good modelling sense, since in the time series analysis a good model should be able to describe the dependence structure of the data adequately, and one important measure of dependence is the autocorrelation function. In other words, a good time series model should be able to produce residuals that are approximately uncorrelated, that is, residuals that are approximately white noise. Note that as in the classical regression case complete independence among the residuals is impossible because of the estimation process. However, the autocorrelations of the residuals should be close to being uncorrelated after taking into account the effect of estimation. As shown in the seminal paper by Box and Pierce (1970), the asymptotic distribution of the residual autocorrelations plays a central role in checking out this feature. From the asymptotic distribution of the residual autocorrelations we can also derive tests for the individual residual autocorrelations and overall tests for an entire group of residual autocorrelations assuming that the model is adequate. These overall tests are often called portmanteau tests, reflecting perhaps that they are in the tradition of the classical chi-square tests of Pearson. Nevertheless, portmanteau tests remain useful as an overall benchmark assuming the same kind of role as the classical chi-square tests. It can also be seen that like the classical chi-square tests, portmanteau tests or their variants can be derived under a variety of situations. Portmanteau tests and the residual autocorrelations are easy to compute and the rationale of using them is easy to understand. These considerations enhance their usefulness in applications.

Model diagnostic checks are often used together with model selection criteria such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). These two approaches actually complement each other. Model diagnostic checks can often suggest directions to improve the existing model while information criteria can be used in a more or less "automatic" way within the same family of models. Through the exposition on diagnostic checking methods, it is hoped that the practitioner should be able to grasp the relative merits of these models and how these different models can be estimated.

1.6.4 Forecasting

One of the objectives of analysing time series is to forecast its future behaviour. That is, based on the observations up to time t, we should be able to predict the value of the variable at a future time point. The method of Minimum Mean Square Error (MMSE) forecasting is widely used when the time series follows a linear model. In this case an l-step ahead forecast at time t becomes the conditional expectation, $E(X_{t+l}|X_t, X_{t-1}, ...)$. In the present study of financial time series, our goal is to forecast the volatility and we have to deal with non-linear models. Hence different approaches are adopted for different models and we will describe them as and when we need such methods.

1.7 Examples for Box-Jenkins Methodology

Example 1: This section illustrates the concepts and ideas just presented by working out a couple of examples. First, we take monthly crude oil price data as an example. The data are plotted in Figure 1.1. In this case, the difference of order one is sufficient to achieve stationarity in mean. The first differenced data are plotted in Figure 1.3.

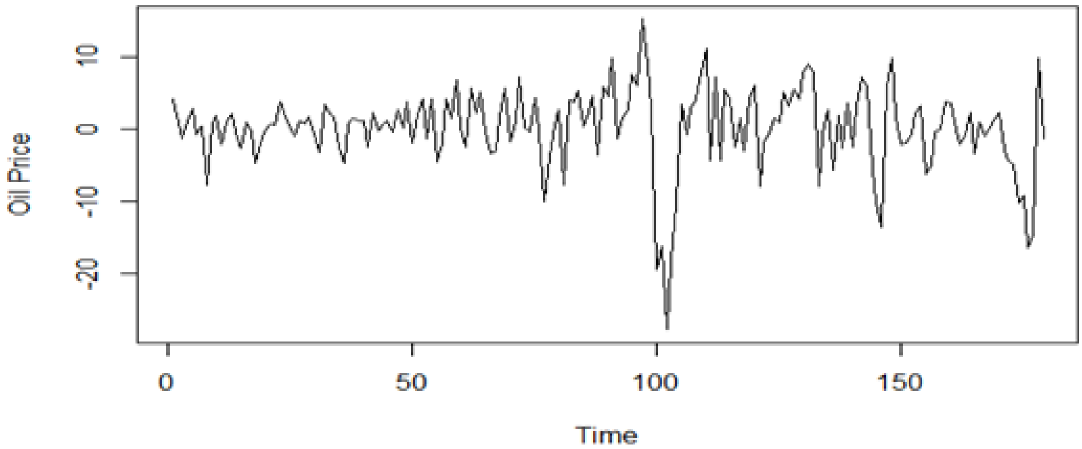

FIGURE 1.3: Stationary series of monthly crude oil price from April 2000 to March 2015

First we analyse the ACF and the PACF. They are plotted together with corresponding confidence intervals in Figure 1.4. The exponentially decaying ACF suggests an AR model. As the sample PACF has a single significant spike at lag 1 indicates that the series is likely to be generated from an AR(1) process.

The least squares fit of this model is:

$$X_t = \underset{(0.0673)}{0.4410}\, X_{t-1} + \hat{\eta}_t.$$

Next, we investigate the information criteria AIC and BIC to identify the orders of the ARMA(p,q) model. We examine all models with $0 \leq p, q \leq 4$. The AIC and the BIC values are reported in Table 1.1. Both criteria reach a minimum at $(p, q) = (1, 0)$ (bold numbers) so that both criteria suggest an AR(1) model.

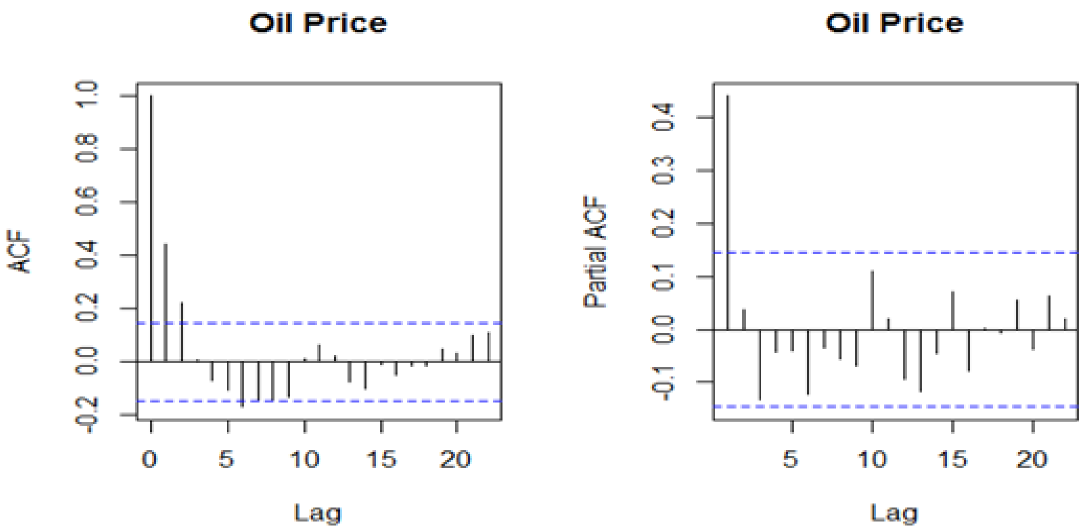

FIGURE 1.4: ACF and PACF plot for monthly crude oil price

Order (p, q)	0	1	2	3	4
0	*AIC*	6.1444	6.1153	6.1061	6.1173
	BIC	6.1622	6.1309	6.1495	6.1785
1	**6.0961**	6.1165	6.1297	6.1206	6.1365
	6.1140	6.1423	6.1533	6.1821	6.1959
2	6.1118	6.1276	6.1257	6.1360	6.1216
	6.1476	6.1714	6.1875	6.1957	6.2093
3	6.1282	6.1299	6.1194	6.1205	6.1251
	6.1622	6.1719	6.1995	6.2286	6.2412
4	6.1227	6.1247	6.1287	6.1319	6.1255
	6.1950	6.2051	6.2172	6.2085	6.2602

TABLE 1.1: AIC and BIC of fitted models for crude oil price data

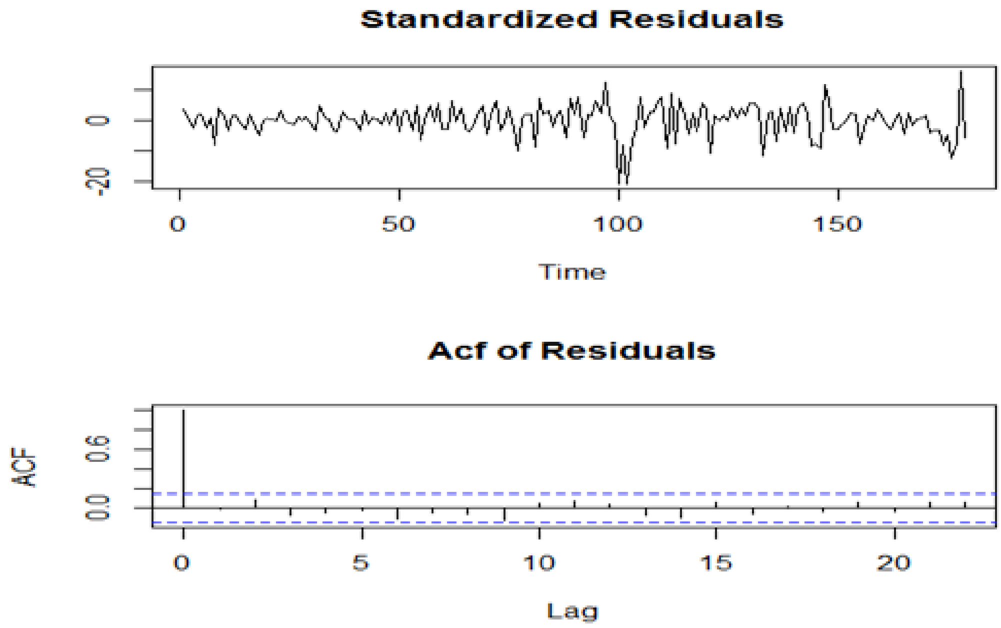

FIGURE 1.5: Plot of residuals and ACF of residuals for monthly crude oil price

The standardized residuals and ACF of residuals are plotted in Figure 1.5. They show no sign of significant autocorrelations so that residual series are practically white noise. We can examine this hypothesis formally by Ljung-Box test. The Ljung-Box statistic for residual series is obtained as 0.2377 which is less than the 5% chi-square critical value 10.117 at degrees of freedom 20. Hence we conclude that there is no significant dependence among the residuals. Thus the model seems adequate for the data.

Example 2: Consider annual rice production data from Section 1.3. A time series plot of the data is given in Figure 1.2. The process shows signs of non-stationarity with changing mean. The series was transformed by taking the first difference of natural logarithm of values to attain stationarity. The time series plot of the

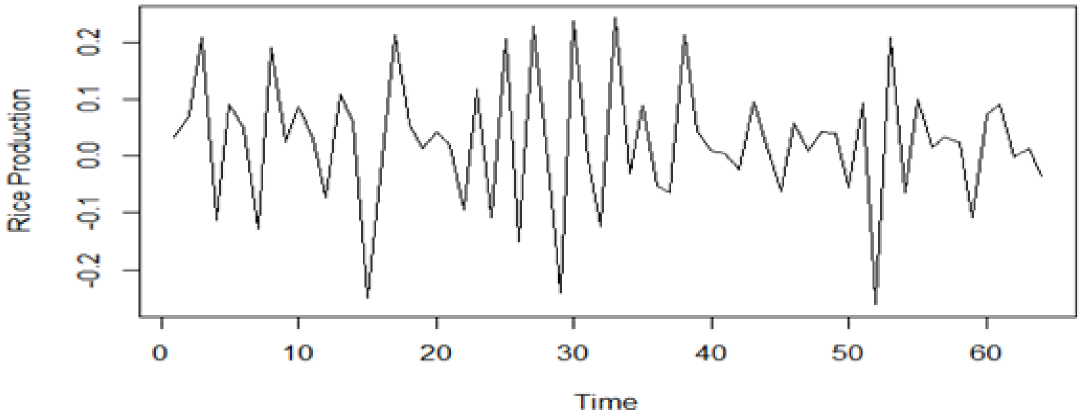

Figure 1.6: Stationary series of annual rice production in India from 1950-51 to 2014-15

transformed series is presented in Figure 1.6.

The plot of ACF and PACF are given in Figure 1.7. The ACF and PACF suggest a MA(1) model for the transformed series.

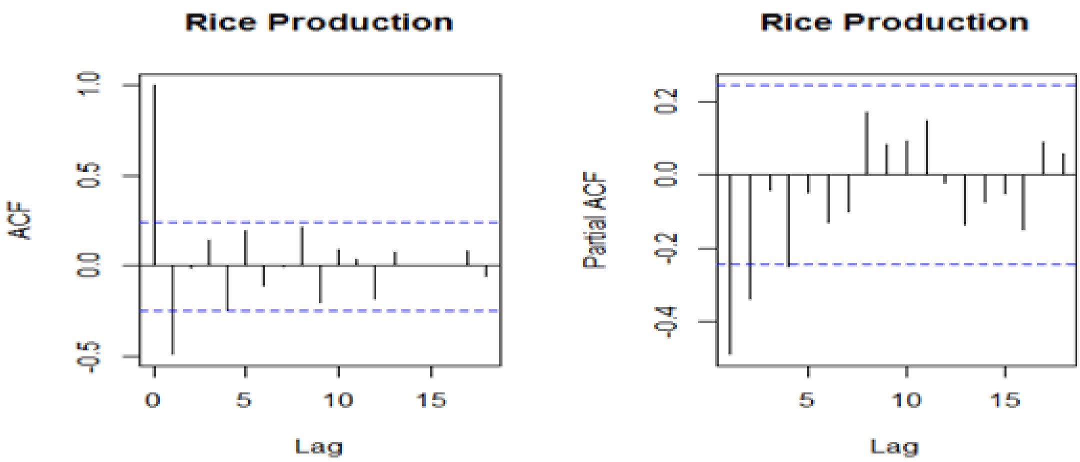

Figure 1.7: ACF and PACF plot for annual rice production

The MA(1) model has fitted representation:

$$X_t = -\underset{(0.0795)}{0.7694}\,\eta_{t-1} + \hat{\eta}_t.$$

The next step in model fitting is diagnostics. This investigation includes the analysis of the residuals as well as model comparisons. The standardized residuals and ACF of residuals are plotted in Figure 1.8. Both the plots suggest that there is no significant dependency in the residuals. The calculated value of Ljung-Box statistic (2.1281) is less than the 5% chi-square critical value 10.117 at degrees of freedom 20, conclude no dependency in residual series.

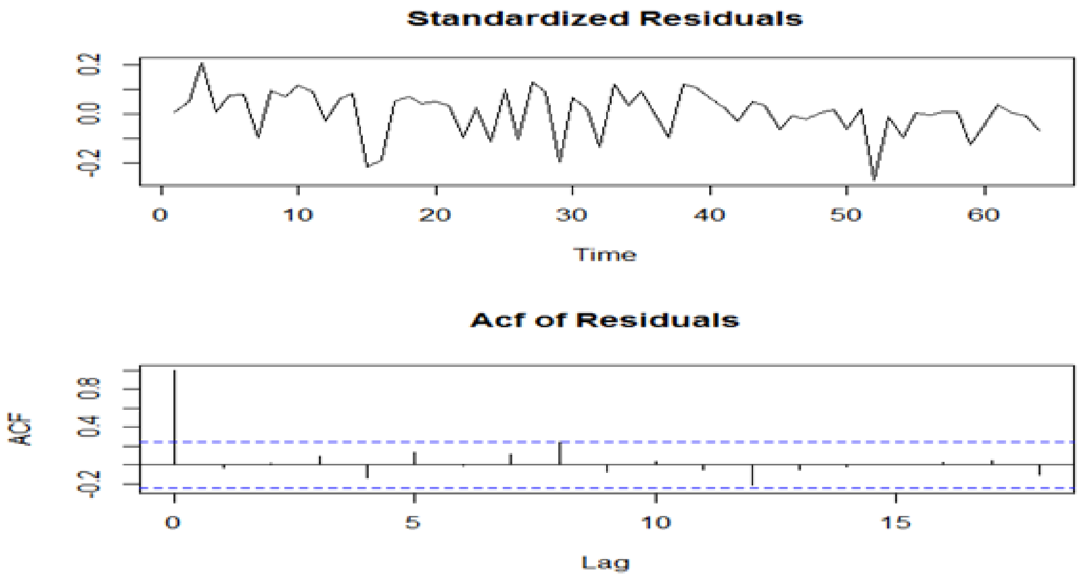

FIGURE 1.8: Plot of residuals and ACF of residuals for annual rice production

Comparing values of AIC and BIC obtained by fitting the different p and q ranging from 0 to 4 in Table 1.2, the AIC and BIC criteria both suggest a MA(1) model. Thus we take the fitted MA(1) model as adequate.

Order (p,q)	0	1	2	3	4
0	*AIC*	**-1.9987**	-1.6352	-1.6137	-1.5843
	BIC	**-1.8250**	-1.5677	-1.5125	-1.4494
1	-1.621	-1.6142	-1.9440	-1.8377	-1.8972
	-1.587	-1.5462	-1.8420	-1.7016	-1.7271
2	-1.615	-1.9291	-1.9108	-1.8062	-1.8653
	-1.546	-1.7961	-1.7736	-1.6346	-1.6595
3	-1.687	-1.6559	-1.8874	-1.8773	-1.9788
	-1.583	-1.5175	-1.7144	-1.6697	-1.7365
4	-1.637	-1.6116	-1.7556	-1.7302	-1.8949
	-1.498	-1.4371	-1.5461	-1.4858	-1.7156

Table 1.2: AIC and BIC for fitted models for rice production data

1.8 Outline of the Thesis

The linear time series models available in the literature are not adequate to model the financial time series. So, new classes of models are introduced to deal with financial time series. Chapter 2 mainly discusses the characteristics of financial time series. The models for financial time series may be broadly classified as observation driven and parameter driven models. In observation driven models, the conditional variance is assumed to be a function of the past observations, which introduces the heteroscedasticity in the model. The famous models such as Autoregressive Conditional Heteroscedastic (ARCH) model of Engle (1982) and Generalized ARCH (GARCH) model of Bollerslev (1986) are examples of these. While in the case of parameter driven models, the conditional variances are generated by some latent processes. The Stochastic Volatility model of Taylor (1986) is the example of parameter driven model. Then, we discuss financial duration concepts and duration models for modelling transaction durations in financial markets. We focus on Autoregressive Conditional Duration model proposed by Engle and Russell (1998) and Stochastic Conditional Duration model proposed by Bauwens and Veredas (2004). We summarize the properties of these models in Chapter 2. One of our objectives in this study is to identify some non-Gaussian time series models and study their suitability for modelling stochastic volatility and conditional durations in finance.

Birnbaum-Saunders (BS) distribution, introduced by Birnbaum and Saunders (1969b), has received considerable attention in the recent years in the context of lifetime modelling. Though the model has been promoted as a life time model, its shape characteristics, tail properties and non-monotone hazard function all suggest that the BS model can be used more generally for modelling non-negative random variables. We introduce a BS Autoregressive Moving Average sequence in Chapter 3, with an idea to develop SV models induced by non-Gaussian volatility sequences. A stationary sequence of random variables with BS marginal distribution is constructed using a Gaussian autoregressive moving average sequence. The parameters of the model are then estimated by maximum likelihood method and the resulting estimators are shown to be consistent and asymptotically normal. A simulation study is carried out in order to assess the performance of the estimators. To illustrate the application of the proposed model, we have analysed two sets of real data - index of Coal production in Eight Core Industries and the number of Foreign Tourist Arrivals in India.

In Chapter 4, we discuss the properties of Birnbaum-Saunders Stochastic Volatility model. The volatility sequences are generated by BS-AR(1) model discussed in Chapter 3. We have employed the moment method and Efficient Importance

Sampling(EIS) method to estimate the model parameters. Simulation studies are carried out to assess the performance of the estimation method and the proposed model is finally used to analyze Rupee/Dollar exchange rate and S&P 500 Opening index data.

The traditional models based on Gaussian distribution are very often not supported by real-life data because of long tails and asymmetry present in these data. Since the class of asymmetric Laplace distributions can account for leptokurtic and skewed data they are natural candidates to replace Gaussian models and processes. In Chapter 5, we propose a stochastic volatility model generated by first order autoregressive process with asymmetric Laplace marginal distribution as an alternative to normal-lognormal SV model. The model parameters are estimated using the method of moments as the likelihood function is intractable. The simulation results indicate that the estimators behave well when the sample size is large. The model is used to analyze two sets of data and found that, it captures the stylized facts of the financial time series.

The durations between market activities such as trades, quotes, etc. provide useful information on the underlying assets while analyzing financial time series. In Chapter 6 we present a brief review of models for such durations and also propose some new conditional duration models based on inverse Gaussian distribution. The non-monotonic nature of the failure rate of inverse Gaussian distribution makes it suitable for modelling the conditional durations in financial time series. First, we proposed an observation drive model – Autoregressive Conditional Duration model based on the inverse Gaussian distribution. Second, a parameter driven model called Stochastic Duration model with inverse Gaussian innovations is constructed.

The model parameters are estimated by the method of maximum likelihood and an EIS method respectively. A simulation experiment is conducted to check the performance of the proposed estimators. Finally a real data analysis is provided to illustrate the practical utility of the models.

Concluding remarks are given in Chapter 7 to summarize the most important contributions of this thesis and some of the problems identified for future research.

Chapter 2

Models for Financial Time Series

2.1 Introduction

Financial time series are well known for their uncertainty, especially the irregularity in the behaviour of certain financial indices such as stock prices, exchange or interest rates, government bond prices, yield of treasury bills and so on, that are prone to time dependent variability. Such variability, otherwise known as volatility can generate very high frequency series of variables which are stochastic in nature, the dynamics of which can best be described by means of stochastic models. As a result of the added uncertainty, statistical theory and methods play an important role in financial time series analysis.

There are two main objectives of investigating financial time series. First, it is important to understand how prices behave. The variance of the time series is particularly relevant. Tomorrow's price is uncertain and it must therefore be described by a probability distribution. This means that statistical methods are the natural way to investigate prices. Usually one builds a model, which is a detailed description of how successive prices are determined. The second objective is to use our knowledge of price behaviour to reduce risk or take better decisions. Time series models may for instance be used for forecasting, option pricing and risk management. This motivates more and more statisticians and econometricians to devote themselves to the development of new (or refined) time series models and methods.

Many finance problems involve the arrival of events such as prices or trades in irregular time intervals, a new direction of modelling is necessary to explain the properties of such data. The durations between market activities such as trades, quotes, etc. provide useful information on the underlying assets while analysing financial time series. Hence it is important to model the dynamic behaviour of such durations in finance.

The objective of this chapter is to understand various aspects of financial time series and list some of the important financial time series models and their useful characteristics. In the next section, we address some of the stylized facts of financial time series which play important role in volatility modelling. Section 2.3 introduces models for volatility and basic properties. In Section 2.4 we discuss about the conditional duration models in finance.

2.2 Stylized facts of Financial Time Series

Financial time series analysis is concerned with the theory and practice of asset valuation over time. One of the objectives of analysing financial time series is to model the volatility and forecast its future values. The volatility is measured in terms of the conditional variance of the random variables involved. The conditional variances in the case of financial time series are not constants. They may be functions of some known or unknown factors. This leads to the introduction of conditional heteroscedastic models for analysing financial time series. In financial markets, the data on price P_t of an asset at time t is available at different time points. However, in financial studies, the experts suggest that the series of returns be used for analysis instead of the actual price series, see Tsay (2005). For a given series of prices $\{P_t\}$, the corresponding series of returns is defined by

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

The advantages of using the return series are, 1) for an investor, the return series is a scale free summary of the investment opportunity, 2) the return series are easier to handle than the price series because of their attractive statistical properties. Further consideration of the attractive statistical properties, suggested that, the log-return series defined by $r_t = \log\left(P_t/P_{t-1}\right)$ is more suitable for analysing the stochastic nature of the market behaviour. Hence, we focus our attention on the modelling and analysis of the log-return series in this thesis and we refer $\{r_t = \log\left(P_t/P_{t-1}\right), t = 1, 2, ...\}$ as financial time series.

Empirical studies on financial time series (See Mandebrot (1963) and Fama (1965)) show that the series $\{r_t\}$ defined above is characterized by the properties such as

1. Absence of autocorrelation in $\{r_t\}$.
2. Significant serial correlation in $\{r_t^2\}$.
3. The marginal distribution $\{r_t\}$ is heavy-tailed.
4. Conditional variance of r_t given the past is not constant.
5. Volatility tends to form clusters, i.e., after a large (small) price change (positive or negative) a large (small) price change tends to occur. This attribute is called volatility clustering.

To get an intuitive feel of these stylized facts, a typical example is shown in Figure 2.1, where Bombay Stock Exchange (BSE) opening index during July 02, 2007 to May 13, 2016 is plotted.

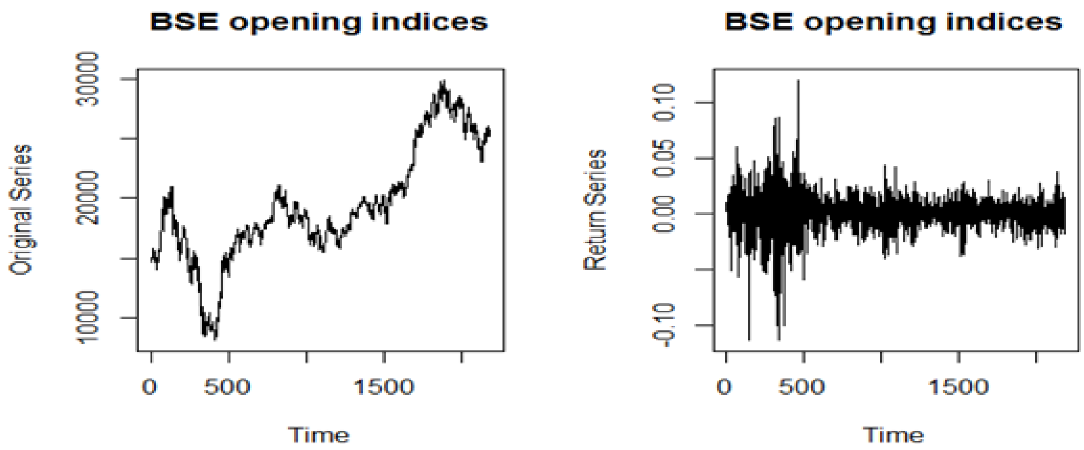

FIGURE 2.1: Time series plot of BSE index and returns

Right panel in Figure 2.1 plots the time series of returns of the indices under study. Time series plot of indices are clearly non-stationary, however daily returns are stationary.

Summary statistics for daily index returns r_t are provided in Table 2.1. These statistics are used in the discussion of some stylized facts related to the probability density function of the return series.

Statistics	BSE index
Observations	2185
Mean	0.0003
Median	0.0004
Maximum	0.1205
Minimum	-0.1138
Std. Dev.	0.0166
Skewness	-0.3853
Kurtosis	10.2901

Table 2.1: Summary statistics for BSE log-returns

As seen in Table 2.1, BSE index returns have excess kurtosis well above 3 indicates leptokurtic and fat tails of returns. The ACF of returns and squared returns are plotted in Figure 2.2. While the autocorrelation of returns are all close to zero, autocorrelation of squared returns are positive and significantly larger than zero. Since the autocorrelation is positive, it can be concluded, that small (positive or negative) returns are followed by small returns and large returns follow large ones again.

Figure 2.3 compares histogram of BSE index return with approximate normal density. It is clear from the figure that the empirical distribution of daily returns does

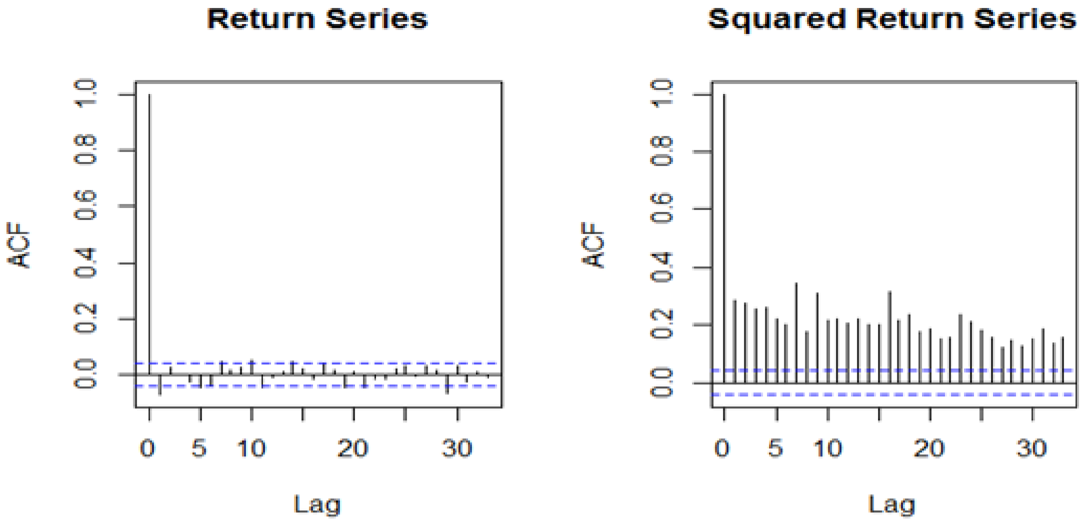

FIGURE 2.2: ACF of returns and squared returns of BSE index

not resemble a Gaussian distribution. The peak around zero appears clearly, but the thickness of the tails is more difficult to visualize.

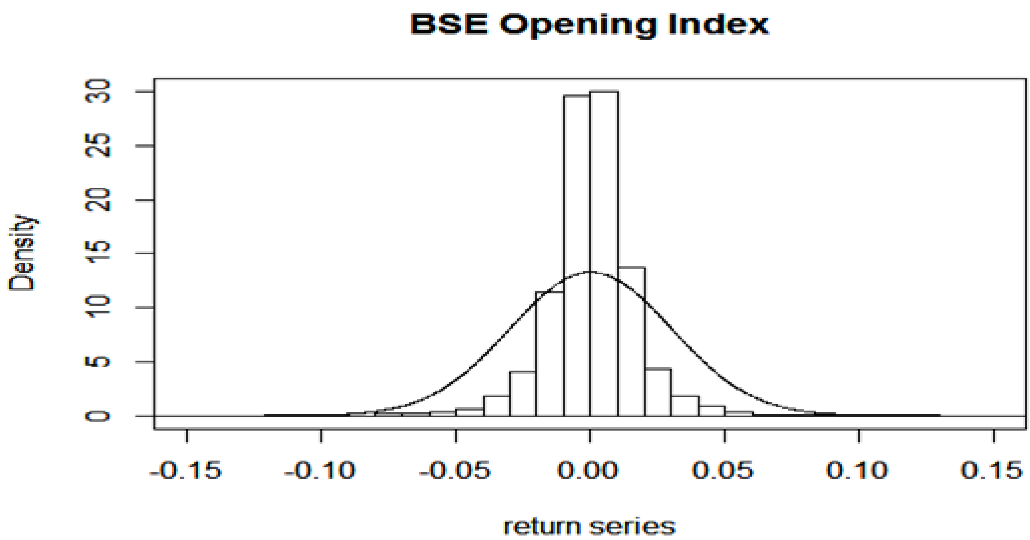

FIGURE 2.3: Histogram of BSE index return and normal approximation

2.3 Models for Volatility

The models described in the previous chapter are often very useful in modelling time series in general. However, they have the assumption of constant error variance. As a result the conditional variance of the observation at any time given the past will remain a constant, a situation referred to as homoscedasticity. This is considered to be unrealistic in many areas of economics and finance as the conditional variances are non-constants. Therefore, two prominent classes of models have been developed by researchers which capture the time-varying autocorrelated volatility process: the autoregressive conditional heteroscedastic (ARCH) model, introduced by Engle (1982), assumes that the conditional variances are some functions of the squares of the past returns and are referred to as the observation driven models. Another class of models to study the price changes is the SV models introduced by Taylor (1986), where the conditional variance at time t is assumed to be a stochastic process in terms of some latent variables, which are referred to as the parameter driven models.

2.3.1 Autoregressive Conditional Heteroscedastic Models

The ARCH model introduced by Engle (1982) was a first attempt in econometrics to capture volatility clustering in time series data. In particular, Engle (1982) used conditional variance to characterize volatility and postulated a dynamic model for conditional variance. We will discuss the properties and some generalizations of the ARCH model in subsequent sections; for a comprehensive review of this class of models we refer to Bollerslev et al. (1992). ARCH models have been widely used in financial time series analysis and particularly in analyzing 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. Specifically, an ARCH(p) model for $\{r_t\}$ is defined by

$$r_t = \sqrt{h_t}\varepsilon_t \ , \quad h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, r_{t-i}^2, \tag{2.1}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with mean zero and variance 1, $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$. If $\{\varepsilon_t\}$ has standardized Gaussian distribution conditional on h_t, r_t follows normal with mean 0 and variance h_t. The Gaussian assumption of ε_t is not critical. We can relax it and allow for more heavy-tailed distributions, such as the Student's t -distribution, as is typically required in finance. Now we describe the properties of a first order ARCH model in detail.

ARCH(1) model and properties:

The structure of the ARCH model implies that the conditional variance h_t of r_t, evolves according to the most recent realizations of r_t^2 analogous to an AR(1) model. Large past squared shocks imply a large conditional variance for r_t. As a consequence, r_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

$$r_t = \sqrt{h_t}\, \varepsilon_t \ , \quad h_t = \alpha_0 + \alpha_1 \, r_{t-1}^2, \tag{2.2}$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$.

1. The unconditional mean of r_t is zero, since

$$E\left(r_t\right) = E\left(E\left(r_t | r_{t-1}\right)\right) = E\left(\sqrt{h_t} E\left(\varepsilon_t\right)\right) = 0.$$

2. The conditional variance of r_t is

$$E\left(r_t^2 | r_{t-1}\right) = E\left(h_t \varepsilon_t^2 | r_{t-1}\right) = h_t E\left(\varepsilon_t^2 | r_{t-1}\right) = h_t = \alpha_0 + \alpha_1 r_{t-1}^2.$$

3. The unconditional variance of r_t is

$$\begin{aligned} Var\left(r_t\right) &= E\left(r_t^2\right) = E\left(E\left(r_t^2 | r_{t-1}\right)\right) \\ &= E\left(\alpha_0 + \alpha_1 r_{t-1}^2\right) \\ &= \alpha_0 + \alpha_1 E\left(r_{t-1}^2\right) \\ &= \frac{\alpha_0}{1-\alpha_1}. \end{aligned}$$

4. Assuming that the fourth moment of r_t are finite, the Kurtosis K of r_t, is given by

$$K = \frac{E\left(r_t^4\right)}{E\left(r_t^2\right)^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3,$$

provided $\alpha_1^2 < 1/3$.

The ARCH model with a conditionally normally distributed r_t leads to heavy tails in the unconditional distribution. In other words, the excess kurtosis of

r_t is positive and the tail of the distribution of r_t is heavier than that of the normal distribution.

5. The autocovariance of r_t is defined by

$$Cov\left(r_t, r_{t-k}\right) = E\left(r_t\, r_{t-k}\right) - E\left(r_t\right) E\left(r_{t-k}\right)$$

$$= E\left(r_t\, r_{t-k}\right) = E\left(\sqrt{h_t}\sqrt{h_{t-k}}\right) E\left(\varepsilon_t \varepsilon_{t-k}\right) = 0.$$

Then the autocorrelation function of r_t is zero. The ACF of $\{r_t^2\}$ is $\rho_{r_t^2}(k) = \alpha_1^k$ and notice that $\rho_{r_t^2}(k) \geq 0$ 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. The $\{r_t\}$ is stationary process for which $0 \leq \alpha_1 < 1$ is satisfied, since the variance of r_t must be positive. These properties continue to hold for general ARCH models, but the formulas become more complicated for higher order ARCH models.

2.3.2 Generalized ARCH (GARCH) Models

The GARCH model is an extension of Engle's work by Bollerslev (1986) that allows the conditional variance to depend on the previous conditional variances and the squares of previous returns. The possibility that estimated parameters in ARCH model do not satisfy the stationarity condition increases with lag. Thus GARCH model is an alternative to ARCH model. The GARCH (p, q) is defined by

$$r_t = \sqrt{h_t}\varepsilon_t \ , \quad h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, r_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \tag{2.3}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with mean 0 and variance 1; $\{\varepsilon_t\}$ is assumed to be independent of $\{h_{t-i}, \, i \geq 1\}$. α_0, α_i and β_j are unknown parameters satisfying $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$. The constraint on $\alpha_i + \beta_i$ implies that the unconditional variance of r_t is finite, whereas its conditional variance h_t evolves over time. As before, ε_t is assumed to be a standard normal distribution.

GARCH (1,1) model and properties:

Let us now consider the GARCH (1,1) model, which is the most popular one for modelling asset-return volatility. We represent this model as

$$r_t = \sqrt{h_t}\varepsilon_t \, , \qquad h_t = \alpha_0 + \alpha_1 \, r_{t-1}^2 + \beta_1 h_{t-1}, \tag{2.4}$$

where $\varepsilon_t \sim N(0, 1)$ and $0 \leq \alpha_1, \beta_1 < 1$, $\alpha_1 + \beta_1 < 1$.

1. The unconditional mean of r_t is zero, since

$$E(r_t) = E(E(r_t | r_{t-1})) = E\left(\sqrt{h_t} E(\varepsilon_t)\right) = 0.$$

2. The conditional variance of r_t is

$$\begin{aligned} E\left(r_t^2|r_{t-1}\right) &= E\left(h_t\varepsilon_t^2|r_{t-1}\right) \\ &= h_t E\left(\varepsilon_t^2|r_{t-1}\right) \\ &= h_t = \alpha_0 + \alpha_1\, r_{t-1}^2 + \beta_1 h_{t-1}. \end{aligned}$$

3. The unconditional variance of r_t is

$$\begin{aligned} Var\left(r_t\right) = E\left(r_t^2\right) = E\left(E\left(r_t^2|r_{t-1}\right)\right) = E\left(\alpha_0 + \alpha_1\, r_{t-1}^2 + \beta_1 h_{t-1}\right) \\ = \alpha_0 + \alpha_1 E\left(r_{t-1}^2\right) + \beta_1 E\left(h_{t-1}\right) \end{aligned}$$

Under stationarity we get

$$Var(r_t) = \frac{\alpha_0}{1-(\alpha_1+\beta_1)}.$$

4. The Kurtosis of r_t, K, is given by

$$K = \frac{3\left[1-(\alpha_1+\beta_1)^2\right]}{1-(\alpha_1+\beta_1)^2 - 2\alpha_1^2} > 3.$$

Consequently, similar to ARCH models, the tail of the marginal distribution of GARCH(1,1) process is heavier than that of a normal distribution if $1 - 2\alpha_1^2 - (\alpha_1+\beta_1)^2 > 0$.

5. The ACF of $\{r_t\}$ is zero and the ACF of $\{r_t^2\}$ is given by

$$\rho_{r_t^2}(k) = (\alpha_1+\beta_1)^{k-1}\ \frac{\alpha_1\left(1-\alpha_1\beta_1-\beta_1^2\right)}{1-2\alpha_1\beta_1-\beta_1^2},\ \ k = 1, 2, ...$$

2.3.3 Stochastic Volatility Models

This is a class of parameter driven model in which the volatility at time t is described as a latent stochastic process, such as the one generated by an autoregressive model. An appealing feature of the SV model is its close relationship to financial economic theories. The univariate SV model proposed by Taylor (1986) is given by,

$$r_t = \varepsilon_t \exp\left(h_t/2\right), \quad h_t = \alpha + \rho h_{t-1} + \eta_t\,; \;|\rho| < 1,\; t = 1, 2, ..., \tag{2.5}$$

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

As η_t is Gaussian, $\{h_t\}$ is a Gaussian autoregressive process. It will be (strictly and covariance) stationary if $|\rho| < 1$ with:

$$\mu_h = E\left(h_t\right) = \frac{\alpha}{1-\rho},$$

$$\sigma_h^2 = V\left(h_t\right) = \frac{\sigma_\eta^2}{1-\rho^2}.$$

As $\{\varepsilon_t\}$ is always stationary, $\{r_t\}$ will be stationary if and only if $\{h_t\}$ is stationary, r_t being the product of two stationary process. All odd moments of r_t vanish and using the property of log-normal distribution, all the even moments of r_t can be obtained if h_t is stationary. In particular the kurtosis is

$$K = \frac{E\left(r_t^4\right)}{E\left(r_t^2\right)^2} = 3\;\exp\left(\sigma_h^2\right)\;\geq 3,$$

which shows that the SV model has fatter tails than the corresponding normal distribution. The dynamic properties of r_t are easy to find. First, as $\{\varepsilon_t\}$ is independent and identically distributed, $\{r_t\}$ is a martingale difference and is a white noise if $|\rho| < 1$. As h_t is a Gaussian AR(1),

$$Cov\left(r_t^2, r_{t-k}^2\right) = E\left(r_t^2 r_{t-k}^2\right) - E\left(r_t^2\right) E\left(r_{t-k}^2\right)$$

$$= E\left(\exp\left(h_t + h_{t-k}\right)\right) - \left(E\left(\exp\left(h_t\right)\right)\right)^2$$

$$= \exp\left(2\mu_h + \sigma_h^2\right)\left(\exp\left(\sigma_h^2 \rho^k\right) - 1\right),$$

and so

$$\rho_{r_t^2}(k) = \frac{Cov\left(r_t^2, r_{t-k}^2\right)}{V\left(r_t^2\right)} = \frac{\exp\left(\sigma_h^2 \rho^k\right) - 1}{3\exp\left(\sigma_h^2\right) - 1} \simeq \frac{\exp\left(\sigma_h^2\right) - 1}{3\exp\left(\sigma_h^2\right) - 1}\rho^k.$$

Note that if $\rho < 0$, $\rho_{r_t^2}(k)$ can be negative, unlike ARCH models. This resembles the autocorrelation function of an ARMA(1,1) process. Thus the SV model behaves in a manner similar to the GARCH(1,1) model. Finally, note that there is no need for non-negativity constraints or for bounded kurtosis constraints on the coefficients. This is a great advantage with respect to GARCH models. A review of the properties of SV models may be found in Taylor (1994) and Tsay (2005).

Despite theoretical advantages, the SV models have not been popular as the ARCH models in practical applications. The main reason is that the likelihood function for the SV model is not easy to evaluate unlike in the case of ARCH models. A variety of estimation procedures have been proposed to overcome this difficulty, including, for example, the Generalized Method of Moments (GMM) used by Melino and Turnbull

(1990), the Quasi Maximum Likelihood (QML) approach followed by Harvey et al. (1994) and Ruiz (1994), the Efficient Method of Moments (EMM) applied by Gallant et al. (1997), and Markov Chain Monte Carlo (MCMC) procedures used by Jacquier et al. (1994) and Kim et al. (1998). For a survey of these estimation procedures, one can refer Ghysels et al. (1996), Broto and Ruiz (2004) and Bauwens et al. (2012).

2.4 Models for Durations

The statistical analysis of sequence of durations between events is well studied in the area of point processes which includes the renewal process as a special case. In the analysis of financial time series the durations between market activities such as trades, quotes, etc. provide important information on the underlying asset. These durations are irregularly time-spaced and they form a sequence of random variables. As a result, the number of financial transactions taken place in any interval forms a point process called financial point process. For a specified asset, longer durations indicate lack of trading activities, which in turn signify a period of no new information. On the other hand arrival of new information often results in heavy trading and hence leads to shorter durations. The dynamic behaviour of durations thus contains useful information about market activities. Furthermore, since financial markets typically take a period of time to uncover the effect of new information, active trading is likely to persist for a period of time, resulting in clusters of short durations. Consequently, durations might exhibit characteristics similar to those of asset volatility, which is an important aspect of financial time series. Such features may be captured in alternative ways through different dynamic models based on either duration or intensity representation of a point process. These considerations motivated Engle and Russell (1998) to introduce a class of dynamic models known as Autoregressive Conditional Duration models, described below.

2.4.1 Autoregressive Conditional Duration (ACD) Models

Let $\{T_i, i \geq 0\}$ be a sequence of times of occurrence of certain financial events and we assume that $0 = T_0 < T_1 < T_2 < ....$ Then the i^{th} duration X_i, the interval between the *i-1*th and i^{th} occurrence of the event is defined by $X_i = T_i - T_{i-1},\ \ i = 1, 2, ...$ The basic ACD model of Engle and Russell (1998) expresses the seasonally adjusted duration X_i in the form of a multiplicative error model as

$$X_i = \psi_i \, \varepsilon_i, \tag{2.6}$$

where $\{\varepsilon_i\}$ is a sequence of independent and identically distributed non-negative random variables with unit mean. Here ψ_i is introduced as the conditional expectation of the duration X_i given the information on the past durations. That is, $\psi_i = E(X_i|F_{i-1})$, where $F_{i-1} = \sigma(X_1, X_2, ..., X_{i-1})$ is the sigma field generated by $(X_1, X_2, ..., X_{i-1})$. Engle and Russell (1998) defined the ACD (p, q) model specifying ψ_i as

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}, \tag{2.7}$$

where p and q are non-negative integers. The following conditions are imposed on the parameters for the stationarity of the sequence:

$$\omega > 0, \quad \alpha_j \geq 0, \quad \beta_j \geq 0, \quad \sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1. \tag{2.8}$$

Different classes of ACD models can be defined either by the choice of the functional form of ψ_i or by the choice of the distribution for ε_i. For example, ACD models with several standard distributions such as Exponential and Weibull (Engle and Russell (1998)), Burr (Grammig and Maurer (2000)), etc. have been studied in the literature, see Pacurar (2008) for a detailed survey. The equations (2.6) and (2.7) define the dynamical structure of the standard ACD models, which can be viewed as an observation driven model like GARCH model described in Section 2.3. However, in the latter case the model was specified for the conditional variance of the returns.

The standard ACD model has been extended in several ways, directed mainly to improving the fitting of the stylized facts of financial durations. The strong similarity between the ACD and GARCH models nurtured the rapid expansion of alternative specifications of conditional durations. An equally important model for analysing financial durations is a class of parameter driven model introduced by Bauwens and Veredas (2004) known as Stochastic Conditional Duration model.

2.4.2 Stochastic Conditional Duration (SCD) Models

In this Section, we analyse a class of parametric models for durations, which are referred to as SCD models proposed by Bauwens and Veredas (2004). In contrast to the ACD model of Engle and Russell (1998), in which the conditional mean of the duration is modelled as a conditionally deterministic function of past information, the SCD model treats the conditional mean of durations as a stochastic latent process, with innovations to the process captured by an appropriate distribution with positive support. As such, the contrast between the two specifications mimics the contrast between the GARCH and SV frameworks for capturing the conditional volatility of financial returns. In particular, as is the case with the SV model, the SCD model presents a potentially more complex estimation problem than its alternative, by augmenting the set of unknowns with a set of unobservable latent factors.

An SCD model of order one is defined by

$$X_i = e^{\psi_i}\varepsilon_i, \psi_i = \omega + \beta\,\psi_{i-1} + u_i\,,\ i = 1, 2, ..., \tag{2.9}$$

where u_i follows independent and identically distributed $N\left(0, \sigma^2\right)$ so that $\{\psi_i\}$ defines a Gaussian AR(1) sequence and ε_i is as defined in the case of (2.6). Unlike in the case of ACD models, the analysis of SCD model is more complicated due to the presence of latent variables, ψ_i, which are not observable. Economically, the latent factor can be interpreted as information flow that cannot be observed directly but drives the duration process. In this sense, the SCD model is the counterpart of the SV model introduced by Taylor (1986).

A difficulty associated with SCD framework is the parameter estimation because no explicit expression for the likelihood function of SCD model is directly available due to the presence of latent structure in ψ_i . The evaluation of the likelihood function of the SCD model requires computing an integral that has the dimension of the sample size. Bauwens and Galli (2009) developed ML estimation based on the efficient importance sampling (EIS) method for computing such integral. Other methods, that are less demanding in computing time, do not evaluate the exact likelihood function. The easiest two techniques are quasi-maximum likelihood (QML) and generalized method of moments (GMM). These techniques provide asymptotically consistent estimators and previous research seems to indicate that the behaviour of the QML estimator is better than the one of GMM in the context of the stochastic volatility model; see Ruiz (1994) and Jacquier et al. (1994). Bauwens and Veredas (2004) used QML based on the transformation of the model into a linear state space representation and the application of the Kalman filter.

Chapter 3

Birnbaum-Saunders Autoregressive Moving Average Processes

3.1 Introduction

Most of the theory behind the analysis of time series is based on the assumption that the innovation sequence which generates the process is normally distributed. If the model for the time series is linear, as in the autoregressive moving average case, then the assumed normality of the innovation implies that the marginal distribution of the observed values will be normal as well. However time series in which observations are non Gaussian nature are very common in many areas. The literature on non-Gaussian time series deals with ARMA processes with several non-normal marginal distributions. When the support of these marginal distributions are restricted to be non-negative, they are used to model the life time data. Let us consider a stationary and invertible first order autoregressive moving average (ARMA(1,1)) model defined by

$$X_t = \rho X_{t-1} + \theta \eta_{t-1} + \eta_t, \quad |\rho| < 1, \ |\theta| < 1, \ t = 1, 2, \dots \ , \tag{3.1}$$

where $\{\eta_t\}$ is a sequence of independent and identically distributed random variables termed as innovations. It is assumed that X_t depends only on the present and the past innovations.

Birnbaum-Saunders (BS) distribution, introduced by Birnbaum and Saunders (1969b), has received considerable attention in the recent years in the context of lifetime modelling This is due to its many attractive properties and its close relationship with the normal distribution. These aspects make the BS distribution a natural and meaningful alternative candidate to the normal model to accommodate positive skewness and non-negative support in the data. Recently, much work has been carried out on BS distribution while modelling non-negative lifetime data; see Kundu et al. (2010) and the references therein for more details. In general, its applications are not only limited to the mentioned area, but its use in business, economics, finance, industry, insurance, inventory, quality control, and toxicology have also been considered, among others; see Jin and Kawczak (2003), Lio and Park (2008), Ahmed et al. (2010), Bhatti (2010), Paula et al. (2012), Marchant et al. (2013) Leiva et al. (2014a,b,c), Wanke and Leiva (2015) and Leiva et al. (2016). A recent text book by Leiva (2016) gives an up to date survey on applications on BS distribution in several areas. However, there is no development on a stationary sequence of BS random variables with an observation at time t depending on its past values. In this chapter, we construct an ARMA process with BS marginal distribution through a Gaussian ARMA process with the primary aim of using it for modelling non-negative variables. The modelling of non-negative random variables play a major role in the study of financial time series, where one has to model the evolution of conditional variances known as stochastic volatility (see Tsay (2005)).

The rest of this chapter is organized as follows. Section 3.2 describes the elementary properties of the BS distribution. In Sections 3.3 to 3.5 we construct the AR(1), MA(1) and ARMA(1,1) models with BS marginal distribution and study their second-order properties. The maximum likelihood method of estimation and its asymptotic properties are discussed in Sections 3.6 and 3.7. A simulation study is carried out in Section 3.8, while Section 3.9 deals with a data analysis for illustrating the results developed in the preceding sections.

3.2 Birnbaum-Saunders distribution

A random variable Y has a BS distribution if it can be expressed as

$$Y = \beta \left[\frac{\alpha}{2} X + \sqrt{\left(\frac{\alpha}{2} X \right)^2 + 1} \right]^2, \tag{3.2}$$

where X is a random variable following the standard normal distribution, i.e., $X \sim N(0, 1)$.

The probability density function of Y is given by

$$f_Y(y;\alpha,\beta) = \frac{1}{2\alpha\beta\sqrt{2\pi}}\left[\left(\frac{\beta}{y}\right)^{1/2} + \left(\frac{\beta}{y}\right)^{3/2}\right] \exp\left(-\frac{1}{2\alpha^2}\left[\frac{y}{\beta} + \frac{\beta}{y} - 2\right]\right), \quad y > 0; \tag{3.3}$$

here, $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively. Hereafter, this distribution will be denoted by $BS(\alpha,\beta)$. The shape of the density function in (3.3) is governed by the parameter α. It can be shown that $f_Y(\cdot\,;\alpha,\beta)$ is a unimodal function and for fixed β, the mode is an increasing function of α. See Leiva (2016) for details.

The transformation in (3.2) is very useful as it enables the determination of the moments of Y through known results on expectations of functions of X. Using (3.2), the mean, variance and the coefficients of skewness and kurtosis can be respectively obtained as

$$E(Y) = \beta\left(1 + \frac{1}{2}\alpha^2\right), Var(Y) = (\alpha\beta)^2\left(1 + \frac{5}{4}\alpha^2\right),$$

$$\gamma = \frac{16\alpha^2(11\alpha^2+6)}{(5\alpha^2+4)^3}, K = 3 + \frac{6\alpha^2(93\alpha^2+41)}{(5\alpha^2+4)^2}. \tag{3.4}$$

It is clear that both the mean and variance increase as α increases. The coefficient of skewness, γ converges to zero as $\alpha \to \infty$, and so the shape of the probability density function in (3.3) become symmetric as $\alpha \to \infty$. Moreover, kurtosis K tends to 3 (that of normal distribution) as $\alpha \to \infty$. More details on BS distribution may be found in Leiva (2016).

3.3 BS-AR(1) Model

Let us consider a stationary process $\{X_t\}$ given by

$$X_t = \rho X_{t-1} + \eta_t, \quad |\rho| < 1, \quad t = 1, 2, \dots \ , \tag{3.5}$$

where X_0 is a standard normal random variable independent of η_1. Then $\{X_t\}$ is a stationary Gaussian AR(1) process with standard normal $(N(0,1))$ marginal distribution with probability density function

$$\phi_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}, -\infty < x < \infty. \tag{3.6}$$

Then the distribution of the innovation η_t is $N(0, 1-\rho^2)$ with probability density function

$$\phi_\eta(x) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\{-x^2/2(1-\rho^2)\}, -\infty < x < \infty. \tag{3.7}$$

Note that $\{X_t\}$ defined by (3.5) is a stationary Markov sequence with one-step transition density function of X_t at x_t , given $X_{t-1} = x_{t-1}$, as

$$\begin{aligned}\phi_{t|t-1}(x_t|x_{t-1}) &= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\{-(x_t - \rho x_{t-1})^2/2(1-\rho^2)\}, -\infty < x_{t-1}, x_t < \infty \\ &= \phi_\eta(x_t - \rho x_{t-1}).\end{aligned} \tag{3.8}$$

The stationary bivariate density function of (X_{t-1}, X_t) is given by

$$\begin{aligned}\phi(x_{t-1}, x_t) &= \phi_\eta(x_t - \rho x_{t-1})\phi_{X_{t-1}}(x_{t-1}) \\ &= \tfrac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\{-(x_t^2 - 2\rho x_t x_{t-1} + x_{t-1}^2)/2(1-\rho^2)\}\end{aligned} \tag{3.9}$$

for $-\infty < x_{t-1}, x_t < \infty$.

Let us denote the joint distribution function corresponding to (3.9) by

$$\Phi_{t-1,t}(x_{t-1}, x_t) = \int_{-\infty}^{x_t} \int_{-\infty}^{x_{t-1}} \phi(u, v) du dv. \tag{3.10}$$

Suppose $\{X_t\}$ is a Gaussian AR(1) sequence generated by (3.5), and we define $\{Y_t\}$ by

$$Y_t = \beta \left[\frac{1}{2}\alpha X_t + \sqrt{\left(\frac{1}{2}\alpha X_t\right)^2 + 1} \right]^2 \quad for\ t = 1, 2, ... \tag{3.11}$$

Then, it is evident that Y_t follows a $BS(\alpha, \beta)$ distribution for every t and $\{Y_t\}$ is a stationary Markov sequence with BS marginal distribution. Consider the joint distribution function of Y_{t-1} and Y_t given by

$$F_{t-1,t}(y_{t-1}, y_t) = P(Y_{t-1} \leq y_{t-1}, Y_t \leq y_t)$$

$$= \Phi_{t-1,t}\left(\frac{1}{\alpha}\left\{ \sqrt{\frac{y_{t-1}}{\beta}} - \sqrt{\frac{\beta}{y_{t-1}}} \right\}, \frac{1}{\alpha}\left\{ \sqrt{\frac{y_t}{\beta}} - \sqrt{\frac{\beta}{y_t}} \right\} \right), \ y_{t-1}, y_t > 0, \tag{3.12}$$

where $\Phi_{t-1,t}(.,.)$ is as defined in (3.10). This is indeed the same bivariate distribution as given in Section 3 of Kundu et al. (2010) with $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$. The probability density function corresponding to the distribution function (3.12) is given by

$$f_{t-1,t}(y_{t-1},y_t) = \frac{1}{2\pi\sqrt{1-\rho^2}}\frac{1}{4\alpha^2\beta^2}\left(\left(\frac{\beta}{y_{t-1}}\right)^{1/2}+\left(\frac{\beta}{y_{t-1}}\right)^{3/2}\right)\left(\left(\frac{\beta}{y_t}\right)^{1/2}+\left(\frac{\beta}{y_t}\right)^{3/2}\right)$$
$$\times\exp\left(-\frac{1}{2(1-\rho^2)}\left\{\begin{array}{c}\frac{1}{\alpha^2}\left[\frac{y_{t-1}}{\beta}+\frac{\beta}{y_{t-1}}-2\right]+\frac{1}{\alpha^2}\left[\frac{y_t}{\beta}+\frac{\beta}{y_t}-2\right]\\ -\frac{2\rho}{\alpha^2}\left[\left(\sqrt{\frac{y_{t-1}}{\beta}}-\sqrt{\frac{\beta}{y_{t-1}}}\right)\right]\left(\sqrt{\frac{y_t}{\beta}}-\sqrt{\frac{\beta}{y_t}}\right)\end{array}\right\}\right).$$

Note that this density function is symmetric in y_t and y_{t-1} and so the Markov sequence is time-reversible. The one-step transition density function of $\{Y_t\}$ is then given by

$$f_{t|t-1}(y_t|y_{t-1}) = \frac{1}{2\alpha\beta\sqrt{2\pi}\sqrt{1-\rho^2}}\left(\left(\frac{\beta}{y_t}\right)^{1/2}+\left(\frac{\beta}{y_t}\right)^{3/2}\right)$$
$$\times\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\alpha}\left(\sqrt{\frac{y_t}{\beta}}-\sqrt{\frac{\beta}{y_t}}\right)-\frac{\rho}{\alpha}\left(\sqrt{\frac{y_{t-1}}{\beta}}-\sqrt{\frac{\beta}{y_{t-1}}}\right)\right]^2\right\}. \tag{3.13}$$

The mean, variance, skewness and kurtosis of Y_t are respectively given by (3.4). The autocorrelation function of a stationary sequence $\{Y_t\}$ is defined by $\rho_Y(k) = Cov(Y_t, Y_{t+k})/Var(Y_t)$. This can be computed by constructing the distribution of (Y_1, Y_{1+k}) using that of (X_1, X_{1+k}). By model (3.5), it is easy to show that the ACF of $\{X_t\}$, $\rho_X(k) = \rho^k,\ k = 0, 1, 2, ...$

For deriving the distribution of (X_1, X_{1+k}), we consider the characteristic function of (X_t, X_{t+k}) and given by,

$$\begin{aligned}\psi(u_1,u_2) &= E\left[\exp\left\{iu_1X_t+iu_2X_{t+k}\right\}\right]\\ &= E\left[\exp\left\{iu_1X_t+iu_2\left(\rho^kX_t+\sum_{j=1}^{k}\rho^{k-j}\eta_{t+j}\right)\right\}\right]\end{aligned}$$

$$
\begin{aligned}
&= E\left[\exp\left\{i\left(u_1+u_2\rho^k\right)X_t\right\}\right]\ E\left[\exp\left\{iu_2\left(\sum_{j=1}^{k}\rho^{k-j}\eta_{t+j}\right)\right\}\right] \\
&= \exp\left\{-\frac{1}{2}\left(u_1+u_2\rho^k\right)^2\right\}\left[\exp\left\{-\frac{1}{2}u_2^2(1-\rho^2)\right\}\left(\rho^{2(k-1)}+\rho^{2(k-2)}+...+1\right)\right] \\
&= \exp\left\{-\frac{1}{2}\left(u_1+u_2\rho^k\right)^2\right\}\exp\left\{-\frac{1}{2}u_2^2\left(1-\rho^{2k}\right)\right\} \\
&= \exp\left\{-\frac{1}{2}\left(u_1^2+u_2^2+2u_1u_2\rho^k\right)\right\},
\end{aligned}
$$

which is the characteristic function of bivariate normal distribution. Therefore, it can be shown that

$$
(X_1, X_{1+k}) \sim N_2\left(0, \begin{bmatrix} 1 & \rho^k \\ \rho^k & 1 \end{bmatrix}\right) \qquad k = 1, 2, ...
$$

Autocorrelation function:

We now derive the ACF of $\{Y_t\}$ in terms of that of $\{X_t\}$ by using the relation in (3.11). Let us first compute the auto-covariance function of lag k given by

$$
Cov(Y_t, Y_{t+k}) = E(Y_tY_{t+k}) - E(Y_t)E(Y_{t+k}).
$$

For this purpose, let us consider

$$
E(Y_tY_{t+k}) = E\left\{\beta\left[\frac{\alpha X_t}{2}+\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right]^2 \times \beta\left[\frac{\alpha X_{t+k}}{2}+\sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2+1}\right]^2\right\}
$$

$$= \beta^2 E \left[1 + \frac{\alpha^2 X_t^2}{2} + \frac{\alpha^2 X_{t+k}^2}{2} + \frac{\alpha^4 X_t^2 X_{t+k}^2}{4} + \alpha X_t \sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1} \right.$$

$$+ \alpha X_{t+k} \sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2 + 1} \; + \frac{1}{2}\alpha^3 X_t X_{t+k}^2 \sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1}$$

$$\left. + \frac{1}{2}\alpha^3 X_t^2 X_{t+k} \sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2 + 1} \; + \alpha^2 X_t X_{t+k} \left(\sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1} \right) \sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2 + 1} \right].$$

Note that

$$E(X_t^2) = E(X_{t+k}^2) = 1$$

$$E(X_t^2 X_{t+k}^2) = 1 + 2\rho^{2k}$$

$$E\left[X_t \sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1} \right] = E\left[X_{t+k} \sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2 + 1} \right]$$

$$= E(odd \;\; function \;\; in \;\; X_t \,) = 0$$

$$E\left[X_t^2 X_{t+k} \sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2 + 1} \right] = E\left[X_t X_{t+k}^2 \sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1} \right]$$

$$= E(odd \;\; function \;\; in \;\; X_t \,) = 0$$

Consequently, we have

$$E(Y_t Y_{t+k}) = \beta^2 \left[1 + \alpha^2 + \frac{\alpha^4}{4}(1 + 2\rho^{2k}) + \alpha^2 I_1 \right], \tag{3.14}$$

where

$$I_1 = E\left[X_t X_{t+k}\left(\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right)\left(\sqrt{\left(\frac{\alpha X_{t+k}}{2}\right)^2+1}\right)\right]$$
$$= E\left[\left\{X_t + \frac{1}{2^3}\,\alpha^2 X_t^3 + \sum_{i=2}^{\infty}(-1)^{i-1}\,\frac{1\cdot 3\ldots(2i-3)}{2^{3i}\,i!}\alpha^{2i}X_t^{2i+1}\right\}\right.$$
$$\left.\times\left\{X_{t+k} + \frac{1}{2^3}\,\alpha^2 X_{t+k}^3 + \sum_{j=2}^{\infty}(-1)^{j-1}\,\frac{1\cdot 3\ldots(2j-3)}{2^{3j}\,j!}\alpha^{2j}X_t^{2j+1}\right\}\right].$$

For non-negative integers m and n, we have [see Kotz et al. (2000), pp 261]

$$E\left(X_t^{2m+1}X_{t+k}^{2n+1}\right) = \frac{(2m+1)!(2n+1)!}{2^{m+n}}\sum_{i=0}^{\min(m,n)}\frac{(2\rho_X)^{2i+1}}{(m-i)!\,(n-i)!\,(2i+1)!} = a_{m,n}\ (say)\,.$$

Therefore,

$$I_1 = a_{0,0} + \frac{1}{2^2}\alpha^2 a_{0,1} + \frac{1}{2^6}\alpha^4 a_{1,1} + \sum_{i=2}^{\infty}(-1)^{i-1}\frac{1\cdot 3\ldots(2i-3)}{2^{3i-1}\ i!}\,\alpha^{2i}a_{0,i}$$

$$+\sum_{i=2}^{\infty}(-1)^{i-1}\frac{1\cdot 3\ldots(2i-3)}{2^{3i+2}\ i!}\,\alpha^{2i+2}a_{1,i} + \sum_{i=2}^{\infty}\sum_{j=2}^{\infty}(-1)^{i+j}\,\frac{1\cdot 3\ldots(2i-3)}{2^{3i}\ i!}\times\frac{1\cdot 3\ldots(2j-3)}{2^{3j}\ j!}\alpha^{2i+2j}a_{i,j} \tag{3.15}$$

and so

$$Cov\,(Y_t, Y_{t+k}) = \alpha^2\beta^2\left(\frac{\alpha^2\rho^{2k}}{2} + I_1\right). \tag{3.16}$$

The ACF of $\{Y_t\}$ is then readily obtained to be

$$\rho_Y(k) = \left(\frac{\alpha^2\rho^{2k}}{2} + I_1\right)\Big/\left(1 + \frac{5}{4}\alpha^2\right),\ \ k = 1, 2, \ldots. \tag{3.17}$$

The plots in Figure 3.1 compare the ACFs of the Gaussian AR(1) and the corresponding BS Markov sequences for selected values of ρ and α.

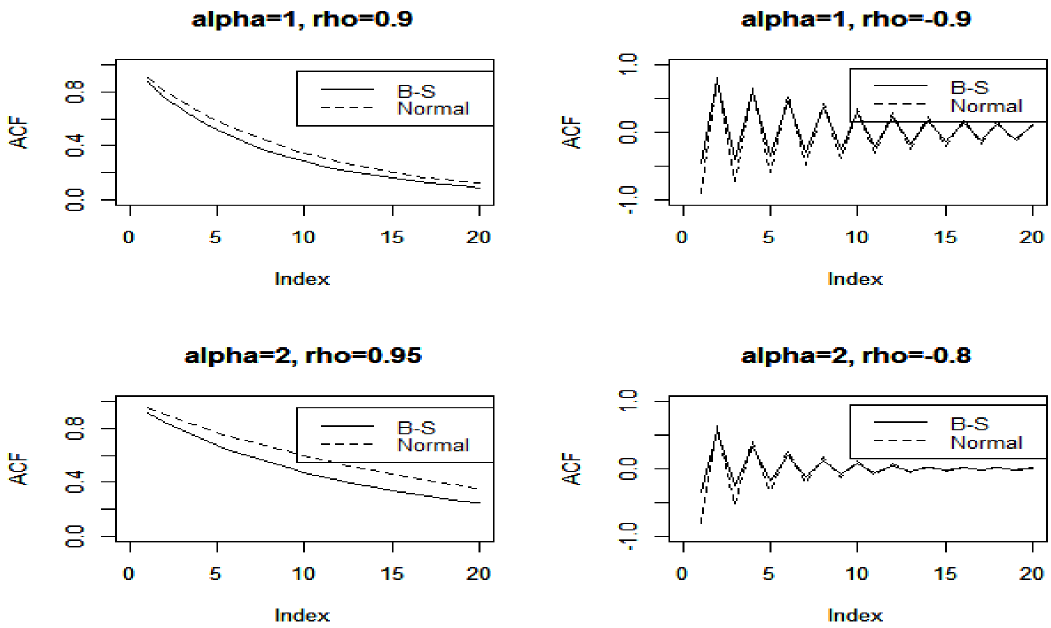

FIGURE 3.1: ACF of the Gaussian AR(1) and the corresponding BS Markov sequence

3.4 BS-MA(1) Model

In this section, we consider a BS model with a different kind of dynamics. Instead of the AR(1) process considered in the previous section, dependence is incorporated through a first order Moving Average process. Consider an invertible MA(1) model with standard normal marginal distribution,

$$X_t = \theta\eta_{t-1} + \eta_t, \quad |\theta| < 1, \quad t = 1, 2, \ldots \ . \tag{3.18}$$

Then the distribution of the innovation η_t is $N(0, 1/(1+\theta^2))$ with probability density function

$$\phi_\eta(x) = \frac{\sqrt{1+\theta^2}}{\sqrt{2\pi}} \exp\left\{-(1+\theta^2)x^2/2\right\}, \ -\infty < x < \infty. \tag{3.19}$$

Then the stationary multivariate density function of $X_1, X_2, ..., X_T$ is given by

$$\phi(x_1, x_2, ..., x_T; \, \Sigma) = \frac{1}{(2\pi)^{\frac{T}{2}} \, |\Sigma|^{\frac{1}{2}}} \ \exp\left\{-\frac{1}{2}X'\Sigma^{-1}X\right\}, \tag{3.20}$$

where $X = (x_1, x_2, ..., x_T)'$ and Σ is the correlation matrix of X and given by

$$\Sigma = \begin{bmatrix} 1 & \rho & 0 & 0 & \dots & 0 \\ \rho & 1 & \rho & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & \rho & 1 \end{bmatrix}, \rho = \frac{\theta}{1+\theta^2}.$$

If we make the transformation (3.11) to $\{X_t\}$ generated by (3.18), then it can be easily seen that $Y_t \sim BS\,(\alpha, \beta)$ distribution for every t and $\{Y_t\}$ is an invertible moving average process with BS marginal distribution. Then the joint density function of $Y_1, Y_2, ..., Y_T$ is given by

$$\begin{aligned} f_Y(y_1, y_2, ..., y_T) = \phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{y_1}{\beta}} - \sqrt{\frac{\beta}{y_1}}\right), \dots, \frac{1}{\alpha}\left(\sqrt{\frac{y_T}{\beta}} - \sqrt{\frac{\beta}{y_T}}\right); \Sigma\right) \\ \times \prod_{n=1}^{T} \frac{1}{2\alpha\beta}\left\{\left(\frac{\beta}{y_n}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_n}\right)^{\frac{3}{2}}\right\} \end{aligned} \tag{3.21}$$

for $y_t, y_{t-1}, ..., y_1 > 0$ and $\phi(\,\cdot\,; \Sigma)$ is as defined in (3.20). This is indeed the same Multivariate Birnbaum-Saunders distribution with Multivariate normal kernel as given in Section 3 of Kundu et al. (2013) with $\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, 2, .., p$.

The Autocovariance function of BS-MA sequence is derived similar to that of BS-AR case discussed in Section 3.3. From model (3.18), it is evident that the ACF of $\{X_t\}$, $\rho_X(k) = \frac{\theta}{1+\theta^2},\ k = 1$ and zero elsewhere.

Also, $(X_1, X_2, ..., X_T) \sim N_T(0, \Sigma)$ with correlation matrix Σ.

We now derive the ACF of $\{Y_t\}$ in terms of that of $\{X_t\}$ by using the relation in (3.11). Then the autocovariace of $\{Y_t\}$ is obtained as

$$Cov(Y_t, Y_{t+k}) = \begin{cases} \alpha^2\beta^2 I_1, & k = 1 \\ 0 & , k > 1 \end{cases} \tag{3.22}$$

and the ACF is obtained as

$$\rho_Y(k) = I_1 \Big/ \left(1 + \frac{5}{4}\alpha^2\right),\ for\ k = 1 \tag{3.23}$$

and zero elsewhere, where I_1 is same as the expression given by (3.15) with ρ_X replaced by $\frac{\theta}{1+\theta^2}$ in the expression for $a_{m,n}$.

The plots in Figure 3.2 compare the ACFs and PACFs of the Gaussian MA(1) and the corresponding BS sequences for selected values of α and θ.

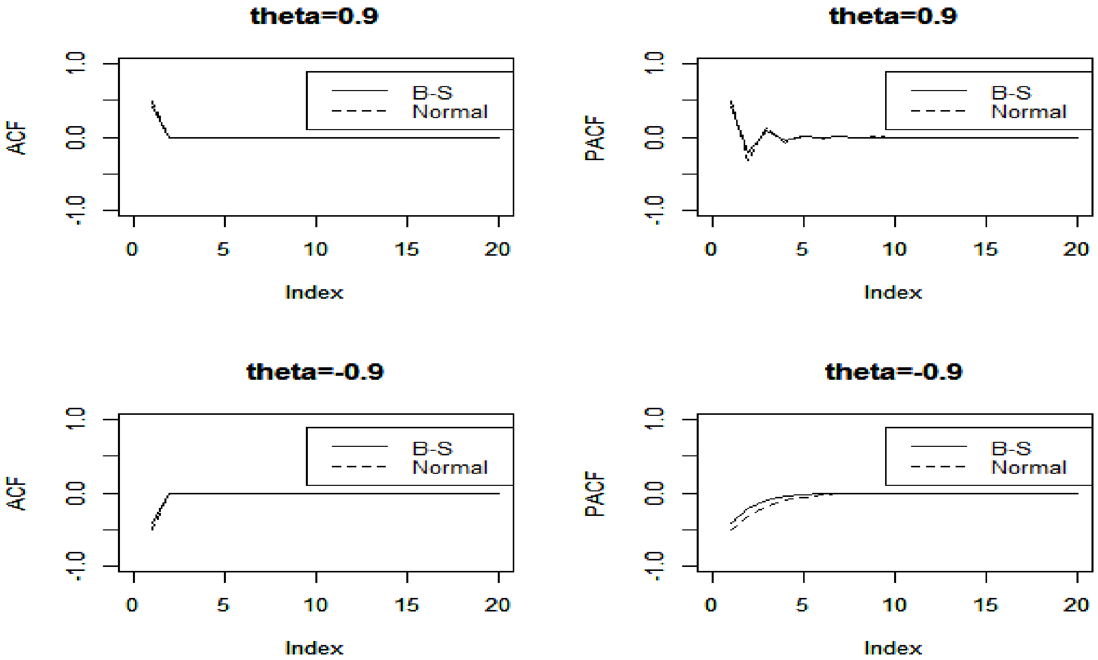

FIGURE 3.2: ACF and PACF of the Gaussian MA(1) and the corresponding BS-MA sequence.

3.5 BS-ARMA(1,1) Model

In this section we define a more flexible ARMA model with BS marginal distribution. To begin with consider an invertible and stationary ARMA(1,1) model

$$X_t = \rho X_{t-1} + \theta \eta_{t-1} + \eta_t, \quad t = 1, 2, ..., \tag{3.24}$$

such that the marginal distribution of $\{X_t\}$ is standard normal. We assume that $|\rho| < 1$ for stationarity, and $|\theta| < 1$ for invertibility. Under these constraints the distribution of η_t is $N(0,(1-\rho^2)/(1+\theta^2+2\theta\rho))$ with probability density function

$$\phi_\eta(x) = \frac{\sqrt{(1+\theta^2+2\theta\rho)}}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left\{-(1+\theta^2+2\theta\rho)x^2/2(1-\rho^2)\right\}, \; -\infty < x < \infty. \tag{3.25}$$

The stationary multivariate density function of $X_1, X_2, ..., X_T$ may be expressed as

$$\phi(x_1, x_2, ..., x_T) = \frac{1}{(2\pi)^{\frac{T}{2}} \, |\Gamma|^{\frac{1}{2}}} \; \exp\left\{-\frac{1}{2}X'\Gamma^{-1}X\right\}, \tag{3.26}$$

where $X = (x_1, x_2, ..., x_T)^{'}$ and Γ is the correlation matrix of X.

If we define Y_t in terms of X_t using the transformation (3.11) for $t = 1, 2, \ldots$ then $\{Y_t\}$ becomes a stationary sequence with finite dimensional distribution specified by the joint density function of $Y_1, Y_2, ..., Y_T$ for $T = 1, 2, \ldots$ as

$$f_Y(y_1, y_2, ..., y_T) = \phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{y_1}{\beta}} - \sqrt{\frac{\beta}{y_1}}\right), \ldots, \frac{1}{\alpha}\left(\sqrt{\frac{y_T}{\beta}} - \sqrt{\frac{\beta}{y_T}}\right); \Gamma\right) \times \prod_{n=1}^{T} \frac{1}{2\alpha\beta}\left\{\left(\frac{\beta}{y_n}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_n}\right)^{\frac{3}{2}}\right\} \tag{3.27}$$

for $y_1, y_2, ..., y_T > 0$ and $\phi(\,\cdot\,;\Gamma)$ is as defined in (3.26). We refer $\{Y_t\}$ as a BS-ARMA(1,1) sequence. In fact (3.27) is the same Multivariate Birnbaum-Saunders density with Multivariate normal kernel as given in Section 3 of Kundu et al. (2013) with $\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, 2, .., p$.

The autocovariance function of BS-ARMA(1,1) sequence is derived similar to that of BS-AR case discussed earlier. The ACF $R_X(k)$ of $\{X_t\}$ defined by (3.24) can derived as,

$$R_X(k) = \begin{cases} \frac{(\theta+\rho)(1+\rho\theta)}{(1+\theta^2+2\rho\theta)}, & k = 1 \\ \rho.R_X(k-1) & , \ k \geq 2 \end{cases}$$

Also, $(X_1, X_2, ..., X_T) \sim N_T\,(0, \Gamma)$ with correlation matrix

$$\Gamma \ = \ \begin{bmatrix} 1 & R & \rho R & \rho^2 R & \dots & \rho^{T-2}R \\ R & 1 & R & \rho^2 R & \dots & \rho^{T-3}R \\ \vdots & & & & & \\ \rho^{T-2}R & \rho^{T-3}R & \rho^{T-4}R & \dots & R & 1 \end{bmatrix},$$

where $R = \frac{(\theta+\rho)(1+\rho\theta)}{(1+\theta^2+2\rho\theta)}$.

Therefore, the autocovariance function of BS-ARMA (1,1) sequence is obtained as

$$Cov\,(Y_t, Y_{t+k}) = \alpha^2\beta^2\left[\frac{\alpha^2}{4}\left(3\rho^{2k} + \frac{(1+\theta^2)(1-\rho^{2k})}{(1+\theta^2+2\theta\rho)} - 1\right) + I_1\right],\ k = 1, 2, ... \tag{3.28}$$

and the ACF is obtained as

$$\rho_Y(k) = \left(I_1 + \frac{\alpha^2}{4}\left(3\rho^{2k} + \frac{(1+\theta^2)(1-\rho^{2k})}{(1+\theta^2+2\theta\rho)} - 1\right)\right) \Big/ \left(1 + \frac{5}{4}\alpha^2\right),\ k = 1.2, ... \tag{3.29}$$

where I_1 is same as the expression given by (3.15) with ρ_X replaced by R_X in the expression for $a_{m,n}$.

The plots in Figure 3.3 compare the ACFs and PACFs of the Gaussian ARMA(1,1) and the corresponding BS sequences for selected values of α, ρ and θ. Figures 3.1-3.3 suggest that the properties of theoretical ACF and PACF of proposed BS models are in line with the Gaussian ARMA models.

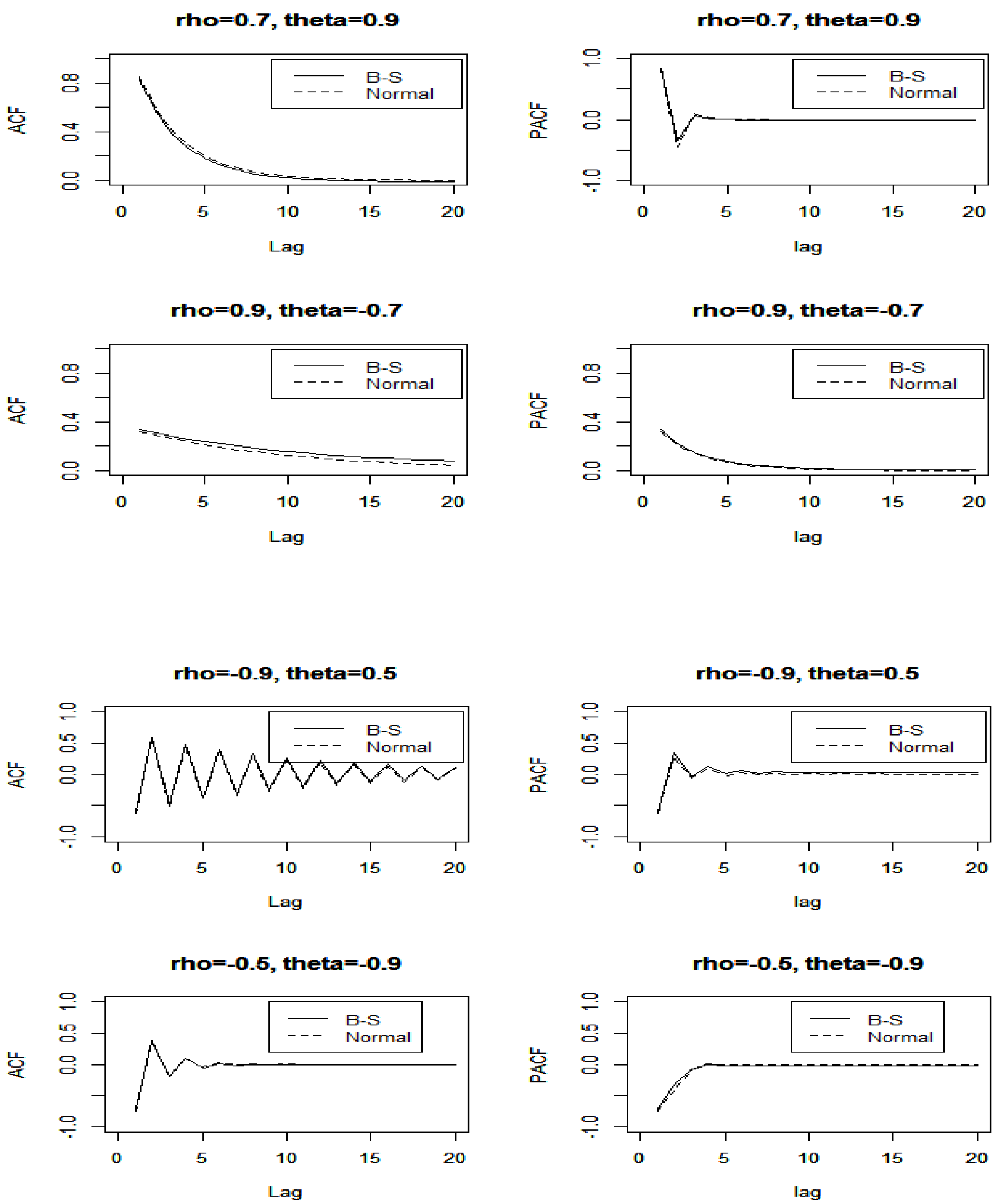

FIGURE 3.3: ACF and PACF of the Gaussian ARMA(1,1) and the corresponding BS-ARMA sequence

3.6 Estimation of Parameters

The maximum likelihood (ML) estimators based on a random sample from the BS distribution were discussed originally by Birnbaum and Saunders (1969a) and their asymptotic distributions were obtained by Engelhardt et al. (1981). However, while estimating the parameters we have to take care of the dependence structure among the observations. When we have explicit density function for the innovation random variables, the methods based on conditional likelihood function may be preferable. Hence we propose the method of conditional likelihood for estimating the BS-ARMA(1,1) model discussed earlier. Let the parameter vector to be estimated be $\Theta = (\alpha, \beta, \rho, \theta)$. Let us begin with the case of an AR(1) model and estimate $\Theta = (\alpha, \beta, \rho)$.

3.6.1 BS-AR(1) Model

An advantage of AR(1) process is that, it is Markovian when the errors are independent. For a Markov sequence with transition density function, the likelihood function based on a realization $(y_1, y_2, ..., y_T)$ conditional on y_1 can be expresses as

$$L(\Theta|y_1, y_2, ..., y_T) = \prod_{t=2}^{T} f_{t|t-1}(y_t|y_{t-1}),$$

where $f_{t|t-1}(.)$ is the one step transition density given by (3.13). The parameter vector to be estimated is $\Theta = (\alpha, \beta, \rho)$. Then, the log-likelihood function (without the additive constant) of the BS Markov sequence is given by

$$\log L = -(T-1)\log(\alpha\beta) - \left(\frac{T-1}{2}\right)\log\left(1-\rho^2\right) + \sum_{t=2}^{T}\log\left[\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right]$$
$$- \frac{1}{2\alpha^2\left(1-\rho^2\right)}\sum_{t=2}^{T}\left[\left(\sqrt{\frac{y_t}{\beta}} - \sqrt{\frac{\beta}{y_t}}\right) - \rho\left(\sqrt{\frac{y_{t-1}}{\beta}} - \sqrt{\frac{\beta}{y_{t-1}}}\right)\right]^2 \quad (3.30)$$

The ML estimators of ρ and α are, respectively, given by

$$\hat{\rho} = \frac{\sum_{t=2}^{T} w_t w_{t-1}}{\sum_{t=2}^{T} w_{t-1}^2}; \qquad \hat{\alpha} = \sqrt{\frac{\sum_{t=2}^{T}\left(w_t - \hat{\rho} w_{t-1}\right)^2}{(T-1)\left(1-\hat{\rho}^2\right)}}, \quad (3.31)$$

where

$$w_t = \sqrt{\frac{y_t}{\hat{\beta}}} - \sqrt{\frac{\hat{\beta}}{y_t}}\ , \quad w_{t-j} = \sqrt{\frac{y_{t-j}}{\hat{\beta}}} - \sqrt{\frac{\hat{\beta}}{y_{t-j}}}.$$

The ML estimator $\hat{\beta}$ of β can be obtained as a positive root of the equation

$$\sum_{t=2}^{T}\left(\frac{\beta - y_t}{\beta + y_t}\right) + \frac{1}{\alpha^2\left(1-\rho^2\right)}\sum_{t=2}^{T}\left(w_t - \rho w_{t-1}\right)\left(v_t - \rho v_{t-1}\right) = 0, \quad (3.32)$$

where $v_t = \sqrt{\frac{y_t}{\beta}} + \sqrt{\frac{\beta}{y_t}}$ *and* $v_{t-j} = \sqrt{\frac{y_{t-j}}{\beta}} + \sqrt{\frac{\beta}{y_{t-j}}}$.

We may use Newton-Raphson algorithm to obtain $\hat{\beta}$. Hence, by using numerical algorithm, we can solve equations (3.31) and (3.32). This algorithm allow us to have the joint iterative procedure:

$$\hat{\rho}^{(m+1)} = \frac{\sum_{t=2}^{T}\left(\sqrt{\frac{y_t}{\hat{\beta}^{(m)}}} - \sqrt{\frac{\hat{\beta}^{(m)}}{y_t}}\right)\left(\sqrt{\frac{y_{t-1}}{\hat{\beta}^{(m)}}} - \sqrt{\frac{\hat{\beta}^{(m)}}{y_{t-1}}}\right)}{\sum_{t=2}^{T}\left(\sqrt{\frac{y_{t-1}}{\hat{\beta}^{(m)}}} - \sqrt{\frac{\hat{\beta}^{(m)}}{y_{t-1}}}\right)^2}\ , \ m = 0, 1, 2, ...$$

$$\hat{\alpha}^{(m+1)} = \sqrt{\frac{\sum_{t=2}^{T}\left(\left(\sqrt{\frac{y_t}{\hat{\beta}^{(m)}}} - \sqrt{\frac{\hat{\beta}^{(m)}}{y_t}}\right) - \hat{\rho}^{(m)}\left(\sqrt{\frac{y_{t-1}}{\hat{\beta}^{(m)}}} - \sqrt{\frac{\hat{\beta}^{(m)}}{y_{t-1}}}\right)\right)^2}{(T-1)\left(1-(\hat{\rho}^{(m)})^2\right)}},$$

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} - \frac{f\left(\hat{\beta}^{(m)}\right)}{f'\left(\hat{\beta}^{(m)}\right)},$$

where $f(\beta)$ is as given in (3.32) and $f'(\beta)$ is its first derivative with respect to β. This iterative procedure needs starting values $\hat{\alpha}^{(0)}$, $\hat{\beta}^{(0)}$ and $\hat{\rho}^{(0)}$. Finding a proper initial value become quite important in this case and this will be discussed later in this section.

Next, we discuss the estimation of BS-MA(1) model.

3.6.2 BS-MA(1) Model

The parameter vector to be estimated is $\Theta = (\alpha, \beta, \theta)$. The conditional log-likelihood function (without the additive constant) corresponding to the the BS-MA sequence is given by

$$\log L = -\frac{1}{2}\log|\Sigma| - \frac{1}{2}V'\Sigma^{-1}V - T\log\alpha - T\log\beta + \sum_{t=1}^{T}\log\left\{\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right\}, \tag{3.33}$$

where

$$V' = \left[\frac{1}{\alpha}\left(\sqrt{\frac{y_1}{\beta}} - \sqrt{\frac{\beta}{y_1}}\right), \ldots, \frac{1}{\alpha}\left(\sqrt{\frac{y_T}{\beta}} - \sqrt{\frac{\beta}{y_T}}\right)\right].$$

Then the ML estimators of the unknown parameters can be obtained by maximizing (3.33) with respect to the parameters α, β and Σ. We observe that

$$\left[\left(\sqrt{\frac{y_1}{\beta}}-\sqrt{\frac{\beta}{y_1}}\right),\ldots,\left(\sqrt{\frac{y_T}{\beta}}-\sqrt{\frac{\beta}{y_T}}\right)\right]' \sim N(0,\, D\Sigma D'), \tag{3.34}$$

where D is the diagonal matrix given by $D = diag\{\alpha, \alpha, ..., \alpha\}$. Therefore, for given β, the ML estimators of α and Σ become

$$\hat{\alpha} = \left[\frac{1}{T}\sum_{t=1}^{T}\left(\sqrt{\frac{y_t}{\hat{\beta}}}-\sqrt{\frac{\hat{\beta}}{y_t}}\right)^2\right]^{\frac{1}{2}}; \qquad \hat{\Sigma} = PQP'; \tag{3.35}$$

here, P is a diagonal matrix given by $P = diag\{1/\hat{\alpha}, ..., 1/\hat{\alpha}\}$ and the elements of the matrix Q are given by

$$q_{ij} = \begin{cases} 1 & , \quad i=j \\ \frac{1}{T}\sum_{i=2}^{T}\left(\sqrt{\frac{y_i}{\beta}}-\sqrt{\frac{\beta}{y_i}}\right)\left(\sqrt{\frac{y_{i-1}}{\beta}}-\sqrt{\frac{\beta}{y_{i-1}}}\right), & |i-j|=1 \\ 0 & , \quad |i-j|>1 \end{cases}$$

for $i,\, j = 1, 2, ..., T$.

Finally, the ML estimators of β can be obtained by maximizing the profile log-likelihood function

$$l(\hat{\alpha},\beta,\hat{\Sigma}) = -\frac{1}{2}\log\left|\hat{\Sigma}\right| - \frac{1}{2}V'\hat{\Sigma}^{-1}V - T\log\hat{\alpha} - T\log\beta + \sum_{t=1}^{T}\log\left\{\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right\} \tag{3.36}$$

with respect to β by using Newton-Raphson numerical maximization algorithm. As in the BS-AR(1) case, the ML estimates of $\Theta = (\alpha, \beta, \theta)$ can be obtained by a numerical iterative procedure.

3.6.3 BS-ARMA(1,1) Model

Let $(y_1, y_2, ..., y_T)$ be a realization from the stationary and invertible BS-ARMA(1,1) model. We propose to obtain the conditional ML estimator of $\Theta = (\alpha, \beta, \rho, \theta)$ based on this realization. Based on the joint density function given in (3.27), log-likelihood function (without the additive constant) corresponding to the the BS-ARMA sequence is given by

$$\log L = -\frac{1}{2}\log|\Gamma| - \frac{1}{2}V'\Gamma^{-1}V - T\log\alpha - T\log\beta + \sum_{t=1}^{T}\log\left\{\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right\}, \tag{3.37}$$

where $V' = \left[\frac{1}{\alpha}\left(\sqrt{\frac{y_1}{\beta}} - \sqrt{\frac{\beta}{y_1}}\right), ..., \frac{1}{\alpha}\left(\sqrt{\frac{y_T}{\beta}} - \sqrt{\frac{\beta}{y_T}}\right)\right]$.

Then the ML estimators of the unknown parameters can be obtained by maximizing (3.37) with respect to the parameters α, β and Γ. We observe that

$$\left[\left(\sqrt{\frac{y_1}{\beta}} - \sqrt{\frac{\beta}{y_1}}\right), ..., \left(\sqrt{\frac{y_T}{\beta}} - \sqrt{\frac{\beta}{y_T}}\right)\right]' \sim N(0,\, D\Gamma D'), \tag{3.38}$$

where D is the diagonal matrix given by $D = diag\{\alpha, \alpha, ..., \alpha\}$. Therefore, for given β, the ML estimators of α and Γ become

$$\hat{\alpha} = \left[\frac{1}{T}\sum_{t=1}^{T}\left(\sqrt{\frac{y_t}{\hat{\beta}}} - \sqrt{\frac{\hat{\beta}}{y_t}}\right)^2\right]^{\frac{1}{2}}; \qquad \hat{\Gamma} = PGP' \tag{3.39}$$

here, P is a diagonal matrix given by $P = diag\{1/\hat{\alpha}, ..., 1/\hat{\alpha}\}$ and the elements of the matrix G are given by

$$g_{ij} = \begin{cases} 1 & , \quad i = j \\ \frac{1}{T}\sum_{i=2}^{T}\left(\sqrt{\frac{y_i}{\beta}} - \sqrt{\frac{\beta}{y_i}}\right)\left(\sqrt{\frac{y_{i-h}}{\beta}} - \sqrt{\frac{\beta}{y_{i-h}}}\right), & |i-j| \geq 1, \; h = 1, 2, ..., (T-1) \end{cases}$$

for $i, \, j = 1, 2, ..., T$.

Finally, the ML estimator of β can be obtained by maximizing the profile log-likelihood function

$$l(\hat{\alpha}, \beta, \hat{\Gamma}) = -\frac{1}{2}\log\left|\hat{\Gamma}\right| - \frac{1}{2}V'\hat{\Gamma}^{-1}V - T\log\hat{\alpha} - T\log\beta + \sum_{t=1}^{T}\log\left\{\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right\} \tag{3.40}$$

with respect to β by using Newton-Raphson numerical algorithm. For computing ML estimates, initial values are required which we discuss next.

Determination of initial values:

Since ML estimators do not have closed-form expressions and need to be obtained by solving non-linear equations, we propose the following modified moment estimators for the unknown parameters by following the approach of Ng et al. (2003) and Kundu et al. (2010). The modified moment estimators can be obtained by equating the moments and inverse moments with the corresponding sample quantities. Let $(y_1, y_2, ..., y_T)$ be a realization from the stationary and invertible BS-ARMA(1,1) model. The sample arithmetic and harmonic means are defined by

$$s_a = \frac{1}{T}\sum_{i=1}^{T} y_i\,, \qquad s_h = \left[\frac{1}{T}\sum_{i=1}^{T} y_i^{-1}\right]^{-1}.$$

As a preliminary analysis, we obtain the modified moment estimators of α and β

$$\tilde{\alpha} = \left\{ 2\left[\left(\frac{s_a}{s_h} \right)^{\frac{1}{2}} - 1 \right] \right\}^{\frac{1}{2}}, \tilde{\beta} = (s_a s_h)^{\frac{1}{2}}.$$

Then, the estimator of ρ is $\tilde{\rho} = r_2/r_1$, where

$$r_h = \frac{\sum_{i=1}^{T} \left(\sqrt{\frac{y_i}{\tilde{\beta}}} - \sqrt{\frac{\tilde{\beta}}{y_i}} \right) \left(\sqrt{\frac{y_{i-h}}{\tilde{\beta}}} - \sqrt{\frac{\tilde{\beta}}{y_{i-1}}} \right)}{\sqrt{\sum_{i=1}^{T} \left(\sqrt{\frac{y_i}{\tilde{\beta}}} - \sqrt{\frac{\tilde{\beta}}{y_i}} \right)^2} \sqrt{\sum_{i=1}^{T} \left(\sqrt{\frac{y_{i-1}}{\tilde{\beta}}} - \sqrt{\frac{\tilde{\beta}}{y_{i-h}}} \right)^2}}, \; h = 1, 2.$$

Finally, solve for $\tilde{\theta}$ in

$$r_1 = \frac{(\theta + \tilde{\rho})(1 + \tilde{\rho}\theta)}{(1 + \theta^2 + 2\tilde{\rho}\theta)}.$$

These estimators can be used effectively as the initial guess in the iterative procedure for computing the ML estimators. The computations are illustrated in Section 3.8

3.7 Asymptotic Properties of the estimators for BS-AR(1) Model

The BS-ARMA sequence $\{Y_t\}$ is generated by a non-linear model and establishing the asymptotic properties of the ML estimators is a challenging problem. However, when $\theta = 0$, $\{Y_t\}$ becomes a stationary Markov sequence as discussed in Section 3.3. Billingsley (1961) has proved the consistency and asymptotic normality of a ML estimator for Markov sequences under some regularity conditions which involve the second and third derivatives of the transition densities. As the likelihood function in our case is differentiable with respect to parameters and all the moments of BS-AR(1) sequence $\{Y_t\}$ are finite, the required regularity conditions hold here. We state the following theorem, whose proof is similar to that of Theorem 2.1 of Billingsley (1961).

Theorem 3.1. *Let $\{Y_t\}$ be the stationary BS-AR(1) sequence defined in Section 3.3 and $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\rho})$ be the ML estimator of $\Theta = (\alpha, \beta, \rho)$ obtained in Section 3.5. Then as $T \to \infty$,*

$$\sqrt{T}\left(\hat{\Theta} - \Theta\right) \xrightarrow{d} N_3\left(0, I^{-1}\right). \tag{3.41}$$

where, $N_3\left(0, I^{-1}\right)$ denotes the trivariate normal distribution with mean vector 0 and covariance matrix I^{-1}, and the matrix I is the Fisher information matrix.

The following computations are useful for evaluating the elements of the Fisher information matrix. Recall that

$$Y_t = \beta\left[\frac{\alpha X_t}{2} + \sqrt{\left(\frac{\alpha X_t}{2}\right)^2 + 1}\right]^2 \quad for\ t = 1, 2, \ldots,$$

and so we can write

$$X_t = \frac{1}{\alpha}\left(\sqrt{\frac{Y_t}{\beta}} - \sqrt{\frac{\beta}{Y_t}}\right).$$

It can be easily shown that

$$E\left[\left(\sqrt{\frac{Y_t}{\beta}} - \sqrt{\frac{\beta}{Y_t}}\right)\left(\sqrt{\frac{Y_{t-1}}{\beta}} - \sqrt{\frac{\beta}{Y_{t-1}}}\right)\right] = \alpha^2\rho,$$

$$E\left(\sqrt{\frac{Y_t}{\beta}}-\sqrt{\frac{\beta}{Y_t}}\right)^2 = E\left(\sqrt{\frac{Y_{t-1}}{\beta}}-\sqrt{\frac{\beta}{Y_{t-1}}}\right)^2=\alpha^2.$$

Now, we derive the expression for $E\left(\sqrt{Y_tY_{t-1}}\right)$ and $E\left(\frac{1}{\sqrt{Y_tY_{t-1}}}\right)$.

$$\begin{aligned}E\left(\sqrt{Y_tY_{t-1}}\right) &= E\left\{\beta\left[\frac{\alpha X_t}{2}+\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right]\left[\frac{\alpha X_{t-1}}{2}+\sqrt{\left(\frac{\alpha X_{t-1}}{2}\right)^2+1}\right]\right\}\\ &= \beta E\left\{\frac{\alpha^2 X_tX_{t-1}}{4}+\frac{\alpha X_t}{2}\sqrt{\left(\frac{\alpha X_{t-1}}{2}\right)^2+1}+\frac{\alpha X_{t-1}}{2}\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right.\\ &\qquad\left.+\left(\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right)\left(\sqrt{\left(\frac{\alpha X_{t-1}}{2}\right)^2+1}\right)\right\}\end{aligned}$$

So, we have

$$E\left(\sqrt{Y_tY_{t-1}}\right)=\beta\left[\frac{\alpha^2\rho}{4}+I_2\right],$$

where $I_2 = E\left[\left(\sqrt{\left(\frac{\alpha X_t}{2}\right)^2+1}\right)\left(\sqrt{\left(\frac{\alpha X_{t-1}}{2}\right)^2+1}\right)\right]$.

Also, the joint distribution of (Y_{t-1}, Y_t) and $\left(\frac{\beta^2}{Y_{t-1}},\frac{\beta^2}{Y_t}\right)$ are same and so we have $E\left(Y_tY_{t-1}\right)=E\left(\frac{\beta^4}{Y_tY_{t-1}}\right)$ and $E\left(\sqrt{Y_tY_{t-1}}\right)=E\left(\frac{\beta^2}{\sqrt{Y_tY_{t-1}}}\right)$, readily implies

$$E\left(\frac{1}{\sqrt{Y_tY_{t-1}}}\right)=\frac{1}{\beta}\left[\frac{\alpha^2\rho}{4}+I_2\right].$$

where the expression of I_2 is obtained similar to that of I_1 in Section 3.3 and is given by

$$I_2 = 1+\frac{1}{2^2}\alpha^2+\frac{1}{2^6}\alpha^4\left(1+2\rho^2\right)+\sum_{i=2}^{n}(-1)^{i-1}\frac{1\cdot 3\ldots(2i-3)}{2^{3i}\,i!}\alpha^{2i+2j}b_{0,i}$$

$$+\sum_{i=2}^{n}(-1)^{i-1}\frac{1\cdot 3\ldots(2i-3)}{2^{3i+3}\,i!}\alpha^{4i+4}b_{1,i}$$

$$+\sum_{i=2}^{n}\sum_{j=2}^{n}(-1)^{i+j}\frac{1\cdot 3\ldots(2i-3)}{2^{3i}\,i!}\times\frac{1\cdot 3\ldots(2j-3)}{2^{3j}\,j!}\alpha^{2i+2j}b_{i,j}$$

and for integers m,n ,

$$b_{m,n}=E\left(X_t^{2m}X_{t+h}^{2n}\right)=\frac{(2m)!(2n)!}{2^{m+n}}\sum_{i=0}^{\min(m,n)}\frac{\left(2\rho^h\right)^{2i}}{(m-i)!\,(n-i)!\,(2i)!}.$$

We have, from (3.13),

$$\log f=-\log\alpha-\log\beta-\frac{1}{2}\log\left(1-\rho^2\right)+\log\left[\left(\frac{\beta}{y_t}\right)^{\frac{1}{2}}+\left(\frac{\beta}{y_t}\right)^{\frac{3}{2}}\right]$$
$$-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{1}{\alpha}\left(\sqrt{\frac{y_t}{\beta}}-\sqrt{\frac{\beta}{y_t}}\right)-\frac{\rho}{\alpha}\left(\sqrt{\frac{y_{t-1}}{\beta}}-\sqrt{\frac{\beta}{y_{t-1}}}\right)\right]^2.$$

Then, the Fisher information matrix is $I(\theta)=((I_{ij}(\theta)))$, where $I_{ij}(\theta)=-E\left(\frac{\partial^2\log f}{\partial\theta_i\partial\theta_j}\right)$ and $\theta=(\theta_1,\theta_2,\theta_3)=(\alpha,\beta,\rho)$. We have obtained the exact expression of $I_{ij}(\theta)$, for $i,j=1,2,3$, and these are obtained as follows:

$$I_{11}=-E\left(\frac{\partial^2\log f}{\partial\alpha^2}\right)=\frac{2}{\alpha^2}\qquad I_{12}=I_{21}=-E\left(\frac{\partial^2\log f}{\partial\alpha\partial\beta}\right)=0,$$
$$I_{13}=I_{31}=-E\left(\frac{\partial^2\log f}{\partial\alpha\partial\rho}\right)=-\frac{\rho}{\alpha\left(1-\rho^2\right)},$$
$$I_{22}=-E\left(\frac{\partial^2\log f}{\partial\beta^2}\right)$$

$$=\frac{1}{\beta^2}\left\{-\frac{1}{2}+\psi\left(\alpha\right)+\frac{1}{\alpha^2\left(1-\rho^2\right)}\left[\left(1+\frac{1}{2}\alpha^2\right)-2\rho\left(\frac{\alpha^2\rho}{4}+I_2\right)+\rho^2\left(1+\frac{1}{2}\alpha^2\right)\right]\right\},$$

where $\psi\left(\alpha\right)=\int_{-\infty}^{\infty}\left(1+g\left(\alpha x_t\right)\right)^{-2}\phi\left(x_t\right)dx_t$, $g\left(u\right)=1+\frac{1}{2}u^2+u\left(1+\frac{u^2}{4}\right)^{\frac{1}{2}}$, and ϕ is the standard normal probability density function. Further, we also have

$I_{23} = I_{32} = -E\left(\frac{\partial^2 \log f}{\partial\rho\partial\beta}\right) = 0$ and $I_{33} = -E\left(\frac{\partial^2 \log f}{\partial\rho^2}\right) = \frac{1+\rho^2}{(1-\rho^2)^2}$.

3.8 Simulation Study

In this section, we illustrate the performance of the proposed estimators using Monte Carlo simulations. We present the following algorithm to generate $\{Y_t\}$ from BS-ARMA(1,1) model:

Step 1: Set values for $\alpha,\ \beta,\ \rho$ and θ;

Step 2: Set $\eta_0 = x_0 = 0$;

Step 3: Generate a random sample $\{\eta_t\}$ from $\eta \sim N\left(0, (1-\rho^2)/(1+\theta^2+2\theta\rho)\right)$, $t = 1, 2, ...$;

Step 4: Generate x_t from $X_t = \rho X_{t-1} + \theta\eta_{t-1} + \eta_t,\quad t = 1, 2, ...$;

Step 5: Compute y_t using the transformation $Y_t = \beta\left[\frac{1}{2}\alpha X_t + \sqrt{\left(\frac{1}{2}\alpha X_t\right)^2 + 1}\right]^2$ for $t = 1, 2, ...$;

Step 6: Repeat Steps 1 to 5 to get a sample of required size.

We generated samples of sizes n =500 and 1000 from BS-AR, BS-MA and BS-ARMA models, with $\alpha = 2,\ \beta = 1$ and for different choices of ρ and θ. The ML estimates were obtained based on the algorithm developed in Section 3.6 using modified moment estimates as the initial values. For each parameter combination, ML estimates were obtained and we then computed the average and the mean squared errors (MSE) over 100 replications. The resulting estimates and the corresponding MSEs (within parentheses) are reported in Tables 3.1-3.3.

n	ρ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
500	0.90	1.9698(0.1304)	1.0254(0.0864)	0.8945(0.0190)
	0.70	2.0082(0.1357)	1.0412(0.0676)	0.7033(0.0374)
	0.50	2.0172(0.0896)	0.9998(0.0966)	0.4910(0.0495)
	0.25	1.9914(0.0754)	1.0116(0.0690)	0.2412(0.0581)
	0.00	1.9938(0.0531)	0.9974(0.0870)	-0.0065(0.0368)
	-0.25	2.0057(0.0505)	1.0257(0.0593)	-0.2527(0.0385)
	-0.50	2.0181(0.0796)	1.0055(0.0454)	-0.5046(0.0382)
	-0.70	2.0014(0.0732)	0.9951(0.0296)	-0.6978(0.0305)
	-0.90	2.0280(0.1046)	1.0009(0.0117)	-0.8994(0.0218)
1000	0.90	2.0039(0.0968)	0.9980(0.0810)	0.9042(0.0146)
	0.70	1.9982(0.0478)	1.0126(0.0513)	0.6978(0.0179)
	0.50	2.0140(0.0491)	0.9961(0.0680)	0.4973(0.0226)
	0.25	2.0108(0.0574)	1.0067(0.0550)	0.2420(0.0231)
	0.00	2.0166(0.0427)	0.9972(0.0374)	0.0035(0.0290)
	-0.25	1.9918(0.0412)	1.0063(0.0345)	-0.2508(0.0281)
	-0.50	1.9968(0.0688)	1.0088(0.0281)	-0.4926(0.0336)
	-0.70	2.0060(0.0577)	0.9933(0.0151)	-0.6996(0.0195)
	-0.90	2.0072(0.0767)	0.9979(0.0136)	-0.8989(0.0151)

TABLE 3.1: The average estimates and the corresponding mean square error for the MLEs of BS-AR(1), when $\alpha = 2$, $\beta = 1$ and for different ρ's

From the simulation results, it is clear that the performance of the ML estimates (in terms of bias and MSE) do not seem to depend on ρ and θ, and depends only on the sample size. We observe that the estimates are slightly biased for sample size n=500 and bias gets smaller when the sample size increases. Overall, the estimates behave quite well and they become more accurate with increasing sample size, as one would expect. Thus, we see that the method of conditional likelihood estimation yields

n	θ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
500	0.90	1.9792(0.1052)	0.9859(0.0778)	0.8878(0.1946)
	0.70	1.9876(0.0933)	0.9883(0.0698)	0.6766(0.1023)
	0.50	1.9885(0.0706)	1.0853(0.0957)	0.4794(0.0537)
	0.25	1.9811(0.0758)	1.0612(0.0482)	0.2235(0.0449)
	0.00	2.0073(0.0701)	0.9941(0.0536)	-0.0077(0.0522)
	-0.25	2.0011(0.0688)	0.9892(0.0367)	-0.2354(0.0464)
	-0.50	1.9818(0.0783)	1.0182(0.0445)	-0.5257(0.0825)
	-0.70	2.0592(0.0683)	0.9833(0.0452)	-0.7165(0.0449)
	-0.90	1.9809(0.1015)	1.0892(0.1068)	-0.8736(0.3056)
1000	0.90	1.9812(0.0916)	0.9871(0.0700)	0.8900(0.1464)
	0.70	1.9881(0.0850)	0.9899(0.0611)	0.6966(0.0927)
	0.50	1.9910(0.0616)	1.0605(0.0811)	0.4897(0.0462)
	0.25	1.9851(0.0651)	0.9912(0.0396)	0.2504(0.0411)
	0.00	2.0021(0.0641)	0.9965(0.0521)	0.0019(0.0498)
	-0.25	1.9952(0.0610)	0.9912(0.0301)	-0.2470(0.0390)
	-0.50	1.9901(0.0569)	1.0102(0.0395)	-0.5121(0.0610)
	-0.70	2.0120(0.0581)	0.9895(0.0359)	-0.7022(0.0391)
	-0.90	1.9932(0.0861)	1.0892(0.9102)	-0.9063(0.2213)

TABLE 3.2: The average estimates and the corresponding mean square error for the MLEs of BS-MA(1), when $\alpha = 2$, $\beta = 1$ and for different θ's

good estimates for the model parameters. The detailed computational algorithm is given in Appendix A.

We also obtain the confidence interval for $\Theta = (\alpha, \beta, \rho)$ and evaluated the coverage probabilities based on the asymptotic distribution of the ML estimators of BS-AR(1) model. Let $\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) = (\hat{\alpha}, \hat{\beta}, \hat{\rho})$ be the ML estimators determined as described earlier. For a given significance level τ, the $100(1-\tau)\%$ asymptotic confidence interval for θ_i is given by $\hat{\theta}_i \pm Z_{\tau/2}\,\sigma_{\hat{\theta}_i}(1/\sqrt{n})$, $i = 1, 2, 3$; where $\sigma_{\hat{\theta}_i}$ is the ith diagonal element of the asymptotic dispersion matrix I^{-1} given in (3.41) with the parameters replaced by the corresponding estimates; here, $Z_{\tau/2}$ is the upper $\tau/2$ percentage point of the standard normal distribution.

n	(ρ,θ)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\theta}$
500	(0.90, 0.00)	1.9812(0.1432)	0.9877(0.1186)	0.8874(0.0465)	0.0195(0.0621)
	(0.70, 0.25)	1.9892(0.1094)	1.0360(0.1126)	0.6982(0.0315)	0.2414(0.0590)
	(0.50, 0.50)	1.9957(0.1036)	1.0090(0.0921)	0.4856(0.0546)	0.5284(0.1070)
	(0.25, 0.70)	2.0080(0.1348)	1.1000(0.1272)	0.2457(0.0438)	0.6881(0.1257)
	(0.00, 0.90)	2.0178(0.0661)	0.9866(0.1040)	0.0586(0.0342)	0.8803(0.1133)
	(-0.25,-0.90)	2.0320(0.1165)	1.0050(0.0770)	-0.2617(0.0846)	-0.8817(0.1127)
	(-0.50,-0.70)	2.0150(0.0671)	0.9896(0.0816)	-0.5206(0.0498)	-0.6887(0.1111)
	(-0.70,-0.50)	1.9878(0.1320)	1.0151(0.0737)	-0.6127(0.0612)	-0.4869(0.0904)
	(-0.90,-0.25)	1.9789(0.1365)	1.0223(0.0785)	-0.8916(0.0561)	-0.2412(0.0954)
1000	(0.90, 0.00)	1.9879(0.1314)	0.9912(0.1186)	0.9096(0.0400)	0.0144(0.0566)
	(0.70, 0.25)	1.9910(0.1013)	1.0120(0.1126)	0.7011(0.0331)	0.2502(0.0414)
	(0.50, 0.50)	1.9951(0.0966)	1.0112(0.0921)	0.4965(0.0467)	0.5190(0.0954)
	(0.25, 0.70)	2.0282(0.1141)	1.0823(0.1272)	0.2520(0.0333)	0.6954(0.1214)
	(0.00, 0.90)	2.0322(0.0624)	0.9897(0.1040)	0.0416(0.0311)	0.9054(0.1069)
	(-0.25,-0.90)	2.0623(0.0922)	1.0171(0.0770)	-0.2519(0.0789)	-0.8962(0.0955)
	(-0.50,-0.70)	2.0261(0.0533)	0.9961(0.0816)	-0.5166(0.0415)	-0.6900(0.0891)
	(-0.70,-0.50)	1.9919(0.1129)	1.0011(0.0737)	-0.6074(0.0546)	-0.4958(0.0783)
	(-0.90,-0.25)	1.9859(0.1096)	1.0009(0.0785)	-0.8955(0.0575)	-0.2516(0.0812)

TABLE 3.3: The average estimates and the corresponding mean square error for the MLEs of BS-ARMA(1,1), when $\alpha = 2$, $\beta = 1$ and for different (ρ, θ) values.

n	ρ	$CP(\alpha)$	$CP(\beta)$	$CP(\rho)$
500	0.90	0.940	0.945	0.936
	0.70	0.933	0.964	0.937
	0.50	0.931	0.961	0.929
	0.25	0.954	0.964	0.965
	0.00	0.955	0.950	0.940
	-0.25	0.921	0.968	0.953
	-0.50	0.934	0.974	0.947
	-0.70	0.947	0.984	0.944
	-0.90	0.948	0.980	0.922
1000	0.90	0.956	0.971	0.964
	0.70	0.960	0.984	0.975
	0.50	0.963	0.960	0.973
	0.25	0.955	0.973	0.988
	0.00	0.976	0.966	0.965
	-0.25	0.982	0.955	0.968
	-0.50	0.984	0.984	0.972
	-0.70	0.986	0.966	0.980
	-0.90	0.980	0.993	0.981

TABLE 3.4: Estimated coverage probabilities ($CP(.)$) for selected values of $\alpha = 2, \beta = 1$ and different values of ρ

We determined the 95% confidence intervals based on the simulated samples of size n from the BS Markov sequences for specified values of the parameters. We repeated this computation for 100 samples and then determined the proportion of times these intervals contained the actual parameters. These proportions give us the estimated coverage probabilities which are presented in Table 3.4. The entries in the first column are the sample sizes and those in the second column are the specified values of ρ. The values in the third, fourth and fifth columns are the estimated coverage probabilities corresponding to the parameters α, β and ρ, respectively. From the table it is observed that the performances of coverage probabilities are quite satisfactory for large sample sizes.

3.8.1 Performance Analysis of the Models

In this subsection, we analyze the performance of the proposed BS ARMA model based on the Akaike Information Criteria (AIC), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) for the simulated data. The RMSE and MAPE are used to assess the forecasting ability of the model. This computation is particularly useful to observe the consequences of fitting a Gaussian ARMA model when the data is actually from a BS-ARMA model. Towards this objective, we generated samples of sizes n =500 and 1000 from BS-ARMA(1,1) model, with $\alpha = 2$, $\beta = 1$ and for different choices of ρ and θ, and then fitted the BS-ARMA(1,1) model vis-à-vis Gaussian ARMA(1,1) model using ML method. Then, we compare the BS and Gaussian models on the basis of some information criterion statistics particularly, the AIC, wherein the model with the smallest information criterion value implies the model with the highest maximized log-likelihood and hence the best fitting model. The Akaike information criterion values and the forecast evaluation statistics using the simulated data for Gaussian ARMA(1,1) and BS-ARMA(1,1) (in parentheses) models are given in Table 3.5. The AIC and the log-likelihood values highlight the fact that the BS-ARMA model fits better the data than Gaussian ARMA model. In order to check the forecast performance of the model, the forecasts are compared against actual observations and we then computed RMSE and MAPE for the BS and Gaussian models. These results are summarized in Table 3.5. The BS-ARMA model gives lower RMSE and MAPE (with in parentheses) values indicating better forecast capability.

n	(ρ, θ)	AIC	Log-like	RMSE	MAPE
500	(0.90, 0.00)	965.26	-479.63	0.2606	16.9690
		(817.18)	(-404.59)	(0.2570)	(7.1100)
	(0.70, 0.25)	1145.10	-569.55	0.3121	31.8820
		(517.14)	(-254.57)	(0.3085)	(10.3685)
	(0.50, 0.50)	1245.20	-619.60	0.3451	37.0936
		(445.72)	(-218.86)	(0.3439)	(12.5100)
	(0.25, 0.70)	1224.48	-609.24	0.3376	41.4342
		(406.14)	(-199.07)	(0.3291)	(13.6194)
	(0.00, 0.90)	1570.50	-782.25	0.4778	57.2870
		(333.82)	(-162.91)	(0.4277)	(14.4635)
	(-0.25,-0.90)	1694.60	-844.30	0.5412	71.7563
		(498.14)	(-245.07)	(0.3362)	(10.4250)
	(-0.50,-0.70)	1663.86	-828.93	0.5247	81.7054
		(491.46)	(-241.73)	(0.3310)	(11.5034)
	(-0.70,-0.50)	1590.58	-792.29	0.4876	77.4590
		(661.74)	(-326.87)	(0.3387)	(8.1411)
	(-0.90,-0.25)	1369.72	-681.86	0.3906	85.0058
		(1079.74)	(-535.87)	(0.2130)	(4.7527)
1000	(0.90, 0.00)	1933.00	-963.50	0.2619	16.1836
		(1690.80)	(-841.40)	(0.2599)	(6.6095)
	(0.70, 0.25)	2456.56	-1225.28	0.3404	25.6785
		(1931.00)	(-961.50)	(0.3387)	(10.9191)
	(0.50, 0.50)	2288.32	-1141.16	0.3129	35.6167
		(1800.50)	(-896.25)	(0.3097)	(11.9672)
	(0.25, 0.70)	2806.14	-1400.07	0.4054	50.9527
		(2004.50)	(-998.25)	(0.3937)	(13.2894)
	(0.00, 0.90)	2743.58	-1368.79	0.3930	54.9803
		(1978.46)	(-985.23)	(0.3710)	(14.1725)
	(-0.25,-0.90)	2939.90	-1466.95	0.4336	62.4163
		(1943.16)	(-967.58)	(0.3528)	(12.0020)
	(-0.50,-0.70)	3191.54	-1592.77	0.4917	75.5837
		(1943.12)	(-967.56)	(0.3521)	(9.6591)
	(-0.70,-0.50)	2725.34	-1359.67	0.3894	53.4320
		(1806.70)	(-899.35)	(0.2599)	(8.2026)
	(-0.90,-0.25)	2590.88	-1292.44	0.3640	70.1784
		(1798.50)	(-895.25)	(0.2134)	(4.8762)

Table 3.5: Forecast evaluation statistics for simulated data

3.9 Data Analysis

To illustrate the application of the proposed models and the associated inferential results, we analyse two monthly time series. The data sets used for this purpose are: (1) the index of coal production in Eight Core Industries (Base: 2004-05=100) obtained from the Office of Economic Adviser, Government of India, with the data consisting of 129 observations from April 2004 to December 2014; (2) The number of Foreign Tourist Arrivals in India obtained from Press Information Bureau, Government of India, with the data consisting of 154 observations from April 2002 to January 2015.

The time series plots of actual data series are given in the left panel of Figure 3.4. As we are interested in modelling a stationary non-negative series, we need to remove the trend and seasonality present in the data. Here, we use the multiplicative seasonal model to obtain seasonally adjusted data. A series is formed by dividing the original data with residuals of the seasonal model and then analysed by BS models. Let us denote the resulting series by $\{Y_t\}$. Table 3.6 presents descriptive

Statistics	Coal Index	Tourist Arrival
Mean	0.9998	0.9994
Median	0.9969	0.9974
Maximum	1.1021	1.2174
Minimum	0.7952	0.7943
Std. Dev.	0.0459	0.0644
Skewness	0.7169	0.2337
Kurtosis	5.4283	4.3317
Sample size	129	154

TABLE 3.6: Summary statistics for transformed data

statistics for $\{Y_t\}$ and its time series plots are given in the right panel of Figure 3.4.

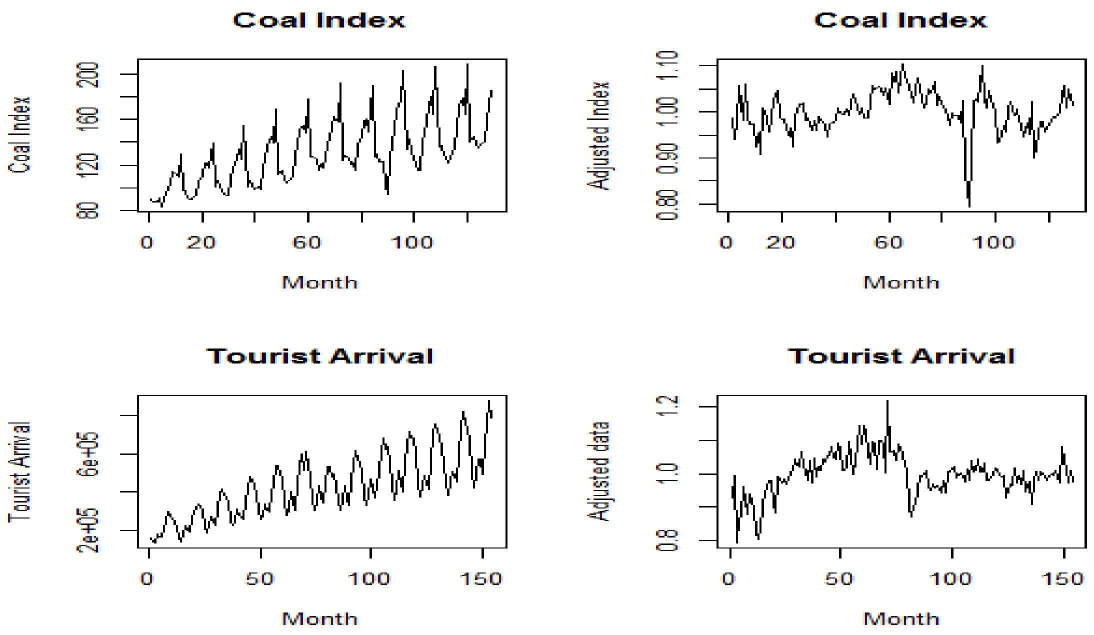

FIGURE 3.4: Time series plot of the actual series and adjusted series

Note that the kurtosis for both the series are greater than three, indicating that a heavy-tailed distribution is more appropriate than the normal distribution.

The sample ACF and PACF of $\{Y_t\}$ are plotted in Figure 3.5, which suggest that the coal index series may have an AR(1) structure and the Tourist Arrival series may be a realization of ARMA(1,1) process. Next, we obtained the modified moment estimates of the parameters of the proposed models for the above data sets and obtained $\tilde{\alpha}$=0.0411, $\tilde{\beta}$=0.9858, $\tilde{\rho}$=0.5821 for BS-AR(1), and $\tilde{\alpha}$=0.0502, $\tilde{\beta}$=0.9917, $\tilde{\rho}$=0.8861, $\tilde{\theta}$=-0.4253 for BS-ARMA(1,1) models. We then estimated the parameters of BS-AR(1) and BS-ARMA(1,1) models by the method of maximum likelihood

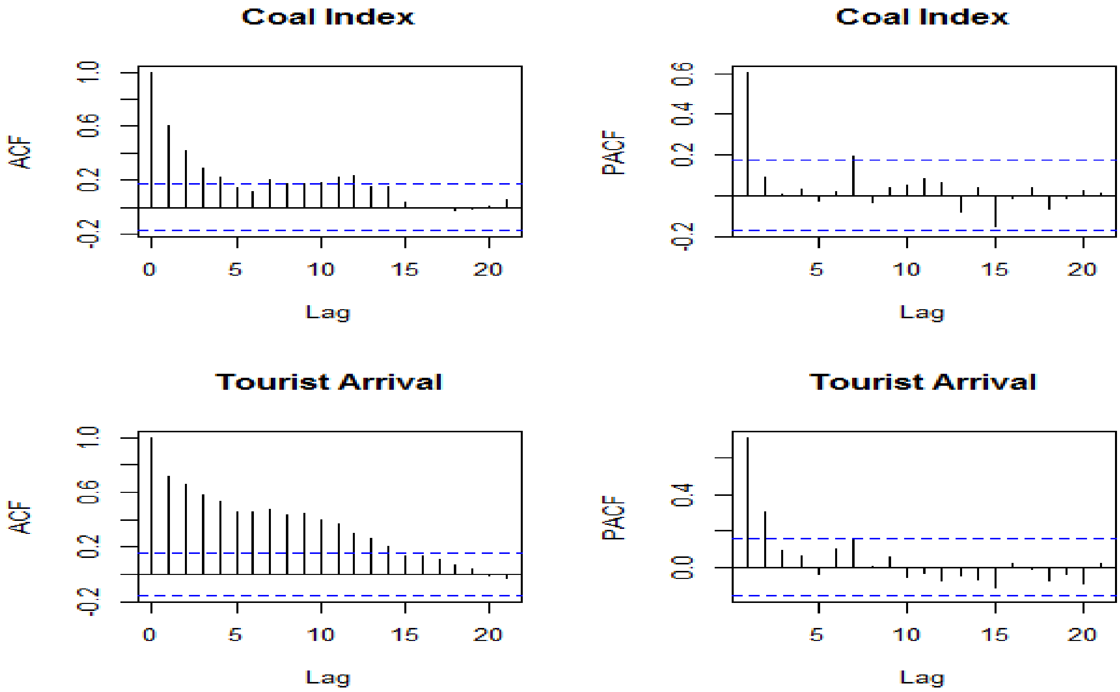

FIGURE 3.5: ACF and PACF of seasonally adjusted and de-trended series

Coal Index	α	β	ρ	θ	σ
BS-AR(1)	0.0469 (0.3015)	0.9990 (0.2261)	0.6031 (0.3368)	–	–
Gaussian AR(1)	–	–	0.5856 (0.0326)	–	0.9994 (0.0419)
Tourist Arrival	α	β	ρ	θ	σ
BS-ARMA(1,1)	0.0535 (0.3622)	1.0004 (0.3353)	0.9080 (0.1197)	0.4310 (0.2797)	–
Gaussian ARMA(1,1)	–	–	0.8819 (0.0455)	0.4211 (0.0568)	0.9968 (0.0658)

TABLE 3.7: Parameter Estimates using ML methods

detailed earlier in Section 3.6 and the obtained results are summarized in Table 3.7. The standard errors of the ML estimates are given inside parentheses. In this case, we see that the modified moment estimates are quite close to the ML estimates.

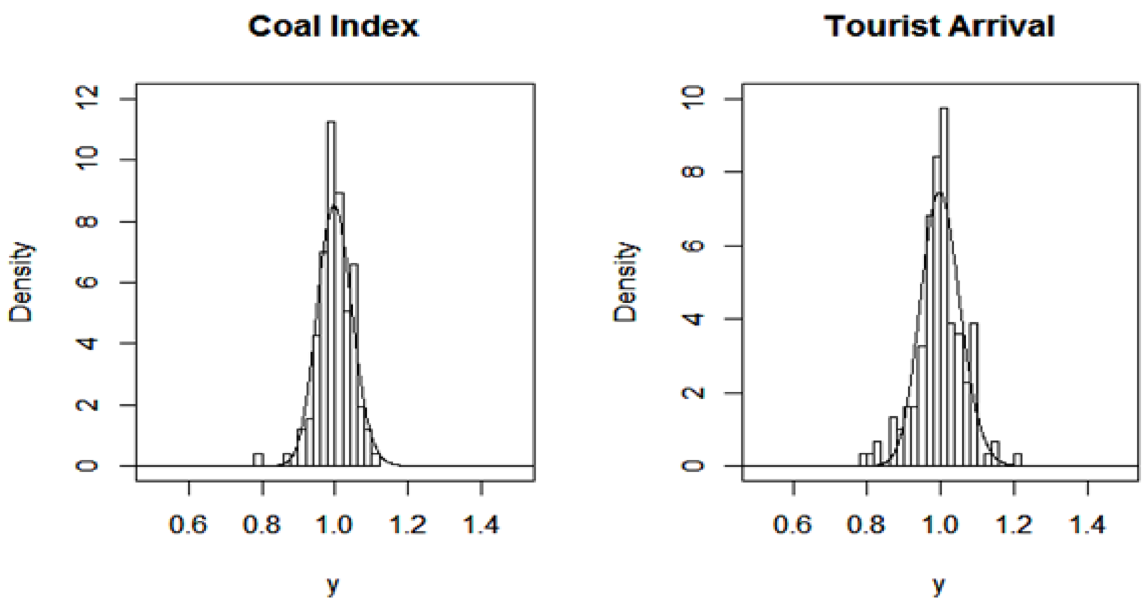

FIGURE 3.6: Histograms of the data with superimposed BS density

In Figure 3.6, we superimpose the histogram of the $\{Y_t\}$ with the BS density curve whose parameters are as given in Table 3.7 to check whether the series follows the BS distribution. These histograms do show that there is a close agreement between the observed and the fitted density functions.

Further, we also fitted Gaussian AR(1) for coal index and Gaussian ARMA(1,1) model for tourist arrival data and the ML estimates are given in Table 3.7. To perform the diagnostic checks on the residuals from the fitted models, the residuals and their ACF are plotted in Figure 3.7. It is observed that the ACF of the resulting residuals from fitted BS and Gaussian models are negligible.

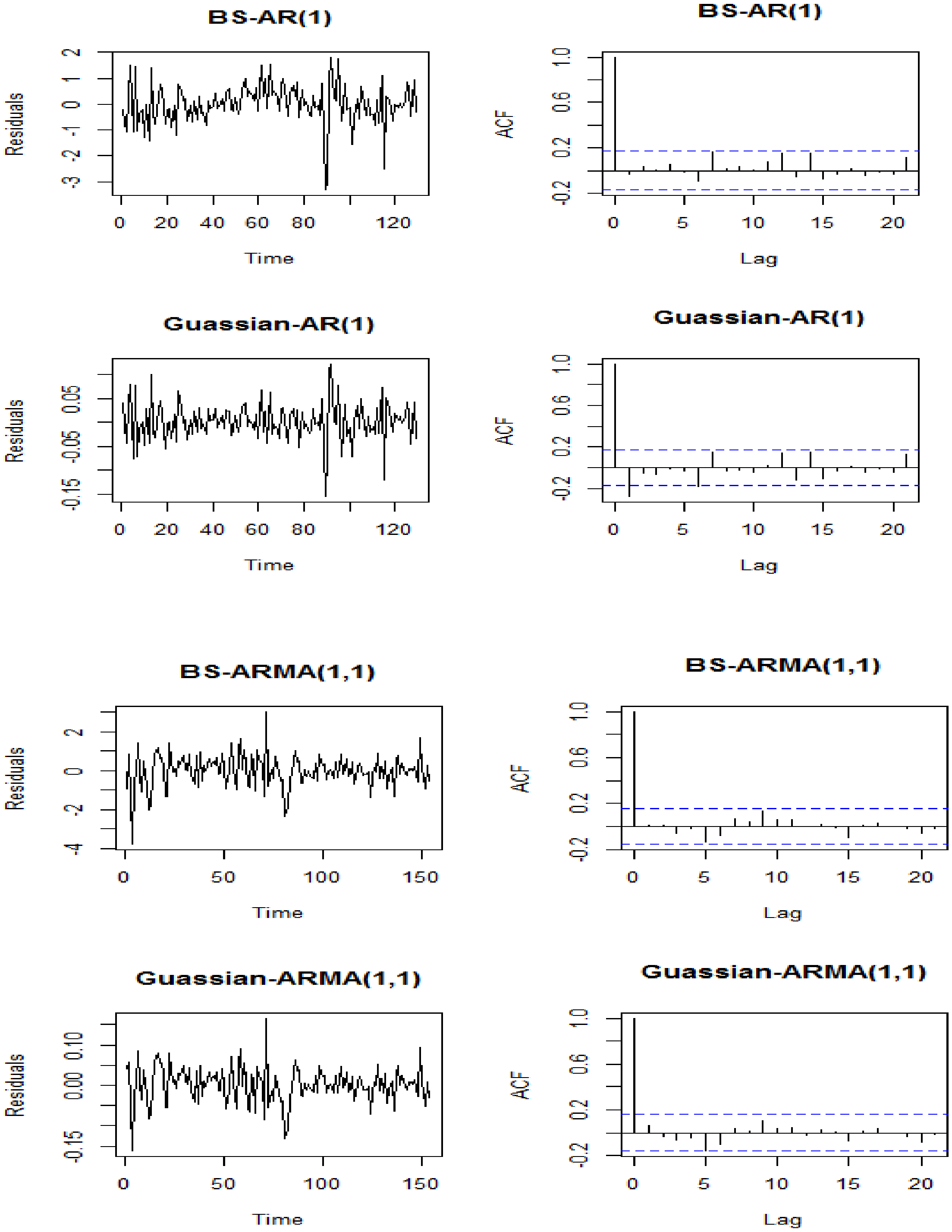

FIGURE 3.7: Time series and ACF plots of the residual series from fitted BS and Gaussian models

Coal Index	AIC	RMSE	MAPE
BS-AR(1)	2.0362	0.0253	1.9885
Gaussian AR(1)	2.1235	0.0698	4.3077
Tourist Arrival	AIC	RMSE	MAPE
BS-ARMA(1,1)	2.4595	0.0339	3.6453
Gaussian ARMA(1,1)	2.5463	0.0766	4.5714

Table 3.8: Model evaluation statistics

We therefore compare the BS and Gaussian models on the basis of their AIC values. The given AIC values in Table 3.8 highlight the fact that the proposed BS models fit better for both data sets than the Gaussian time series models.

Next, we evaluate these models based on forecast performance. We performed the out-sample forecast exercise for coal index and tourist arrival data using the BS and Gaussian ARMA models. For coal index data, we first use the sample period from April 2004 to June 2014 and evaluated one-step ahead forecast for the month July 2014 using BS-AR(1) and Gaussian AR(1) model. Next, we use the sample period April 2004 to July 2014 to predict coal index in August 2014. We repeat this rolling forecast until we obtain six forecasted values. Similarly, we forecasted tourist arrivals using BS-ARMA(1,1) and Gaussian ARMA(1,1) models for six months. Table 3.8 shows the RMSE and MAPE of forecasts for the competing BS and Gaussian ARMA models. The proposed BS-AR(1) and BS-ARMA(1,1) models clearly outperform the Gaussian ARMA models based on both RMSE and MAPE.

The results of this chapter are reported in the paper Rahul *et al.* (2017a).

Chapter 4

Modelling Stochastic Volatility using Birnbaum-Saunders Markov Sequences

4.1 Introduction

The SV model introduced by Taylor (1986) is used to account for the well-documented autoregressive behaviour in the volatility of financial time series. The literature in this area mainly deals with the models with normal-lognormal distributions. An exponentiated Gaussian autoregressive sequence provides a Markov dependent sequence of log-normal random variables to describe the conditional variances, see Tsay (2005). Taylor (1994) suggested several alternative models to describe the evolution of conditional variances while modelling stochastic volatilities. As quoted by Shephard (1996), volatility models provide an excellent testing ground for the development of new non-linear and non-Gaussian time series techniques. A number of autoregressive models are introduced for non-negative random variables in the context of non-Gaussian time series modelling. In principle, one can very well use these autoregressive models to describe the evolution of time-dependent volatilities.

In this chapter, we study the properties of Birnbaum-Saunders Stochastic Volatility (BS-SV) model where the volatilities are generated by a stationary Markov sequence with BS marginal distribution discussed in Section 3.3. The BS distribution received great attention in recent years in the context of life time modelling. Different aspects of BS model have been studied including estimation, regression, diagnostics and applications by researchers. The BS distribution is typically applied to positive data with varying degrees of asymmetry and kurtosis and can be used as an alternative to the log-normal and log-skew-normal models. However, applications of this model for modelling volatility in financial time series context have not received much attention. Only the works by Jin and Kawczak (2003), Fox et al. (2008) and Bhatti (2010) have indirectly considered the use of this distribution in finance. This chapter is such an attempt to use the BS model in the context of financial time series to model stochastic volatility.

Rest of this chapter is organized as follows. The construction of BS-SV model and its second order properties are discussed in Section 4.2. In Section 4.3, we estimate the unknown parameters of the model by Method of Moments (MM) and Efficient Importance Sampling (EIS) method. Section 4.4 presents some numerical results of the estimators via simulation. In Section 4.5, we apply our model to two daily returns data.

4.2 BS-SV Model and Properties

Let r_t be the return at time t. Define the SV model

$$r_t = \sqrt{h_t}\varepsilon_t,$$
$$h_t = \beta\left[\frac{1}{2}\alpha X_t + \sqrt{\left(\frac{1}{2}\alpha X_t\right)^2 + 1}\right]^2,$$
$$X_t = \rho X_{t-1} + \eta_t \;\; ; \;\; |\rho| < 1,\; t = 1, 2, ..., \tag{4.1}$$

with $\{X_t\}$ be a stationary Gaussian AR(1) sequence with standard normal marginal distribution. Where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed standard normal random variables. We assume that the sequence $\{\varepsilon_t\}$ is independent of h_t and η_t for every t. This is a stochastic volatility model for the return series $\{r_t\}$ whose volatilities are generated by a stationary Markov sequence of BS random variables with marginal probability density function

$$f\left(h_t;\alpha,\beta\right) = \frac{1}{2\alpha\beta\sqrt{2\pi}}\left[\left(\frac{\beta}{h_t}\right)^{1/2} + \left(\frac{\beta}{h_t}\right)^{3/2}\right]\exp\left(-\frac{1}{2\alpha^2}\left[\frac{h_t}{\beta} + \frac{\beta}{h_t} - 2\right]\right), \tag{4.2}$$

where $h_t > 0$, $\alpha, \beta > 0$.

The r^{th} row moment about origin zero of $\{h_t\}$ is given by Rieck (1999) as

$$E\left(h_t^r\right) = \frac{\beta^r\left[K_{r+\frac{1}{2}}(\alpha^{-2}) + K_{r-\frac{1}{2}}(\alpha^{-2})\right]}{2K_{\frac{1}{2}}(\alpha^{-2})},$$

where $K_\nu(z)$ is the modified Bessel function of the third kind with ν representing its order and z the coefficient of argument. That is,

$$K_\nu(z) = \frac{1}{2}\int_{-\infty}^{\infty} \exp\left[-z\cosh(t) - \nu t\right] dt.$$

Since the sequence $\{\varepsilon_t\}$ follows standard normal distribution, the odd moments of $\{r_t\}$ are zero and its even moments are given by

$$E\left(r_t^{2r}\right) = \frac{\beta^r \left[K_{r+\frac{1}{2}}(\alpha^{-2}) + K_{r-\frac{1}{2}}(\alpha^{-2})\right]}{2K_{\frac{1}{2}}(\alpha^{-2})} \prod_{j=1}^{r}(2j-1), \; r = 1, 2, \ldots.$$

Then $Var\left(r_t\right) = \beta\left(1 + \frac{\alpha^2}{2}\right)$ and the kurtosis of r_t becomes

$$K = 3 + \frac{3\left(\alpha^2 + \frac{5}{4}\alpha^4\right)}{\left(1 + \alpha^2 + \frac{1}{4}\alpha^4\right)} > 3.$$

The structure of the model (4.1) implies that the ACF of $\{r_t\}$ is zero and that of $\{r_t^2\}$ is significant. The variance of the squared return series is obtained as

$$Var\left(r_t^2\right) = \beta^2\left(2 + 5\alpha^2 + \frac{17}{4}\alpha^4\right).$$

Autocorrelation function:

Let us first compute the lag k autocovariance function of $\{r_t^2\}$ given by

$$Cov\left(r_t^2, r_{t-k}^2\right) = E\left(r_t^2 r_{t-k}^2\right) - E\left(r_t^2\right) E\left(r_{t-k}^2\right)$$

$$= E\left(h_t\varepsilon_t^2 h_{t-k}\varepsilon_{t-k}^2\right) - E\left(h_t\varepsilon_t^2\right) E\left(h_{t-k}\varepsilon_{t-k}^2\right)$$
$$= Cov\left(h_t,\, h_{t-k}\right).$$

The expression for $Cov(h_t,\, h_{t-k})$ is derived in Chapter 3 and it is obtained as

$$Cov(r_t^2,\, r_{t-k}^2) = \beta^2\left(\frac{1}{2}\alpha^4\rho^{2k} + \alpha^2 I_1\right). \tag{4.3}$$

Hence the lag k autocorrelation of the squared sequence $\{r_t^2\}$ is

$$\rho_k\left(r_t^2\right) = \frac{\left(\frac{1}{2}\alpha^4\rho^{2k} + \alpha^2 I_1\right)}{\left(2 + 5\alpha^2 + \frac{17}{4}\alpha^4\right)}, \tag{4.4}$$

where I_1 is the same expression given in (3.15) of Chapter 3.

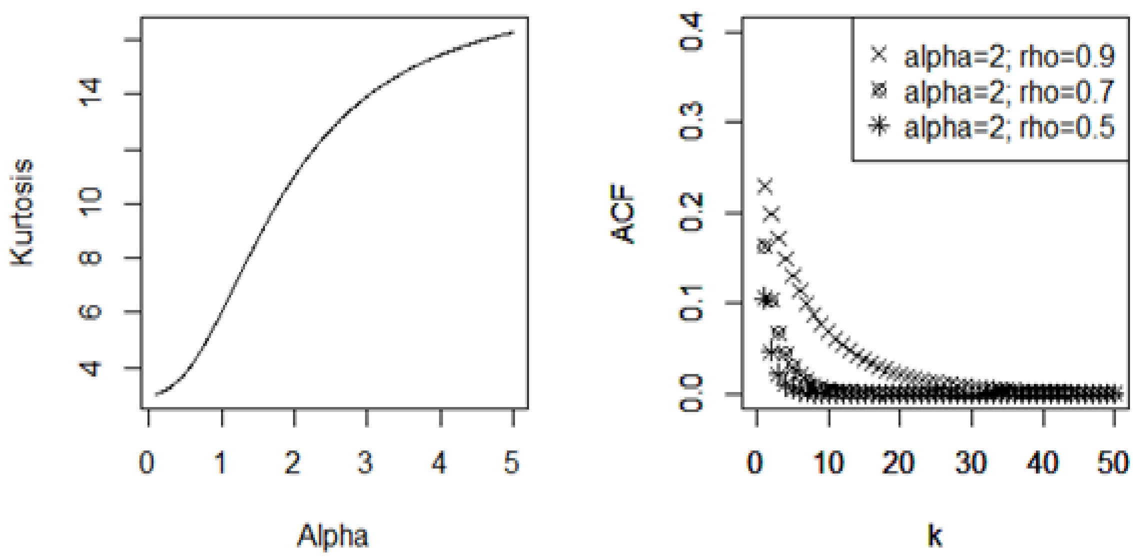

FIGURE 4.1: The plot of kurtosis of return and the ACF of squared return

The ACF is an exponentially decreasing function of the lags for different values of the parameters, as can be seen in Figure 4.1. By choosing different values for α, one can get a distribution with larger kurtosis as shown in Figure 4.1.

4.3 Estimation of Parameters

Estimation of parameters in SV model is difficult because no explicit expression for the likelihood function of SV model is available. It is possible to express the likelihood function as n-fold integrals, but to obtain estimates one may have to use simulation techniques, like simulated maximum likelihood, method of simulated moments or Markov Chain Monte Carlo (MCMC) techniques. Some of the references for simulation-based maximum likelihood methods are Danielsson and Richard (1993), Danielsson (1994), Shephard and Pitt (1997), Durbin and Koopman (1997), Kim et al. (1998), Sandmann and Koopman (1998), Liesenfeld and Richard (2003, 2006), Richard and Zhang (2007). The other methods include the Generalized Method of Moments (GMM) (Melino and Turnbull (1990)), Quasi Maximum Likelihood (QML) approach (Harvey et al. (1994) and Ruiz (1994)), Efficient Method of Moments (EMM) (Gallant et al. (1997)). For an overview of such estimation methods of SV models, see Shephard (1996), Broto and Ruiz (2004) and Tsay (2005). We use the GMM estimation proposed by Melino and Turnbull (1990) and Efficient Importance Sampling (EIS) method by Richard and Zhang (2007) to estimate the parameters.

4.3.1 Parameter Estimation by Method of Moments

The moment estimators are not efficient in general. However, we have computed the moment estimates of the parameters based on simulated samples and found that they are slightly biased. Let $(r_1, r_2, \ldots, r_T)$ be a realization of length T from the SV model (4.1), $\Theta = (\alpha, \beta, \rho)$ be the parameter vector to be estimated. We use the moments

$$E(r_t^2) = \beta\left(1 + \frac{\alpha^2}{2}\right), \qquad E(r_t^4) = 3\beta^2\left(1 + 2\alpha^2 + \frac{3}{2}\alpha^4\right),$$
$$E(r_t^2 r_{t-1}^2) = \beta^2\left(1 + \alpha^2 + \frac{\alpha^4}{4}(1 + 2\rho^2) + \alpha^2 I_1\right)$$

to estimate the parameters.

If we define

$$f(r_t,\, r_{t-1}, \Theta) = \begin{pmatrix} r_t^2 - \beta\left(1 + \frac{\alpha^2}{2}\right) \\ r_t^4 - 3\beta^2\left(1 + 2\alpha^2 + \frac{3}{2}\alpha^4\right) \\ r_t^2 r_{t-1}^2 - \beta^2\left(1 + \alpha^2 + \frac{\alpha^4}{4}(1 + 2\rho^2) + \alpha^2 I_1\right) \end{pmatrix}, \tag{4.5}$$

where I_1 is given by (3.15), then the moment estimator may be obtained by solving

$$\frac{1}{T}\sum_{t=1}^{T} f(r_t,\, r_{t-1}, \Theta) = 0.$$

The resulting moment equations for α, β and ρ are expressed as

$$\frac{\bar{Y}_2^2}{\bar{Y}_4} = \frac{\left(1 + \frac{\hat{\alpha}^2}{2}\right)^2}{3\left(1 + 2\hat{\alpha}^2 + \frac{3}{2}\hat{\alpha}^4\right)}; \qquad \hat{\beta} = \frac{\bar{Y}_2}{\left(1 + \frac{\hat{\alpha}^2}{2}\right)}$$

and

$$\bar{Y}_{22} = \hat{\beta}\left(1 + \hat{\alpha}^2 + \frac{\hat{\alpha}^4}{4}(1 + 2\hat{\rho}^2) + \hat{\alpha}^2 \hat{I}_1\right),$$

where

$$\bar{Y}_2 = (1/T)\sum_{t=1}^{T} r_t^2, \qquad \bar{Y}_{22} = (1/T)\sum_{t=1}^{T} r_t^2 r_{t-1}^2, \qquad \bar{Y}_4 = (1/T)\sum_{t=1}^{T} r_t^4.$$

These equations have to be solved by numerical methods and are illustrated using simulated samples in Section 4.4. The algorithm for computing I_1 is given in Appendix B. In our further analysis, we will use these estimators as initial values for iterative methods.

4.3.2 Parameter Estimation by Efficient Importance Sampling

The likelihood-based inference requires elimination of latent variables, from the likelihood function. Let $R = (r_1, r_2, ..., r_T)$ be a vector of observations from the model and $H = (h_1, h_2, ..., h_T)$ be the vector of associated latent variables. If we denote the joint density function of (R, H) by $f(R, H; \Theta)$, then the likelihood function of the parameter vector $\Theta = (\alpha, \beta, \rho)$ based on the observations is given by

$$L(\Theta; R) = \int f(R, H; \Theta)\, dH = \iint ... \int f(r_1, r_2, ..., r_T, h_1, h_2, ..., h_T)\, dh_1 dh_2 ... dh_T. \tag{4.6}$$

So the maximum likelihood method of estimation involves the computation of the likelihood function by evaluating this multiple integral and then maximizing the resulting function with respect to the parameters. One of the common methods used in such situations is to obtain the Monte Carlo (MC) estimates of the likelihood function based on the observations simulated from the auxiliary variables. However, such procedures lead to inefficient estimators. Richard and Zhang (2007) proposed EIS to overcome this efficiency problem.

To obtain an MC estimate of $L(\Theta; R)$, we need to decompose the above joint density function $f(R, H; \Theta)$ sequentially as

$$f(r_t, h_t; \Theta) = \prod_{t=1}^{T} g(r_t|h_t)p(h_t|h_{t-1}). \tag{4.7}$$

In our case, $g(.)$ and $p(.)$ are respectively given by

$$g(r_t|h_t) = \frac{1}{\sqrt{2\pi h_t}}\, exp\left\{-\frac{r_t^2}{2h_t}\right\} \tag{4.8}$$

and

$$p(h_t|h_{t-1}) = \frac{1}{2\alpha\beta\sqrt{2\pi}\sqrt{1-\rho^2}}\left(\left(\frac{\beta}{h_t}\right)^{1/2} + \left(\frac{\beta}{h_t}\right)^{3/2}\right) \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\alpha}\left(\sqrt{\frac{h_t}{\beta}} - \sqrt{\frac{\beta}{h_t}}\right) - \frac{\rho}{\alpha}\left(\sqrt{\frac{h_{t-1}}{\beta}} - \sqrt{\frac{\beta}{h_{t-1}}}\right)\right]^2\right\}. \tag{4.9}$$

A natural MC estimate of the likelihood function is given by

$$\hat{L}(\Theta; R) = \frac{1}{S}\sum_{j=1}^{S}\left[\prod_{t=1}^{T} g(r_t|\tilde{h}_t^{(j)})\right], \tag{4.10}$$

where $\tilde{h}_t^{(j)}$ denotes a draw from the density $p(h_t|h_{t-1}^{(j)})$. This estimator is highly inefficient since MC variance increases with the sample size.

The EIS procedure constructs a sequence of samplers that exploits the sample information on the $\{h_t\}$as conveyed by $\{r_t\}$. Let $\{m(h_t|h_{t-1}, a_t)\}_{t=1}^T$ denote such a sequence of auxiliary samplers indexed by the auxiliary parameters $\{a_t = (a_{1,t}, a_{2,t})\}_{t=1}^T$. For any given values of the auxiliary parameters, the likelihood function is rewritten as

$$L(\Theta; R) = \int \left[\prod_{t=1}^{T} \frac{f(r_t, h_t|r_{t-1}, h_{t-1}, \Theta)}{m(h_t|h_{t-1}, a_t)} \prod_{t=1}^{T} m(h_t|h_{t-1}, a_t) \right] dh, \tag{4.11}$$

and the corresponding importance sampling MC estimate of the likelihood is given by

$$\tilde{L}(\Theta; R) = \frac{1}{S} \sum_{j=1}^{S} \left\{ \prod_{t=1}^{T} \left[\frac{f(r_t, \tilde{h}_t^{(j)}(a_t)|r_{t-1}, \tilde{h}_{t-1}^{(j)}(a_{t-1}), \Theta)}{m(\tilde{h}_t^{(j)}(a_t)|\tilde{h}_{t-1}^{(j)}(a_{t-1}), a_t)} \right] \right\}, \tag{4.12}$$

where $\left\{ \tilde{h}_t^{(j)}(a_t) \right\}_{t=1}^T$ are trajectories drawn from the auxiliary samplers.

The EIS aims at selecting values of the auxiliary parameters $\{a_t\}_{t=1}^T$ which provide a good match between the product in the numerator and that in the denominator of (4.12) in order to minimize the MC sampling variance of $\tilde{L}$. This minimization problem can be decomposed in to a sequence of sub-problems for each element t of the sequence of observations, provided that the elements depending on the lagged values h_{t-1} are transferred back to the $(t-1)^{th}$ minimization sub-problem. More precisely, if we decompose $m(.|.)$ in the product of a function of h_t and h_{t-1} and one of h_{t-1} only, such that

$$m\left(h_t|h_{t-1}, a_t\right) = \frac{k\left(h_t, a_t\right)}{\chi\left(h_{t-1}, a_t\right)},$$

where $\chi\left(h_{t-1},a_t\right)=\int k\left(h_t,a_t\right)dh_t$.

Now the EIS requires solving a back-recursive sequence of low-dimensional least-squares problems of the form:

$$\hat{a}_t(\Theta)=\arg\min_{a_t}\sum_{j=1}^{S}\left\{\ln\left[f(r_t,\tilde{h}_t^{(j)}|r_{t-1},\tilde{h}_{t-1}^{(j)},\Theta)\chi(\tilde{h}_t^{(j)},\hat{a}_{t+1})\right]-c_t-\ln(k(\tilde{h}_t^{(j)},a_t))\right\}^2, \tag{4.13}$$

where c_t are unknown constants to be estimated jointly with a_t. If the density kernel $k\left(h_t,a_t\right)$ is chosen within the exponential family of distributions, the EIS least-squares problems become linear in a_t. Finally, the EIS estimate of the likelihood function for a given value of Θ is obtained by substituting $\hat{a}_t$ for a_t using the following algorithm.

Step 1: Use the natural sampler $m(h_t|h_{t-1},a_t)$ to draw S trajectories of the latent variable $\{\tilde{h}_t^{(j)}\}_{t=1}^T$ as in (4.10).

Step 2: The draws obtained in step 1 are used to solve for each t (in the order from T to 1) the least squares problems described in (4.13), which takes the form of the auxiliary linear regression:

$$-\tfrac{1}{2}\log h_t-\tfrac{1}{2}\log(2\pi)-\tfrac{r_t^2}{2h_t}+\ln\chi(\tilde{h}_t^{(j)},\hat{a}_{t+1})$$

$$=a_{0,t}+a_{1,t}\tilde{h}_t^{(j)}+a_{2,t}(\tilde{h}_t^{(j)})^2+\upsilon_t^{(j)},\quad j=1,2,...,S,$$

where $\upsilon_t^{(j)}$ is the error term.

Step 3: Use the estimated auxiliary parameters $\hat{a}_t$ to obtain S trajectories $\left\{\tilde{h}_t^{(j)}(\hat{a}_t)\right\}_{t=1}^T$ from the auxiliary sampler $m(h_t|h_{t-1},\hat{a}_t)$.

Step 4: Return to Step 2, this time using the draws obtained with the auxiliary sampler. Steps 2, 3 and 4 are usually iterated a small number of times (from 3 to 5), until a reasonable convergence of the parameters $\hat{a}_t$ is obtained.

Once the auxiliary trajectories have attained a reasonable degree of convergence, the simulated samples can be plugged in formula (4.12) to obtain an EIS estimate of the likelihood function. This procedure is embedded in a numerical maximization algorithm that converges to a maximum of the likelihood function. The same random numbers were also employed for each of the likelihood evaluations required by the maximization algorithm. The number of draws used (S in Eq. (4.12)) for all computations in this section is equal to 100. EIS-ML estimates are finally obtained by maximizing $\tilde{L}\left(\theta; X, a\right)$ with respect to θ. For detailed presentation of the algorithm, see Appendix C.

4.4 Simulation Study

This section illustrates our estimation procedure using the simulated data from BS-SV model. We conducted several repeated simulation experiments with different ρ-values, by fixing α= 2 and β= 1. The trajectories of 500, 1000 and 3000 observations from a SV data generating process were simulated 100 times and the EIS-ML estimates were obtained using moment estimates as initial values. The MM estimates are presented in Table 4.1 and EIS-ML estimates are in Table 4.2 with corresponding mean square error in parentheses.

n	ρ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
500	0.90	2.0989(0.6558)	1.3152(0.6268)	0.9450(0.6622)
	0.70	1.9268(0.6563)	1.1559(0.4945)	0.7275(0.5952)
	0.50	1.7638(0.6915)	1.2145(0.4897)	0.4847(0.4621)
	0.25	1.9093(0.6732)	1.1194(0.4325)	0.2978(0.4433)
	0.00	1.9115(0.6219)	1.2234(0.4203)	0.1234(0.3314)
	-0.25	1.9186(0.6654)	1.1245(0.4896)	-0.3051(0.4476)
	-0.50	1.7834(0.7651)	1.1521(0.5123)	-0.5689(0.6123)
	-0.70	1.9167(0.6754)	1.1656(0.4987)	-0.7934(0.5123)
	-0.90	2.0862(0.6578)	1.2565(0.5051)	-0.9562(0.6340)
1000	0.90	1.9367(0.6334)	1.1808(0.4872)	0.9312(0.6622)
	0.70	1.9398(0.5892)	1.1672(0.4092)	0.7287(0.5952)
	0.50	1.8873(0.4907)	1.1783(0.4367)	0.4769(0.4621)
	0.25	1.9462(0.6234)	1.1098(0.4469)	0.2765(0.4433)
	0.00	1.9346(0.6092)	1.1456(0.4064)	0.0997(0.3314)
	-0.25	1.9419(0.6571)	1.1273(0.4764)	-0.2876(0.4476)
	-0.50	1.8995(0.7075)	1.1183(0.5082)	-0.5561(0.6123)
	-0.70	1.9319(0.5767)	1.1519(0.4337)	-0.7409(0.5123)
	-0.90	2.0510(0.5976)	1.1190(0.4278)	-0.9420(0.6340)
3000	0.90	1.9545(0.6066)	1.0967(0.4562)	0.9267(0.5901)
	0.70	1.9581(0.5906)	1.0877(0.4278)	0.7261(0.5783)
	0.50	1.9256(0.4893)	1.1980(0.4084)	0.4729(0.4563)
	0.25	1.9686(0.6024)	1.0835(0.5011)	0.2710(0.4419)
	0.00	1.9382(0.6063)	1.1052(0.4178)	0.0884(0.3882)
	-0.25	1.9576(0.6327)	1.0728(0.4267)	-0.2765(0.4370)
	-0.50	1.9124(0.6529)	1.0639(0.4271)	-0.5394(0.5922)
	-0.70	1.9412(0.5672)	1.1092(0.4093)	-0.7337(0.5092)
	-0.90	1.9610(0.5571)	1.1076(0.4124)	-0.9374(0.5955)

TABLE 4.1: The average estimates and the corresponding mean square error for the MMEs, when α=2, β=1 and for different ρ's.

n	ρ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
500	0.90	1.9429(0.0776)	0.9493(0.1125)	0.8580(0.0576)
	0.70	1.9501(0.0774)	0.9462(0.1277)	0.6590(0.0522)
	0.50	1.9545(0.0873)	0.9420(0.1126)	0.4820(0.0614)
	0.25	1.9513(0.0793)	0.9521(0.1161)	0.2160(0.0507)
	0.00	1.9468(0.0808)	0.9520(0.1195)	-0.0440(0.0709)
	-0.25	1.9420(0.0942)	0.9524(0.1026)	-0.2820(0.0650)
	-0.50	1.9581(0.0915)	0.9480(0.0950)	-0.5640(0.0578)
	-0.70	1.9400(0.0874)	0.9540(0.1043)	-0.7240(0.0656)
	-0.90	1.9525(0.0727)	0.9559(0.0985)	-0.9355(0.0558)
1000	0.90	1.9626(0.0601)	0.9722(0.0815)	0.8850(0.0471)
	0.70	1.9759(0.0662)	0.9729(0.0866)	0.6812(0.0403)
	0.50	1.9812(0.0712)	0.9810(0.0796)	0.4901(0.0521)
	0.25	1.9788(0.0608)	0.9789(0.0811)	0.2396(0.0488)
	0.00	1.9821(0.0699)	0.9809(0.0785)	-0.0221(0.0532)
	-0.25	1.9760(0.0671)	0.9755(0.0711)	-0.2718(0.0519)
	-0.50	1.9789(0.0615)	0.9801(0.0762)	-0.5355(0.0477)
	-0.70	1.9810(0.0711)	0.9882(0.0815)	-0.7188(0.0452)
	-0.90	1.9729(0.0655)	0.9759(0.0795)	-0.9128(0.0410)
3000	0.90	1.9829(0.0521)	0.9923(0.0802)	0.8906(0.0424)
	0.70	1.9876(0.0556)	0.9847(0.0716)	0.6890(0.0475)
	0.50	1.9890(0.0579)	0.9907(0.0689)	0.4955(0.0545)
	0.25	1.9835(0.0600)	0.9844(0.0709)	0.2431(0.0388)
	0.00	1.9901(0.0628)	0.9879(0.0691)	-0.0560(0.0532)
	-0.25	1.9859(0.0571)	0.9937(0.0669)	-0.2517(0.0519)
	-0.50	1.9914(0.0550)	0.9965(0.0681)	-0.5188(0.0427)
	-0.70	1.9945(0.0609)	0.9899(0.0679)	-0.7054(0.0400)
	-0.90	1.9899(0.0533)	0.9902(0.0602)	-0.9059(0.0379)

TABLE 4.2: The average estimates and the corresponding mean square error for the EIS-MLEs, when α=2, β=1 and for different ρ's.

From the above tables, we observe that the MM estimates are slightly biased. When the sample size is large, the estimates perform reasonably well and there is a marginal reduction in bias and root mean square errors. But EIS-ML method provides estimates which are closer to the true parameter values and mean square error of estimates are remarkably small.

4.5 Data Analysis

We apply the BS-SV model to analyse the daily returns for (1) the rate of exchange on the Rupee/Dollar from July 25, 1998 to May 22, 2015 obtained from Database on Indian Economy, Reserve Bank of India and (2) the opening index of Standard and Poors 500 (S&P 500) from January 02, 2008 to May 22, 2015 obtained from Yahoo Finance. The time series plots of these data are given in Figure 4.2.

Denoting the daily price index by p_t, the returns are transformed into continuously compounded rates centred around their sample mean:

$$r_t = 100\left[\ln\left(\frac{p_t}{p_{t-1}}\right) - \left(\frac{1}{T}\right)\sum_{t=1}^{T}\ln\left(\frac{p_t}{p_{t-1}}\right)\right],\ t = 1, 2, ..., T.$$

The left panel show the plots of actual data series and the continuously compounded return series are on the right panels. The descriptive statistics of the return series are reported in Table 4.3, where $Q(20)$ and $Q^2(20)$ are the Ljung-Box statistic for return and squared return series with lag 20. The corresponding χ^2 table value at 5% significance level is 10.117. Hence the test suggests that the return series is serially uncorrelated whereas the squared return series has significant serial correlation. The

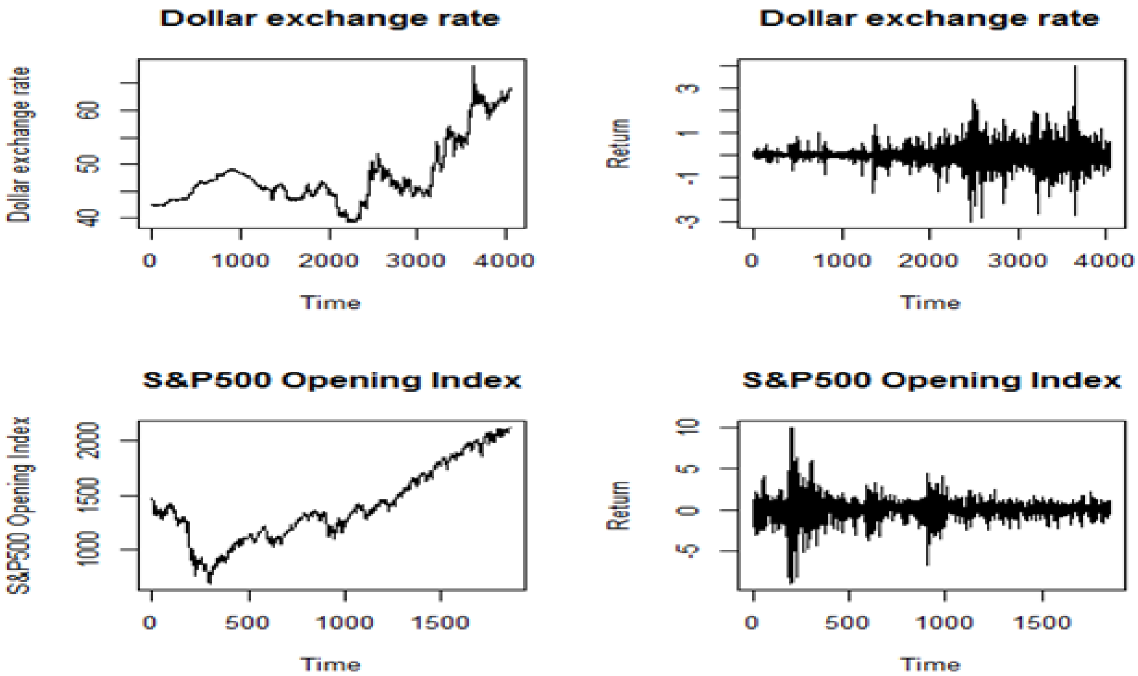

FIGURE 4.2: Time series plot the data and the return

Statistic	Dollar Exchange rate	S&P500 Opening Index
Sample size	4051	1861
Minimum	-3.0164	-9.1349
Maximum	4.0100	10.1193
Std. Dev.	0.4225	1.3496
Kurtosis	11.3384	12.7281
$Q(20)$	1.7088	1.4721
$Q^2(20)$	87.1655	68.1816

TABLE 4.3: Descriptive statistics of the return series

kurtosis of the returns for all the series is greater than three which implies that the distribution of the returns is leptokurtic in nature.

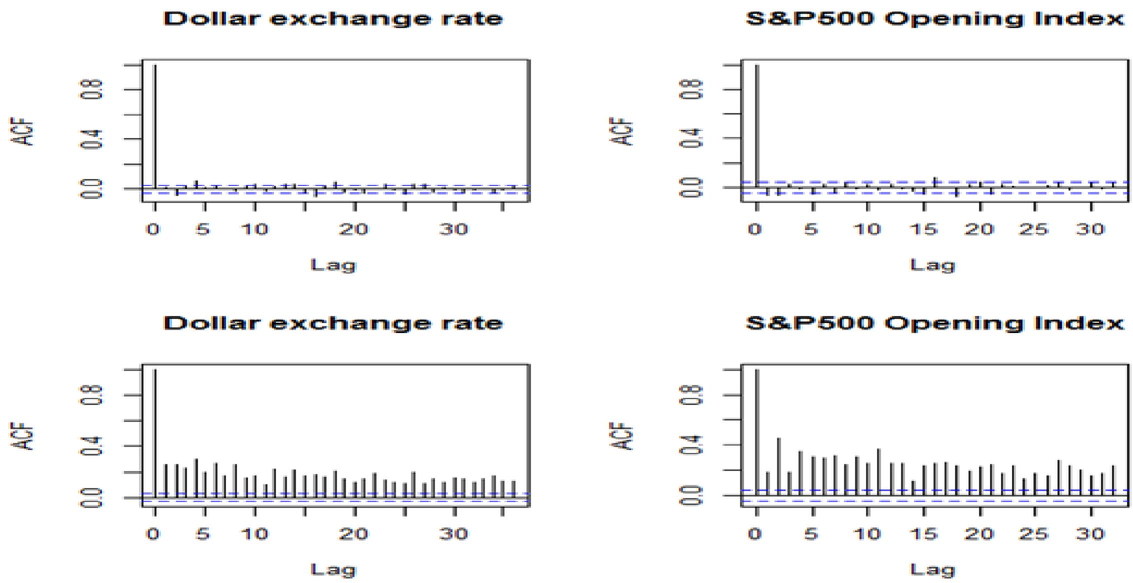

FIGURE 4.3: ACF of the returns and the squared returns

From the ACF of the returns plotted in Figure 4.3, it is observed that serial correlations in the return series are insignificant where as the ACF of the squared returns in the bottom panel remains positive and decays very slowly.

In Table 4.4, we present the parameter estimates for both the return series. Once the estimates of parameters are obtained, the next stage is the model diagnostic checking. That is, we need to check whether the assumptions on the model (4.1) are satisfied with respect to the data we have analysed. Note that the model (4.1) is in terms of the volatilities h_t, which are unobservable. This aspect makes the diagnosis problem difficult. One of the methods suggested in such cases is to employ Kalman filtering by rewriting the model (4.1) in the state-space form. For more details on Kalman filtering method and associated theory, one can refer Jacquier et al. (1994) and Tsay (2005). This method helps in estimating the unobservable volatility h_t.

Parameters	Dollar Exchange rate	S&P 500 Opening Index
α	2.9411	2.5030
β	0.1721	0.4404
ρ	0.9101	0.7942
$Q^*(20)$	0.6820	0.3428
$Q^{2*}(20)$	5.0400	1.1539

TABLE 4.4: Estimates of parameters and Ljung-Box statistic for residuals

The state space representation of the BS-SV model given in (4.1) can be written as

$$\log(r_t^2) = -1.27 + \log h_t + \nu_t, \quad E(\nu_t) = 0,\ V(\nu_t) = \frac{\pi^2}{2} \tag{4.14}$$

and

$$h_t = \beta\left[\frac{1}{2}\alpha X_t + \sqrt{\left(\frac{1}{2}\alpha X_t\right)^2 + 1}\right]^2;\ X_t = \rho X_{t-1} + \eta_t,\ \ t = 1, 2, \dots\ ,$$

where η_t is normally distributed with mean zero and variance $(1-\rho^2)$. If the distribution of ν_t is approximated by a normal distribution then the preceding system (4.13) becomes a standard dynamic linear model, to which the Kalman filter can be applied. Let $\bar{X}_{t|t-1}$ be the prediction of X_t based on the information available at time $t-1$ and $\Omega_{t|t-1}$ be the variance of the prediction. Here we are making an assumption that update that uses the information at time t as $\bar{X}_{t|t}$ and the variance of the update as $\Omega_{t|t}$. The equations that recursively compute the predictions and updating are given by

$$\bar{X}_{t|t-1} = \rho\bar{X}_{t-1|t-1}$$

$$\Omega_{t|t-1} = \rho^2 \Omega_{t-1|t-1} + (1-\rho^2)$$

and

$$\bar{X}_{t|t} = \bar{X}_{t|t-1} + \frac{\Omega_{t|t-1}}{f_t}\left[\log(y_t^2) + 1.27 - \log \bar{X}_{t|t}\right]$$

$$\Omega_{t|t} = \Omega_{t|t-1}(1 - \frac{\Omega_{t|t-1}}{f_t}),$$

where $f_t = \Omega_{t|t-1} + \frac{\pi^2}{2}$.

Then the residuals are calculated by the equation $\hat{\varepsilon}_t = r_t \, \bar{h}_t^{-0.5}$ and use this sequence for the model diagnosis. The system is initialized at the unconditional values, $\Omega_0 = 1$ and $X_0 = 1$. The parameters α, β and ρ in the above system are replaced by their respective estimates which are given in Table 4.4. The correlograms of the residuals are given in Figure 4.4 below suggest that the model performs quite well. Yet,

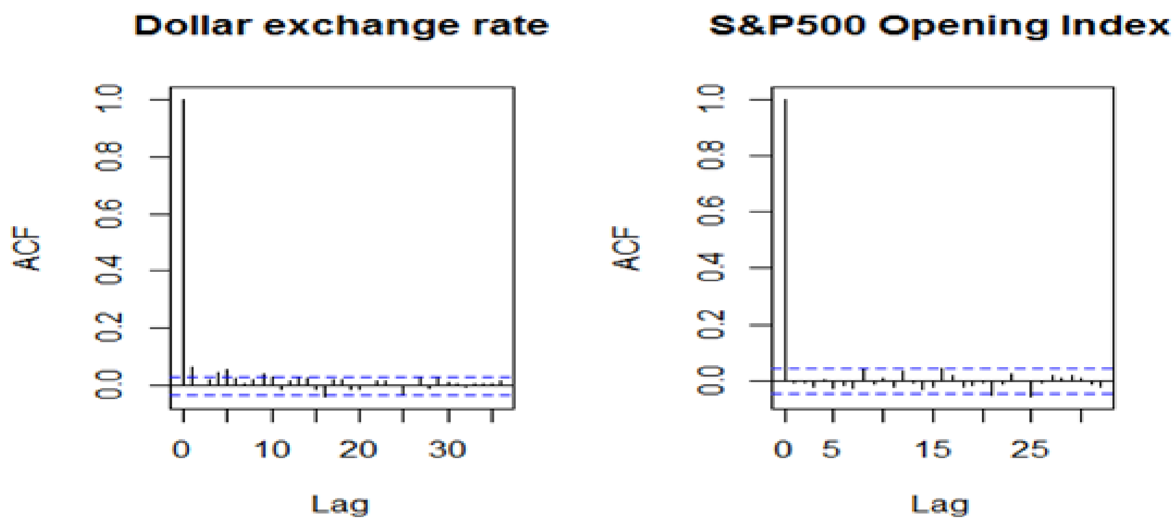

FIGURE 4.4: ACF of the residuals

we have to check formally the serial correlations of the series $\{\hat{\varepsilon}_t\}$ and $\{\hat{\varepsilon}_t^2\}$. The Ljung-Box statistics for the residuals $Q^*(20)$ and the squared residuals $Q^{2*}(20)$ are calculated with lag 20 and are given in Table 4.4. From Table 4.4, the Ljung-Box statistic for both the series is less than the 5% chi-square critical value 10.117 at degrees of freedom 20. Hence we conclude that there is no significant dependence among the residuals and squared residuals. In Figure 4.5 we superimpose the stan-

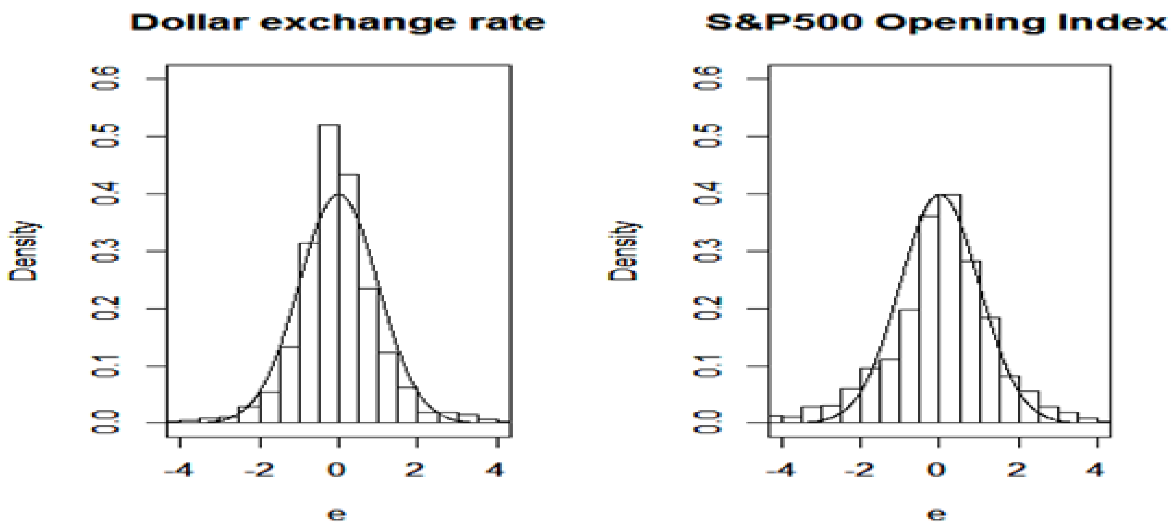

Figure 4.5: Histogram of residuals with superimposed standard normal density

dard normal density on the histogram of the residuals to check whether the series follows standard normal distribution. The figure clearly shows that the standard normal distribution is a good fit for the residuals in both cases. Hence, we conclude that BS-SV model is adequate for above data sets.

Chapter 5

Asymmetric Laplace Stochastic Volatility Model

5.1 Introduction

It is known that return distributions of financial time series data, such as stock and foreign exchange returns, exhibit departure from the normality assumption as they are often skewed and have heavier tails than the normal distribution. Data also exhibit time varying volatility and volatility clustering over time. Accounting for these characteristics of data is crucial to make appropriate decisions for risk management. These aspects motivated the researchers to develop two main classes of models that capture the time-varying auto-correlated volatility process: the ARCH model, introduced by Engle (1982) and the SV model, introduced by Taylor (1986). In ARCH model, the time-varying variance is assumed to be a deterministic function of the lagged values of the squared errors. For a comprehensive survey on this model and its various generalizations, such as Generalized ARCH (GARCH) by Bollerslev (1986), see Shephard (1996) and Tsay (2005). In SV model, the volatility at time t is assumed to be a stochastic process in terms of some latent variables.

The asymmetric Laplace(AL) distribution demonstrates flexibility in fitting data with heavy tails and skewness, which make it a promising candidate for financial data modelling. Kozubowski and Podgorski (2000) and Kotz et al. (2012) studied many properties of asymmetric Laplace distributions. Jayakumar and Kuttykrishnan (2007) introduced a time series model using an asymmetric Laplace distribution for modelling data from financial contexts. Jose and Thomas (2011) developed a first order stationary autoregressive process with generalized Laplace marginal distribution. Although the theory and applications of asymmetric Laplace distributions is well developed and there is considerable literature in recent years, their applications in modelling stochastic volatility in financial time series is not developed. We consider SV model with log-volatility process have an asymmetric Laplace marginal distribution, rather than the Gaussian distribution.

The next section briefly discusses the asymmetric Laplace distribution and its properties. The construction AL-SV model and its second order properties are described in Section 5.3 and 5.4 of this chapter. We discussed the estimation procedure by the method of moments in Section 5.5. The asymptotic properties of estimators are established in Section 5.6. A simulation study is carried out in Section 5.7. In Section 5.8, we present the results on data analysis using our model.

5.2 Asymmetric Laplace distribution

A random variable X is said to have an asymmetric Laplace distribution with parameters $\theta \in \mathrm{R}$, $\kappa > 0$ and $\sigma \geq 0$ ($AL(\theta, \kappa, \sigma)$) if its probability density function is (cf: Kotz et al. (2012)):

$$f(x;\theta,\kappa,\sigma) = \frac{\sqrt{2}}{\sigma}\frac{\kappa}{1+\kappa^2}\begin{cases} \exp\left(-\frac{\sqrt{2}\kappa}{\sigma}\left|x-\theta\right|\right), & if\; x \geq \theta \\ \exp\left(-\frac{\sqrt{2}}{\kappa\sigma}\left|x-\theta\right|\right), & if\; x < \theta \end{cases} \tag{5.1}$$

or, the distribution function of the X is the form

$$F(x;\theta,\kappa,\sigma) = \begin{cases} 1-\frac{1}{1+\kappa^2}\exp\left(-\frac{\sqrt{2}\kappa}{\sigma}\left|x-\theta\right|\right) & , if\; x \geq \theta \\ \frac{\kappa^2}{1+\kappa^2}\exp\left(-\frac{\sqrt{2}}{\kappa\sigma}\left|x-\theta\right|\right) & , if\; x < \theta\ . \end{cases} \tag{5.2}$$

Hence the characteristic function of $AL(\theta, \kappa, \sigma)$ is obtained as

$$\psi_X(t) = E\left(e^{itX}\right) = \frac{e^{i\theta t}}{1+\frac{1}{2}\sigma^2 t^2 - i\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)t}. \tag{5.3}$$

Using (5.3), the mean, variance and the coefficients of skewness and kurtosis can be respectively obtained as

$$E(X) = \theta + \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right), Var(X) = \frac{\sigma^2}{2}\left(\frac{1}{\kappa^2}+\kappa^2\right),$$

$$\gamma = 2\frac{1/\kappa^3-\kappa^3}{\left(1/\kappa^2+\kappa^2\right)^{3/2}}, \qquad K = 6 - \frac{12}{\left(1/\kappa^2+\kappa^2\right)^2}.$$

The absolute value of γ is bounded by two, and as κ increases within the interval $(0, \infty)$, then the corresponding value of γ decreases monotonically from 2 to -2.

Similarly, the distribution is leptokurtic and K varies from 3 (the least value for the symmetric Laplace distribution with $\kappa = 1$) to 6 (the greatest value attained for the limiting exponential distribution when $\kappa \to 0$ (see Kotz et al. (2012)).

The asymmetric Laplace random variable X also admits the representation of the form

$$X \stackrel{d}{=} \rho X + (1-\rho)\theta + \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} I_1 E_1 - \kappa I_2 E_2\right), \ \rho \in [0,1], \tag{5.4}$$

where I_1, I_2 are dependent Bernoulli random variables taking on values of either zero or one with probabilities,

$$P\left(I_1 = 0,\ I_2 = 0\right) = \rho^2, \ \ P\left(I_1 = 1,\ I_2 = 1\right) = 0,$$

$$P\left(I_1 = 1,\ I_2 = 0\right) = (1-\rho)\left(\rho + \frac{1-\rho}{1+\kappa^2}\right),$$

$$P\left(I_1 = 0,\ I_2 = 1\right) = (1-\rho)\left(\rho + \frac{(1-\rho)\kappa^2}{1+\kappa^2}\right),$$

E_1 and E_2 are standard exponential variables, with all variables being mutually independent.

As shown in Kotz et al. (2012), all asymmetric Laplace laws are self decomposable for all values of the parameters. Gaver and Lewis (1980) proved that only self decomposable distributions can be marginal distributions of a first order autoregressive process. Hence the asymmetric Laplace distribution can be the marginal distribution of an AR(1) process. Various authors studied autoregressive models with non-Gaussian marginal distribution extensively in recent years due to wide applications of such models in socio-economic fields. Using the results in Jose and

Thomas (2011), a first order autoregressive process with asymmetric Laplace distribution is constructed in the next section.

5.3 First order Asymmetric Laplace Autoregressive Process

The first order asymmetric Laplace AR process is constituted by $\{h_t,\, t \geq 1\}$, where h_t satisfies the equation,

$$h_t = \rho h_{t-1} + \eta_t\,; \quad \rho \in [0,1), \qquad t > 0, \tag{5.5}$$

where $\{h_t\}$ is a stationary Markov process with asymmetric Laplace marginal distribution with location parameter θ, shape parameter κ and scale parameter σ ($AL(\theta, \kappa, \sigma)$) and $\{\eta_t\}$ is a independent and identically distributed random variables independent of $h_{t-\tau}$ for all $\tau \geq 1$. The basic problem is to find the distribution of $\{\eta_t\}$ such that $\{h_t\}$ has the asymmetric Laplace distribution $AL(\theta, \kappa, \sigma)$ as the stationary marginal distribution. The following theorem proved by Jose and Thomas (2011) summarizes the result in this context.

Theorem 5.1. *The stationary marginal distribution of $\{h_t\}$ in model (5.5) is asymmetric Laplace marginal distribution with parameters θ, κ and σ iff the distribution of η_t is specified as a convolution of the form $\eta_t \stackrel{d}{=} U + (I_1 E_1 - I_2 E_2)$, as in (5.7) provided $\eta_0 \stackrel{d}{=} h_0$.*

Proof. In terms of characteristic function the model in (5.5) can be rewritten as

$$\psi_h(t) = \psi_h(\rho t)\, \psi_\eta(t).$$

Under stationarity assumption, $\psi_\eta(t) = \frac{\psi_h(t)}{\psi_h(\rho t)}$.

Substituting the characteristic function given by (5.3) we get,

$$\psi_\eta(t) = \frac{e^{i\theta t}}{e^{i\theta\rho t}} \frac{\left(1 + i\frac{\sigma\kappa}{\sqrt{2}}\rho t\right)\left(1 - i\frac{\sigma}{\sqrt{2}\kappa}\rho t\right)}{\left(1 + i\frac{\sigma\kappa}{\sqrt{2}} t\right)\left(1 - i\frac{\sigma}{\sqrt{2}\kappa} t\right)}$$

$$= e^{i\theta(1-\rho)t}\left[\rho + (1-\rho)\frac{1}{\left(1 + i\frac{\sigma\kappa}{\sqrt{2}} t\right)}\right]\left[\rho + (1-\rho)\frac{1}{\left(1 - i\frac{\sigma}{\sqrt{2}\kappa} t\right)}\right]. \tag{5.6}$$

This implies that η_t has a convolution structure of the following form

$$\eta_t \overset{d}{=} U + (I_1 E_1 - I_2 E_2)\,, \tag{5.7}$$

where U is degenerate at $\theta(1-\rho)$. E_1 and E_2 are independent exponential random variables with means $\sigma/\sqrt{2}\kappa$ and $\sigma\,\kappa/\sqrt{2}$ respectively and (I_1, I_2) is such that

$$P\left(I_1 = 0,\ I_2 = 0\right) = \rho^2,\ \ P\left(I_1 = 1,\ I_2 = 1\right) = (1-\rho)^2\,,$$

$$P\left(I_1 = 1,\ I_2 = 0\right) = P\left(I_1 = 0,\ I_2 = 1\right) = \rho\,(1-\rho)\,.$$

Further (I_1, I_2) independent of E_1 and E_2.

The converse part can be provided by the method of induction. We assume that h_{t-1} follows $AL(\theta,\kappa,\sigma)$ with characteristic function (5.3).

$$\psi_{h_t}(t) = \psi_{h_{t-1}}(\rho t)\psi_{\eta_t}(t)$$

$$= \left[\frac{e^{i\theta\rho t}}{\left(1+i\frac{\sigma\kappa}{\sqrt{2}}\rho t\right)\left(1-i\frac{\sigma}{\sqrt{2}\kappa}\rho t\right)}\right]\left[\frac{e^{i\theta(1-\rho)t}\left(1+i\frac{\sigma\kappa}{\sqrt{2}}\rho t\right)\left(1-i\frac{\sigma}{\sqrt{2}\kappa}\rho t\right)}{\left(1+i\frac{\sigma\kappa}{\sqrt{2}}t\right)\left(1-i\frac{\sigma}{\sqrt{2}\kappa}t\right)}\right]$$

$$= e^{i\theta t}\left(\frac{1}{1+i\frac{\sigma\kappa}{\sqrt{2}}t}\right)\left(\frac{1}{1-i\frac{\sigma}{\sqrt{2}\kappa}t}\right),$$

which is same as the asymmetric Laplace characteristic function. This shows that $\{h_t\}$ is strictly stationary with asymmetric Laplace marginals provided $\eta_0 \stackrel{d}{=} h_0$ which follows $AL(\theta,\kappa,\sigma)$. Hence the theorem. □

The mean and variance of η_t are given by

$E(\eta_t) = (1-\rho)\left[\theta + \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]$ and $Var(\eta_t) = (1-\rho^2)\left[\frac{\sigma^2}{2}\left(\frac{1}{\kappa}-\kappa\right)^2+\sigma^2\right]$.

Hence the second order properties of the process $\{h_t\}$ are summarized below.

$E(h_t) = \theta + \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)$, $Var(h_t) = \frac{\sigma^2}{2}\left(\frac{1}{\kappa^2}+\kappa^2\right)$ and the ACF, $\rho_k(h_t) = \rho^k$, $k = 1, 2, ...$

The regression of h_t on h_{t-1} is given by

$$\begin{aligned} E(h_t\,|h_{t-1}) &= \rho h_{t-1} + E(\eta_t\,|h_{t-1}) \\ &= \rho h_{t-1} + (1-\rho)\left[\theta + \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right] \\ &= g(\Theta\,;h_{t-1})\ ,\ \ \Theta = (\theta,\kappa,\sigma)'\,. \end{aligned}$$

Next we discuss the construction of SV model generated by first order AL autoregressive process discussed in this section.

5.4 Asymmetric Laplace SV Model

Let $\{r_t\}$ be a sequence of returns on certain financial asset and the volatilities are generated by a Markov sequence $\{\exp(h_t)\}$ of non-negative random variables. Define the SV model

$$\begin{aligned} r_t &= \exp\left(h_t/2\right)\varepsilon_t, \\ h_t &= \rho h_{t-1} + \eta_t, \;\; t = 1, 2, ..., \;\; 0 \leq \rho < 1 \end{aligned} \tag{5.8}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed standard Laplace random variables with mean zero and variance one. We assume that the sequence $\{\varepsilon_t\}$ is independent of h_t and η_t for every t. Here we assume that for every t, the volatility, h_t is an asymmetric Laplace random variables. Since the sequence $\{\varepsilon_t\}$ follows standard Laplace distribution, the odd moments of r_t are zero and its even moments are given by

$$E\left(r_t^{2r}\right) = \frac{(2r)!\; e^{r\theta}}{2^r\left(1 - \frac{1}{2}r^2\sigma^2 - \frac{r\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right)}, \;\; r = 1, 2, ... \tag{5.9}$$

Then

$$Var\left(r_t\right) = \frac{e^{\theta}}{\left(1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right)}$$

and the kurtosis of r_t becomes

$$K = 6\ \frac{\left[1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]^2}{\left[1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]}. \tag{5.10}$$

By choosing different values for σ and κ, one can get a distribution with larger kurtosis as shown in Figure 5.1.

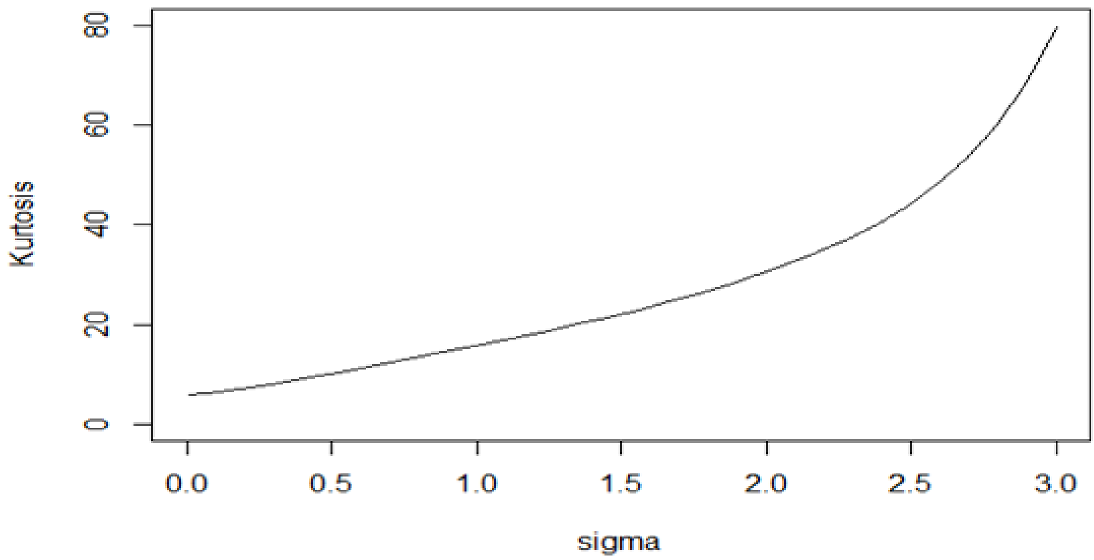

FIGURE 5.1: The plot of kurtosis of r_t

The variance and covariance function of the squared return series are obtained as

$$Var(r_t^2) = E(r_t^4) - \left(E(r_t^2)\right)^2$$

$$= \frac{e^{2\theta}\left\{6\left[1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]^2 - \left[1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]\right\}}{\left[1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]\left[1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]^2}$$

$$Cov\left(r_t^2,\ r_{t-k}^2\right) = E\left(r_t^2 r_{t-k}^2\right) - E\left(r_t^2\right) E\left(r_{t-k}^2\right).$$

For this purpose, we consider

$$
\begin{aligned}
E\left(r_t^2 r_{t-k}^2\right) &= E\left(e^{h_t}\varepsilon_t^2 e^{h_{t-k}}\varepsilon_{t-k}^2\right) = E\left(e^{h_t}e^{h_{t-k}}\right)\\
&= E\left(e^{\left\{\rho^k h_{t-k}+\sum_{i=1}^{k}\rho^{k-i}\eta_{t-(k-i)}\right\}}e^{h_{t-k}}\right)\\
&= E\left(e^{(1+\rho^k)h_{t-k}}\right)E\left(e^{\rho^{k-1}\eta_{t-(k-1)}}\right)E\left(e^{\rho^{k-2}\eta_{t-(k-2)}}\right)...E\left(e^{\rho\eta_{t-1}}\right)E\left(\eta_t\right)\\
&= E\left(e^{(1+\rho^k)h_{t-k}}\right)\frac{E\left(e^{h_t}\right)}{E\left(e^{\rho^k h_{t-k}}\right)}\\
&= \frac{e^{2\theta}\left[1-\frac{1}{2}\sigma^2\rho^{2k}-\frac{\sigma\rho^k}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-\frac{\sigma^2\left(1+\rho^k\right)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho^k)\right]\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}
\end{aligned}
$$

and so

$$
Cov\left(r_t^2,\ r_{t-k}^2\right) = \frac{e^{2\theta}\sigma^2\rho^k\left[1+\frac{1}{2}\left(\frac{1}{\kappa}-\kappa\right)^2+\frac{\sigma^2\rho^k}{4}+\frac{\sigma}{2\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\left(1+\rho^k\right)\right]}{\left[1-\frac{\sigma^2\left(1+\rho^k\right)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho^k)\right]\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]^2}.
$$

Hence, the lag k autocorrelation of the squared sequence $\{r_t^2\}$ is

$$
\rho_{r_t^2}(k) = C(\sigma,\kappa)\times\frac{\sigma^2\rho^k\left[1+\frac{1}{2}\left(\frac{1}{\kappa}-\kappa\right)^2+\frac{\sigma^2\rho^k}{4}+\frac{\sigma}{2\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\left(1+\rho^k\right)\right]}{\left[1-\frac{\sigma^2\left(1+\rho^k\right)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho^k)\right]}, \quad (5.11)
$$

where

$$
C(\sigma,\kappa) = \frac{\left[1-2\sigma^2-\sqrt{2}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left\{6\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]^2-\left[1-2\sigma^2-\sqrt{2}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]\right\}}.
$$

The ACF is an exponentially decreasing function of the lags for different values of the parameters, as can be seen in Figure 5.2.

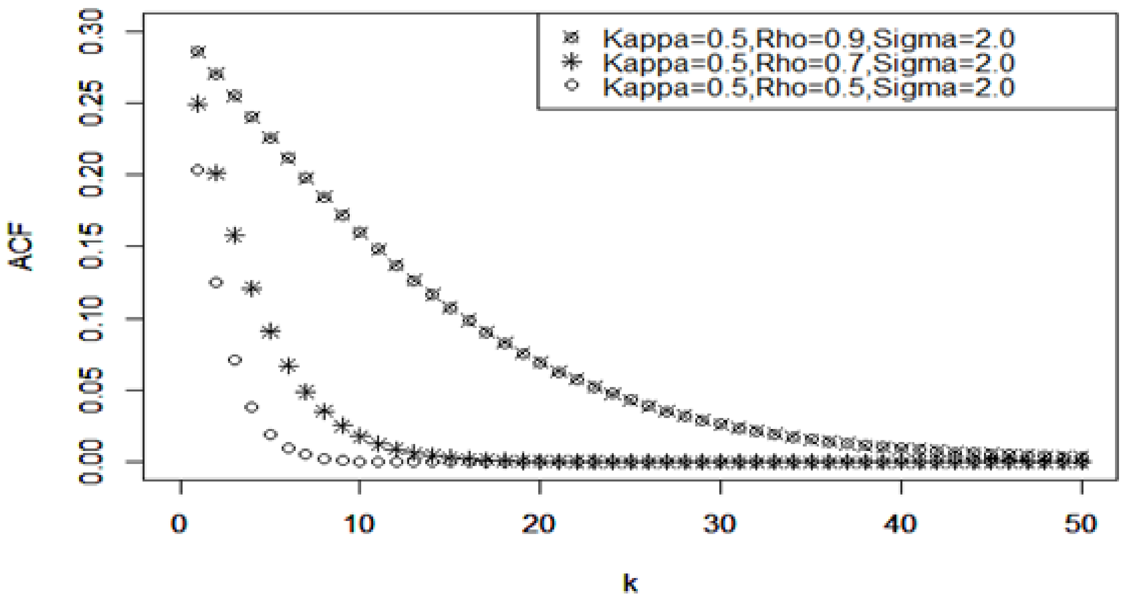

FIGURE 5.2: The ACF of squared returns for different combinations of parameters

5.5 Parameter Estimation

One of the difficulties with the statistical inference for SV models is that the likelihood function involves the unobservable Markov dependent latent variables. These variables have to be integrated out using multiple integrals and this complicates the parameter estimation by the method of maximum likelihood. A number of methods are proposed for estimating the parameters of a SV model and a comprehensive survey may be seen in Tsay (2005). For the SV model described above, we adopt the method of moments to estimate the parameters. Let $(r_1, r_2, ..., r_T)$ be a realization of length T from the AL-SV model (5.8) and $\Theta = (\theta, \kappa, \sigma, \rho)'$ be the parameter vector to be estimated. We use the moments

$$E\left(r_t^2\right) = \frac{e^{\theta}}{\left[1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]}; \qquad E\left(r_t^4\right) = \frac{6\, e^{2\theta}}{\left[1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]},$$

$$E\left(r_t^6\right) = \frac{90\, e^{3\theta}}{\left[1 - \frac{9}{2}\sigma^2 - \frac{3}{\sqrt{2}}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]},$$

$$E\left(r_t^2\, r_{t-1}^2\right) = \frac{e^{2\theta}\left[1 - \frac{\sigma^2\rho^2}{2} - \frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]}{\left[1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]\left[1 - \frac{\sigma^2(1+\rho)^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)(1+\rho)\right]}$$

to estimate the parameters.

We define

$$f\left(r_t\,,\, r_{t-1},\, \theta\right) = \begin{pmatrix} r_t^2 - \frac{e^{\theta}}{\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]} \\ r_t^4 - \frac{6\, e^{2\theta}}{\left[1-2\sigma^2-\sqrt{2}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]} \\ r_t^6 - \frac{90\, e^{3\theta}}{\left[1-\frac{9}{2}\sigma^2-\frac{3}{\sqrt{2}}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]} \\ r_t^2\, r_{t-1}^2 - \frac{e^{2\theta}\left[1-\frac{\sigma^2\rho^2}{2}-\frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\left[1-\frac{\sigma^2(1+\rho)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right]} \end{pmatrix} \tag{5.12}$$

$$= \begin{pmatrix} r_t^2 - c_1 \\ r_t^4 - c_2 \\ r_t^6 - c_3 \\ r_t^2\, r_{t-1}^2 - c_4 \end{pmatrix}, \quad (Say).$$

Then the moment estimator $\hat{\Theta} = \left(\hat{\theta}, \hat{\kappa}, \hat{\sigma}, \hat{\rho}\right)'$ of Θ may be obtained by solving $\frac{1}{T}\sum_{t=1}^{T} f\left(r_t, r_{t-1}, \Theta\right) = 0$. The resulting moment equations for $\Theta = \left(\theta, \kappa, \sigma, \rho\right)'$ are expressed as

$$\hat{\sigma}^2 = \frac{2e^{\hat{\theta}}\bar{Y}_4 - 6e^{2\hat{\theta}}\bar{Y}_2 - \bar{Y}_2\bar{Y}_4}{\bar{Y}_2\bar{Y}_4},$$

$$\bar{Y}_4 = \frac{6\,e^{2\theta}}{\left[1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]},$$

$$\bar{Y}_6 = \frac{90\,e^{3\theta}}{\left[1 - \frac{9}{2}\sigma^2 - \frac{3}{\sqrt{2}}\sigma\left(\frac{1}{\kappa} - \kappa\right)\right]},$$

$$\frac{\bar{Y}_{22}}{\bar{Y}_2} = \frac{e^{\theta}\left[1 - \frac{\sigma^2\rho^2}{2} - \frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)\right]}{\left[1 - \frac{\sigma^2(1+\rho)^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)(1+\rho)\right]},$$

where $\bar{Y}_2 = (1/T)\sum_{t=1}^{T} r_t^2$, $\bar{Y}_4 = (1/T)\sum_{t=1}^{T} r_t^4$, $\bar{Y}_6 = (1/T)\sum_{t=1}^{T} r_t^6$ and $\bar{Y}_{22} = (1/T)\sum_{t=1}^{T} r_t^2 r_{t-1}^2$. These equations have to be solved by numerical methods and computational details are given in Appendix D.

5.6 Asymptotic Properties of Estimators

To prove that the moment estimators are consistent and asymptotically normal (CAN), we refer to Hansen (1982) under the following assumptions. Hansen (1982) proved results for GMM estimators. This theorem also valid for method of moment estimators.

(i) $\{r_t : -\infty < t < \infty\}$ is stationary and ergodic sequence.

(ii) The parameter space Θ is an open subset of R^q that contains the true parameter υ_0.

(iii) $f\left(.\,,\upsilon\right)$ and $\frac{\partial f}{\partial \upsilon}$ are Borel measurable for each $\upsilon \in \Theta$ and $\frac{\partial f(r,.)}{\partial \upsilon}$ is continuous on Θ for each $r \in R^q$.

(iv) $\frac{\partial f_1}{\partial \upsilon}$ is first moment continuous at υ_0, $D = E\left[\frac{\partial}{\partial \upsilon} f\left(r_t\,,\,\upsilon_0\right)\right]$ exists, is finite, and has full rank.

(v) Let $\omega_t = f\left(r_t\,,\upsilon_0\right)\,,\;\; -\infty < t < \infty$ and

$$\vartheta_j = E\left(\omega_0|\omega_{-j},\,\omega_{-j-1},\ldots\right) - E\left(\omega_0|\omega_{-j-1},\,\omega_{-j-2},\ldots\right),\; j \geq 0.$$

The assumptions are that $E\left(\omega_0\,,\,\omega_0^{'}\right)$ exists and is finite, $E\left(\omega_0\,|\omega_{-j},\,\omega_{-j-1},\ldots\right)$ converges in mean square to zero and $\sum_{j=0}^{\infty} E\left(\vartheta_j^{'}\vartheta_j\right)^{1/2}$ is finite. Now we have the following result, proved by Hansen (1982).

Theorem 5.2. *Suppose that the sequence $\{r_t : -\infty < t < \infty\}$ satisfies the assumptions (i) - (v). Then $\left\{\sqrt{T}\left(\hat{\Theta} - \Theta\right), T \geq 1\right\}$ converges in distribution to a normal random vector with mean 0 and dispersion matrix $\left[D\,S^{-1}D^{'}\right]^{-1}$, where D is as given in (iv) and $S = \sum_{k=-\infty}^{\infty}\Gamma_{(k)},\, \Gamma_{(k)} = E\left(\omega_t\,\omega_{t-k}^{'}\right)$.*

The sequence $\{r_t\}$ given in (5.8) is stationary, ergodic and has finite moments, due the fact that $\{h_t\}$ holds these properties. Therefore, the regularity conditions listed above hold good for our AL-SV model.

For computing the elements of the asymptotic dispersion matrix, the following observations become useful.

$$m_{2;2}^{(k)} = E\left(r_t^2 r_{t-k}^2\right) = \frac{E\left(e^{(\rho^k+1)h_{t-k}}\right) E\left(e^{h_t}\right)}{E\left(e^{\rho^k h_{t-k}}\right)};$$

$$m_{2;4}^{(k)} = E\left(r_t^2 r_{t-k}^4\right) = 6\ \frac{E\left(e^{(\rho^k+2)h_{t-k}}\right) E\left(e^{h_t}\right)}{E\left(e^{\rho^k h_{t-k}}\right)};$$

$$m_{2;6}^{(k)} = E\left(r_t^2 r_{t-k}^6\right) = 90\ \frac{E\left(e^{(\rho^k+3)h_{t-k}}\right) E\left(e^{h_t}\right)}{E\left(e^{\rho^k h_{t-k}}\right)};$$

$$m_{2;2,2}^{(k)} = E\left(r_t^2 r_{t-k}^2 r_{t-k-1}^2\right) = \frac{E\left(e^{(\rho^{k+1}+\rho+1)h_{t-k-1}}\right) E\left(e^{(\rho^k+1)h_{t-k}}\right) E\left(e^{h_t}\right)}{E\left(e^{(\rho^{k+1}+\rho)h_{t-k-1}}\right) E\left(e^{\rho^k h_{t-k}}\right)};$$

$$m_{4;2}^{(k)} = E\left(r_t^4 r_{t-k}^2\right) = 6\frac{E\left(e^{(2\rho^k+1)h_{t-k}}\right) E\left(e^{2h_t}\right)}{E\left(e^{2\rho^k h_{t-k}}\right)};$$

$$m_{4;4}^{(k)} = E\left(r_t^4 r_{t-k}^4\right) = 36\frac{E\left(e^{(2\rho^k+2)h_{t-k}}\right) E\left(e^{2h_t}\right)}{E\left(e^{2\rho^k h_{t-k}}\right)};$$

$$m_{4;6}^{(k)} = E\left(r_t^4 r_{t-k}^6\right) = 540\frac{E\left(e^{(2\rho^k+3)h_{t-k}}\right) E\left(e^{2h_t}\right)}{E\left(e^{2\rho^k h_{t-k}}\right)};$$

$$m_{4;2,2}^{(k)} = E\left(r_t^4 r_{t-k}^2 r_{t-k-1}^2\right) = 6\ \frac{E\left(e^{(2\rho^{k+1}+\rho+1)h_{t-k-1}}\right) E\left(e^{(2\rho^k+1)h_{t-k}}\right) E\left(e^{2h_t}\right)}{E\left(e^{(2\rho^{k+1}+\rho)h_{t-k-1}}\right) E\left(e^{2\rho^k h_{t-k}}\right)};$$

$$m_{6;2}^{(k)} = E\left(r_t^6 r_{t-k}^2\right) = 90\ \frac{E\left(e^{(3\rho^k+1)h_{t-k}}\right) E\left(e^{3h_t}\right)}{E\left(e^{3\rho^k h_{t-k}}\right)};$$

$$m_{6;4}^{(k)} = E\left(r_t^6 r_{t-k}^4\right) = 540\frac{E\left(e^{(3\rho^k+2)h_{t-k}}\right) E\left(e^{3h_t}\right)}{E\left(e^{3\rho^k h_{t-k}}\right)};$$

$$m_{6;2,2}^{(k)} = E\left(r_t^6 r_{t-k}^2 r_{t-k-1}^2\right) = 90\ \frac{E\left(e^{(3\rho^{k+1}+\rho+1)h_{t-k-1}}\right) E\left(e^{(3\rho^k+1)h_{t-k}}\right) E\left(e^{3h_t}\right)}{E\left(e^{(3\rho^{k+1}+\rho)h_{t-k-1}}\right) E\left(e^{3\rho^k h_{t-k}}\right)};$$

$$m_{2,2;2}^{(k)} = E\left(r_t^2 r_{t-1}^2 r_{t-k}^2\right) = \frac{E\left(e^{(\rho^k+\rho^{k-1}+1)h_{t-k}}\right) E\left(e^{(\rho+1)h_{t-1}}\right) E\left(e^{h_t}\right)}{E\left(e^{(\rho^k+\rho^{k-1})h_{t-k}}\right) E\left(e^{\rho h_{t-1}}\right)};$$

$$m_{2,2;4}^{(k)} = E\left(r_t^2 r_{t-1}^2 r_{t-k}^4\right) = 6\,\frac{E\left(e^{(\rho^k+\rho^{k-1}+2)h_{t-k}}\right)E\left(e^{(\rho+1)h_{t-1}}\right)E\left(e^{h_t}\right)}{E\left(e^{(\rho^k+\rho^{k-1})h_{t-k}}\right)E\left(e^{\rho h_{t-1}}\right)};$$

$$m_{2,2;6}^{(k)} = E\left(r_t^2 r_{t-1}^2 r_{t-k}^6\right) = 90\,\frac{E\left(e^{(\rho^k+\rho^{k-1}+3)h_{t-k}}\right)E\left(e^{(\rho+1)h_{t-1}}\right)E\left(e^{h_t}\right)}{E\left(e^{(\rho^k+\rho^{k-1})h_{t-k}}\right)E\left(e^{\rho h_{t-1}}\right)};$$

$$\begin{aligned} m_{2,2;2,2}^{(k)} &= E\left(r_t^2 r_{t-1}^2 r_{t-k}^2 r_{t-k-1}^2\right) \\ &= \frac{E\left(e^{(\rho^{k+1}+\rho^k+\rho+1)h_{t-k-1}}\right)E\left(e^{(\rho^k+\rho^{k-1}+1)h_{t-k}}\right)E\left(e^{(\rho+1)h_{t-1}}\right)E\left(e^{h_t}\right)}{E\left(e^{(\rho^{k+1}+\rho^k+\rho)h_{t-k-1}}\right)E\left(e^{(\rho^k+\rho^{k-1})h_{t-k}}\right)E\left(e^{\rho h_{t-1}}\right)}; \end{aligned}$$

where

$$E\left(e^{rh_t}\right) = \frac{e^{r\theta}}{1-\frac{\sigma^2 r^2}{2}-\frac{\sigma r}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)}.$$

Let $\Gamma_{(k)} = \begin{pmatrix} \gamma_{11}^{(k)} & \gamma_{12}^{(k)} & \gamma_{13}^{(k)} & \gamma_{14}^{(k)} \\ \gamma_{21}^{(k)} & \gamma_{22}^{(k)} & \gamma_{23}^{(k)} & \gamma_{24}^{(k)} \\ \gamma_{31}^{(k)} & \gamma_{32}^{(k)} & \gamma_{33}^{(k)} & \gamma_{34}^{(k)} \\ \gamma_{41}^{(k)} & \gamma_{42}^{(k)} & \gamma_{43}^{(k)} & \gamma_{44}^{(k)} \end{pmatrix}, \quad k = 0, \pm 1, \pm 2, ...$

and $\Gamma_{(k)} = \Gamma_{(-k)},\ k = 1, 2, ...$. Then the 4×4 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

$$\begin{aligned} \gamma_{11}^{(0)} &= m_{2;2}^{(0)} - c_1^2, \\ \gamma_{12}^{(0)} &= \gamma_{21}^{(0)} = m_{2;4}^{(0)} - c_1 c_2, \\ \gamma_{13}^{(0)} &= \gamma_{31}^{(0)} = m_{2;6}^{(0)} - c_1 c_3, \\ \gamma_{14}^{(0)} &= \gamma_{41}^{(0)} = m_{2;2,2}^{(0)} - c_1 c_4, \end{aligned}$$

$$\gamma_{22}^{(0)} = m_{4;4}^{(0)} - c_2^2,$$

$$\gamma_{23}^{(0)} = \gamma_{32}^{(0)} = m_{4;6}^{(0)} - c_2c_3,$$

$$\gamma_{24}^{(0)} = \gamma_{42}^{(0)} = m_{4;2,2}^{(0)} - c_2c_4,$$

$$\gamma_{33}^{(0)} = m_{6;6}^{(0)} - c_3^2,$$

$$\gamma_{34}^{(0)} = \gamma_{43}^{(0)} = m_{6;2,2}^{(0)} - c_3c_4,$$

$$\gamma_{44}^{(0)} = m_{2,2;2,2}^{(0)} - c_4,$$

where c_1, c_2, c_3 and c_4 as in (5.12).

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

$$\gamma_{11}^{(k)} = m_{2;2}^{(k)} - c_1^2; \gamma_{12}^{(k)} = m_{2;4}^{(k)} - c_1c_2; \gamma_{13}^{(k)} = m_{2;6}^{(k)} = c_1c_3; \gamma_{14}^{(k)} = m_{2;2,2}^{(k)} - c_1c_4,$$

$$\gamma_{21}^{(k)} = m_{4;2}^{(k)} - c_1c_2; \gamma_{22}^{(k)} = m_{4;4}^{(k)} - c_2^2; \gamma_{23}^{(k)} = m_{4;6}^{(k)} - c_2c_3; \gamma_{24}^{(k)} = m_{4;2,2}^{(k)} - c_2c_4,$$

$$\gamma_{31}^{(k)} = m_{6;2}^{(k)} - c_1c_3; \gamma_{32}^{(k)} = m_{6;4}^{(k)} - c_2c_3; \gamma_{33}^{(k)} = m_{6;6}^{(k)} - c_3^2; \gamma_{34}^{(k)} = m_{6;2,2}^{(k)} - c_3c_4,$$

$$\gamma_{41}^{(k)} = m_{2,2;2}^{(k)} - c_1c_4; \gamma_{42}^{(k)} = m_{2,2;4}^{(k)} - c_2c_4; \gamma_{43}^{(k)} = m_{2,2;6}^{(k)} - c_3c_4; \gamma_{44}^{(k)} = m_{2,2;2,2}^{(k)} - c_4^2,$$

The 4×4 matrix D is evaluated using the form $D = E\left(\frac{d}{d\upsilon} f\left(r_t,\, r_{t-1}\,, \upsilon\right)\right)$ and its elements are:

$$D_{11} = -\frac{e^{\theta}}{1 - \frac{\sigma^2}{2} - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa} - \kappa\right)}$$

$$D_{12} = -\frac{12e^{2\theta}}{1 - 2\sigma^2 - \sqrt{2}\sigma\left(\frac{1}{\kappa} - \kappa\right)}$$

$$D_{13} = -\frac{270e^{3\theta}}{1 - \frac{9}{2}\sigma^2 - \frac{3}{\sqrt{2}}\sigma\left(\frac{1}{\kappa} - \kappa\right)}$$

$$D_{14} = -\frac{2e^{2\theta}\left[1-\frac{\sigma^2\rho^2}{2}-\frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-\frac{\sigma^2(1+\rho)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right]\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]},$$

$$D_{21} = \frac{e^{\theta}\sigma\left(\frac{1}{\kappa^2}+1\right)}{\sqrt{2}\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{22} = \frac{6\sqrt{2}e^{2\theta}\sigma\left(\frac{1}{\kappa^2}+1\right)}{\left[1-2\sigma^2-\sqrt{2}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{23} = \frac{270e^{3\theta}\sigma\left(\frac{1}{\kappa^2}+1\right)}{\sqrt{2}\left[1-\frac{9}{2}\sigma^2-\frac{3}{\sqrt{2}}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{24} = -\frac{\frac{\sigma\rho e^{2\theta}}{\sqrt{2}}\left(\frac{1}{\kappa^2}+1\right)}{\{d_1(\sigma,\kappa,\rho)\ \ d_2(\sigma,\kappa,\rho)\}},$$

$$+\frac{e^{2\theta}d_3(\sigma,\kappa,\rho)\left\{d_1(\sigma,\kappa,\rho)\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa^2}+1\right)+d_2(\sigma,\kappa,\rho)\frac{\sigma(1+\rho)}{\sqrt{2}}\left(\frac{1}{\kappa^2}+1\right)\right\}}{\{d_1(\sigma,\kappa,\rho)\ \ d_2(\sigma,\kappa,\rho)\}^2},$$

$$D_{31} = \frac{e^{\theta}\left[\sigma+\frac{1}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{32} = \frac{6e^{2\theta}\left[4\sigma+\sqrt{2}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-2\sigma^2-\sqrt{2}\sigma\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{33} = \frac{90e^{3\theta}\left[9\sigma+\frac{3}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left[1-\frac{9}{2}\sigma^2-\frac{3\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]^2},$$

$$D_{34} = -\frac{e^{2\theta}\left[\sigma\rho^2+\frac{\rho}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\{d_1(\sigma,\kappa,\rho)\ \ d_2(\sigma,\kappa,\rho)\}}$$

$$+\frac{e^{2\theta}d_3(\sigma,\kappa,\rho)\left\{d_1(\sigma,\kappa,\rho)\left[\sigma+\frac{1}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]+d_2(\sigma,\kappa,\rho)\left[\sigma(1+\rho)^2+\frac{1}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right]\right\}}{\{d_1(\sigma,\kappa,\rho)\ \ d_2(\sigma,\kappa,\rho)\}^2},$$

$$D_{41} = D_{42} = D_{43} = 0,$$

$$D_{44} = -\frac{e^{2\theta}\left[-\sigma^2\rho - \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]}{\left\{\left[1-\frac{\sigma^2(1+\rho)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right]\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\right\}}$$
$$+\frac{e^{2\theta}\left[1-\frac{\sigma^2\rho^2}{2}-\frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\left[\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\left[-\sigma(1+\rho)-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\right]}{\left\{\left[1-\frac{\sigma^2(1+\rho)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right]\left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right]\right\}^2},$$

where

$$\begin{aligned}
d_1(\sigma,\kappa,\rho) &= \left[1-\frac{\sigma^2(1+\rho)^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)(1+\rho)\right];\\
d_2(\sigma,\kappa,\rho) &= \left[1-\frac{\sigma^2}{2}-\frac{\sigma}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right];\\
d_3(\sigma,\kappa,\rho) &= \left[1-\frac{\sigma^2\rho^2}{2}-\frac{\sigma\rho}{\sqrt{2}}\left(\frac{1}{\kappa}-\kappa\right)\right].
\end{aligned}$$

Hence the asymptotic dispersion matrix becomes $\frac{1}{T}\Sigma$, where

$$\Sigma = \left[D\,S^{-1}D'\right]^{-1}.$$

5.7 Simulation Study

We carry out a simulation study to evaluate the performance of the proposed estimators with sample sizes 1000 and 3000. First, we generate a sample of size T from AL Markov sequence specified in (5.5) using the innovation random variable described in (5.7). Then simulate the sequence $\{r_t\}$ using (5.8) model. We use this simulated sample to obtain the estimates of the parameters by solving the moment equations given in the Section 5.5. For each specified value of parameter, we repeat the experiment 1000 times for computing the estimates and then averaged them over the repetitions. The average estimates and the corresponding RMSEs (within parentheses) based on the simulated samples are reported in Table 5.1 and 5.2.

ρ	σ	$\hat{\rho}$	$\hat{\sigma}$	$\hat{\kappa}$	$\hat{\theta}$
0.90	0.5	0.8650(0.1071)	0.4662(0.1835)	1.9626(0.1901)	0.9622(0.1715)
0.75		0.7112(0.1103)	0.4652(0.1752)	1.9659(0.2062)	0.9529(0.1866)
0.50		0.4701(0.1021)	0.4591(0.1854)	1.9512(0.2012)	0.9610(0.1696)
0.25		0.2296(0.1288)	0.4611(0.1752)	1.9588(0.2108)	0.9712(0.1811)
0.00		0.0421(0.1332)	0.4616(0.1899)	1.9521(0.2099)	0.9509(0.1885)
0.90	1.0	0.8710(0.1150)	0.9616(0.1899)	1.9550(0.2005)	0.9755(0.1785)
0.75		0.7132(0.1291)	0.9654(0.1821)	1.9688(0.2156)	0.9652(0.1755)
0.50		0.4735(0.1315)	0.9687(0.1921)	1.9569(0.2102)	0.9725(0.1792)
0.25		0.2345(0.1388)	0.9569(0.1825)	1.9491(0.2155)	0.9795(0.1804)
0.00		0.0302(0.1302)	0.9599(0.1847)	1.9545(0.2088)	0.9605(0.1808)
0.90	1.5	0.8720(0.1235)	1.4650(0.1665)	1.9478(0.2805)	0.9755(0.1865)
0.75		0.7129(0.1197)	1.4626(0.1769)	1.9560(0.2171)	0.9655(0.1711)
0.50		0.4713(0.1256)	1.4561(0.1864)	1.9589(0.2215)	0.9701(0.1865)
0.25		0.2360(0.1298)	1.4621(0.1902)	1.9510(0.2011)	0.9650(0.1715)
0.00		0.0592(0.1351)	1.4598(0.1892)	1.9529(0.2055)	0.9659(0.1895)

TABLE 5.1: The average estimates and the corresponding mean square error of moment estimates based on sample of size n=1000, when κ=2, θ=1 and for different values of ρ and σ

From the above tables, we observe that the estimates are slightly biased. When the sample size is large, the estimators behave reasonably well and there is a significant reduction in bias of the estimates. Hence we claim that the method of moment estimation yields good estimates for the parameters involved.

ρ	σ	$\hat{\rho}$	$\hat{\sigma}$	$\hat{\kappa}$	$\hat{\theta}$
0.90	0.5	0.8835(0.1015)	0.4782(0.1566)	1.9784(0.1821)	0.9805(0.1645)
0.75		0.7321(0.1078)	0.4693(0.1589)	1.9677(0.1845)	0.9869(0.1689)
0.50		0.4788(0.1101)	0.4745(0.1610)	1.9609(0.1894)	0.9798(0.1599)
0.25		0.2387(0.1178)	0.4792(0.1638)	1.9704(0.1812)	0.9840(0.1649)
0.00		0.0301(0.1278)	0.4823(0.1702)	1.9833(0.1904)	0.9799(0.1633)
0.90	1.0	0.8802(0.1067)	0.9788(0.1678)	1.9677(0.1789)	0.9809(0.1701)
0.75		0.7387(0.1137)	0.9756(0.1572)	1.9745(0.1862)	0.9769(0.1566)
0.50		0.4805(0.1089)	0.9721(0.1606)	1.9782(0.1890)	0.9788(0.1629)
0.25		0.2380(0.1024)	0.9820(0.1635)	1.9692(0.1798)	0.9846(0.1634)
0.00		0.0278(0.1178)	0.9788(0.1649)	1.9722(0.1893)	0.9766(0.1589)
0.90	1.5	0.8817(0.1123)	1.4820(0.1559)	1.9655(0.1972)	0.9809(0.1630)
0.75		0.7355(0.1078)	1.4783(0.1587)	1.9684(0.1867)	0.9840(0.1572)
0.50		0.4811(0.1200)	1.4845(0.1588)	1.9730(0.1793)	0.9895(0.1635)
0.25		0.2380(0.1189)	1.4734(0.1649)	1.9738(0.1827)	0.9741(0.1585)
0.00		0.0455(0.1278)	1.4751(0.1611)	1.9746(0.1845)	0.9793(0.1644)

TABLE 5.2: The average estimates and the corresponding mean square error of moment estimates based on sample of size n=3000, when κ=2, θ=1 and for different values of ρ and σ

5.8 Data Analysis

To illustrate the application of the proposed model and the associated inferential results, we analyse two sets of financial data. The data sets used for this purpose are: (1) the daily exchange rate of Rupee/Pound Sterling, with the data consisting of 2399 observations from January 02, 2007 to December 15, 2016 obtained from Database on Indian Economy, Reserve Bank of India; (2) the daily average price of crude oil futures (USD/1 Barrel) traded in Multi Commodity Exchange of India Ltd (MCX), India with the data consisting of 1789 observations January 04, 2010 to December 16, 2016.

The time series plots of these data are given in Figure 5.3. The left panel show the plots of actual data series and the return series are on the right panels.

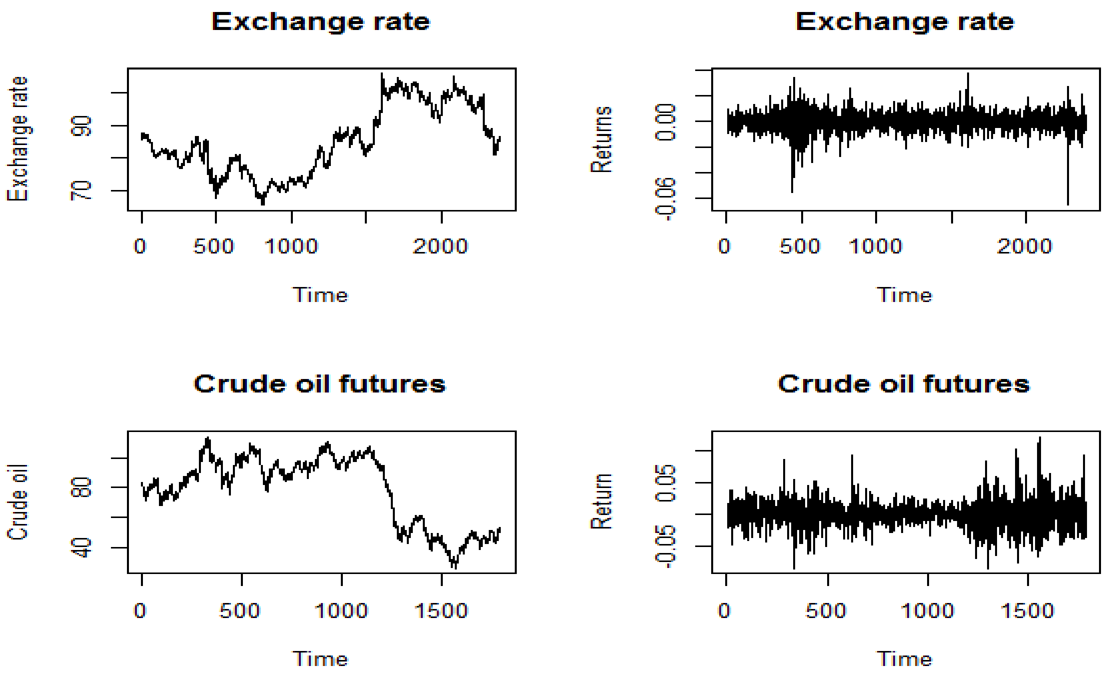

FIGURE 5.3: Time series plot of the original data and the returns

Table 5.3 summarizes the descriptive statistics of the return series, including the mean, median, standard deviation, skewness, kurtosis and Jarque-Bera statistic. The return series are slightly skewed and the kurtosis is well above three, indicating that the return distribution is asymmetric and leptokurtic for both the series. The Jarque-Bera statistic is calculated for the test of joint hypothesis of zero skewness and excess kurtosis and statistic value clearly indicates the return data is non-normal. $Q(20)$ and $Q^2(20)$ are the Ljung-Box statistic for return and squared return series with lag 20. The corresponding χ^2 table value at 5% significance level is

10.117. Hence the test suggests that the return series is serially uncorrelated whereas the squared return series has significant serial correlation.

Statistic	Exchange rate	Crude oil futures
Sample Size	2398	1789
Mean	0.0001	-0.0002
Minimum	-0.0655	-0.0867
Maximum	0.0374	0.1232
Std. Dev.	0.0071	0.0212
Skewness	-0.6628	0.3535
Kurtosis	9.9279	5.8937
$Q(20)$	3.2262	0.0723
$Q^2(20)$	11.7232	20.2069
Jarque-Bera	4971.2720	661.4428

TABLE 5.3: Summary statistics of return series

From the ACF of the returns plotted in Figure 5.4 (left panel), it is observed that serial correlations in the return series are insignificant where as the ACF of the squared returns in the right panel are significant and declines with increasing lags very slowly. In Table 5.4, we present the parameter estimates for each of the return series. The values of the $\hat{\rho}$ in the Table suggest that there is a significant persistence of volatility in the above data series.

Parameter	Exchange rate	Crude oil futures
$\hat{\theta}$	0.0003	-0.1030
$\hat{\kappa}$	0.5122	0.4971
$\hat{\sigma}$	0.2177	0.1076
$\hat{\rho}$	0.7810	0.8620

TABLE 5.4: Parameter estimates using method of moments

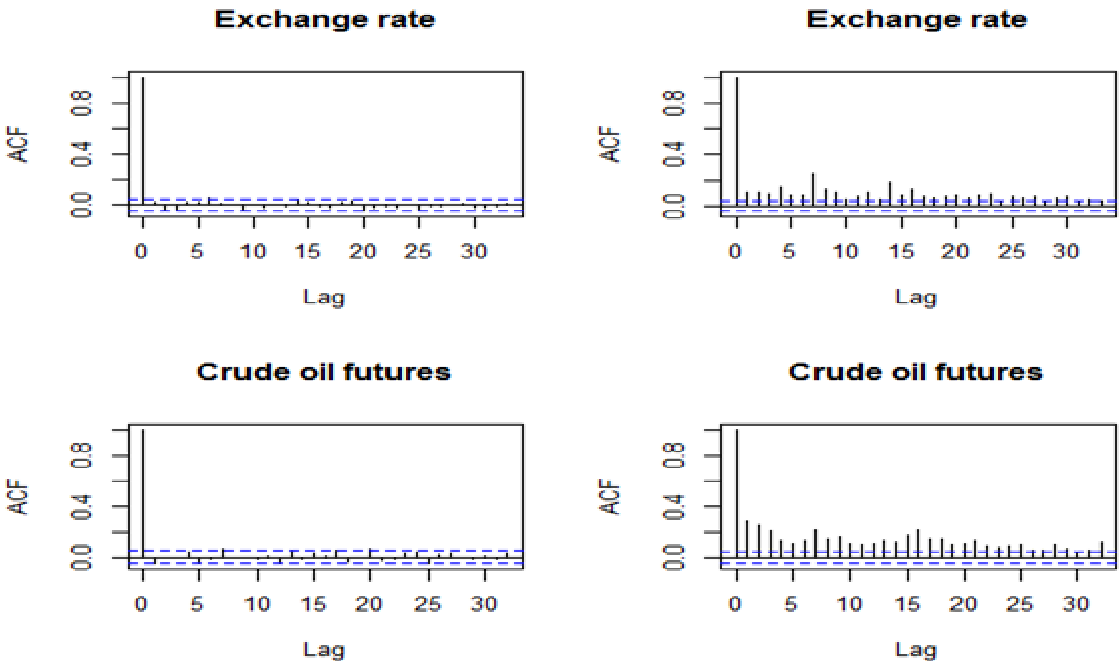

FIGURE 5.4: ACF of the returns and the squared returns

We now perform a diagnostic check of the model based on the residuals. For the AL-SV model, once the estimates of the parameters are obtained, the unobservable component $\{h_t\}$ is estimated using an approximate Kalman filtering (for details see Jacquier et al. (1994). We define the residuals of the model as $\hat{\varepsilon}_t = r_t\, exp(-\hat{h}_t/2)$, where $\hat{h}_t$ is the estimator of h_t provided by the Kalman filter at the MM estimate. If the fitted model is adequate, then $\{\hat{\varepsilon}_t\}$ should behave as an independent and identically distributed sequence of random variables with the assumed distribution. Since the model assumes that the residuals are independent, any dependence on either the residuals or their squares indicates misspecification of the model. In particular, if the fitted model is adequate, both series should have no autocorrelations.

Figure 5.5 gives the sample autocorrelation of the residuals for the fitted AL-SV model. From the figures, the residuals appear to be random and their ACFs fail to indicate any significant serial dependence. Further, we also checked the significance

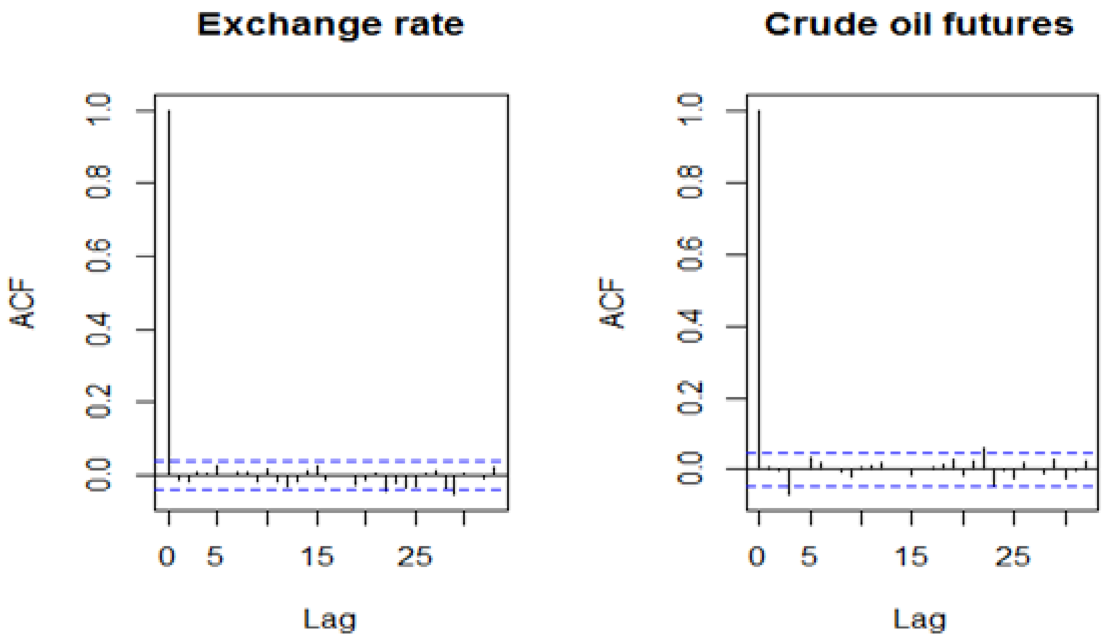

FIGURE 5.5: ACF of the residuals

of ACF in the residues by computing the Ljung-Box statistic for the series $\{\hat{\varepsilon}_t\}$ and $\{\hat{\varepsilon}_t^2\}$, which are summarized in the Table 5.5. All these values are less than the 5% chi-square critical value 10.117 at degrees of freedom 19. Hence we conclude that there is no significant serial dependence among the residuals and the squared residuals.

Data	Ljung-Box statistic	
	Residuals	Squared Residuals
Exchange rate	2.4065	0.0506
Crude oil futures	0.8789	0.2328

TABLE 5.5: Ljung-Box Statistic for the residuals and squared residuals

Thus the data analysis illustrates that the proposed model is capable of capturing the stylized features of financial return series.

The results of this chapter are reported in Balakrishna and Rahul (2017b).

Chapter 6

Inverse Gaussian distribution for Modelling Conditional Durations in Finance

6.1 Introduction

In traditional time series analysis investigators are concerned with the sequence of observations collected at equally spaced intervals. This was also our objective in the last five Chapters of this thesis. That is, in this case, the time process is considered as being non-stochastic. The general time series theory of Autoregressive Moving Average (see Box and Jenkins (1976)) or some of its modifications (see Brockwell and Davis (1991)) can be used in the modelling and forecasting of such situations. Although many financial data may be treated as time series, the standard techniques of time series analysis cannot be employed here directly due to the rapid variation of the time intervals. Since many finance problems involve the arrival of events such as prices or trades in irregular time intervals, a new direction of modelling is necessary to explain the properties of such data.

In order to model the time durations between two successive events, Engle and Russell (1998) introduced the ACD model. Similar to the GARCH model for volatility, the ACD model catches duration clustering and is widely used for calculating expected duration. As mentioned by Hautsch (2012), the model can be directly applied to any other positive valued (continuous) process, such as trading volumes (Manganelli (2005)) , market depth, bid-ask spreads or the number of trades (if they are sufficiently continuous). The basic idea is to (dynamically) parameterize the conditional duration mean rather than the intensity function itself.

Our objective in this chapter is to propose some conditional duration models based on inverse Gaussian distribution and study their properties. The motivation for this approach is: (i) inverse Gaussian distribution is a member of the natural exponential family of distributions and can be considered an alternative to exponential, log-Normal, log-logistic, Frechet and Weibull distributions, among others. Moreover, the inverse Gaussian has a hazard function which is non-monotonic; (ii) it is also likely to prove useful in statistical applications as a flexible and tractable model for fitting duration data, right-skewed unimodal data; (iii) it is a flexible closed form distribution that can be applied to model heavy-tailed processes (for example, it has been applied in many applications in studies of life times, reaction times, reliability and number of event occurrences in fields such as economics, and agricultural science).

Next section contains a brief review of the available ACD models in the literature. In section 6.3 we introduce the IG-ACD model and discuss the properties of the proposed model. The maximum likelihood method of estimation of IG-ACD model is discussed in Section 6.4. Sections 6.5 generalize the IG-ACD model to extended generalized inverse Gaussian (EGIG) ACD model and list the special cases of EGIG-ACD model. Section 6.6 discusses IG- SCD model and its properties. Section 6.7 briefly illustrates the efficient importance sampling method for maximum likelihood estimation of IG-SCD model and Section 6.8 contains the results of the simulation study. Finally Section 6.9 deals with a data analysis for illustrating the methods discussed in the previous Sections.

6.2 Review of ACD Models

Engle and Russell (1998) introduced the most popular ACD model that assumes that the error term follows the standard exponential distribution. The simplest and often very successful member of ACD family is the exponential ACD model. The exponential ACD model denoted by EACD (1,1), may be presented as

$$X_i = \psi_i \varepsilon_i, \qquad \psi_i = \omega + \alpha X_{i-1} + \beta \psi_{i-1}, \tag{6.1}$$

where ε_i follows the standard exponential distribution. We have $E(\varepsilon_i) = 1,\ Var(\varepsilon_i) = 1$, and $E(\varepsilon_i^2) = 2$.

Taking the expectation of the model, we obtain

$$E(X_i) = E(\psi_i \varepsilon_i) = E(\psi_i)E(\varepsilon_i) = E(\psi_i),$$

$$E(\psi_i) = \omega + \alpha E(X_{i-1}) + \beta E(\psi_{i-1}).$$

Under the weak stationarity assumption, $E(X_i) = E(X_{i-1})$, so that

$$E(X_i) = E(\psi_i) = \frac{\omega}{1 - \alpha - \beta} = \mu_x. \tag{6.2}$$

Consequently, $0 \leq \alpha + \beta < 1$ for a weakly stationary process $\{X_i\}$.

We have $E(X_i^2) = 2E(\psi_i^2)$.

Again, under weak stationarity,

$$E(\psi_i^2) = \frac{\mu_x^2[1 - (\alpha + \beta)^2]}{1 - 2\alpha^2 - \beta^2 - 2\alpha\beta},$$

$$Var(X_i) = \frac{\mu_x^2[1 - \beta^2 - 2\alpha\beta]}{1 - 2\alpha^2 - \beta^2 - 2\alpha\beta}. \tag{6.3}$$

From these results, for the EACD(1,1) model to have a finite variance, we need $1 > 2\alpha^2 + \beta^2 + 2\alpha\beta$. Similar results can be obtained for the general EACD (p,q) model, but the algebra involved becomes tedious. One may refer Engle and Russell (1998) for details.

The EACD model has several nice features. For instance, it is simple in theory and in ease of estimation. But the model also encounters some weaknesses. For example, the use of the exponential distribution implies that the model has a constant hazard function. As stated in Tsay (2009) transaction duration in finance is inversely related to trading intensity, which in turn depends on the arrival of new information, making it hard to justify that the hazard function of duration is constant over time. To overcome this weakness, alternative innovation distributions have been proposed in the literature. Engle and Russell (1998) entertain the Weibull distribution for ε_i. A feature of Engle and Russell's linear ACD specification with exponential or Weibull errors is that the implied conditional hazard functions are restricted to being constant, increasing or decreasing. Zhang et al. (2001), Hamilton and Jorda (2002) and Bauwens and Veredas (2004) questioned whether this assumption is an adequate one. As an alternative to the Weibull distribution used in the original ACD model, Lunde (1999) employs a formulation based on the generalized Gamma (GG) distribution, while Grammig and Maurer (2000) and Hautsch (2001) utilize the Burr and generalized F (GF) distributions respectively. Bhatti (2010) introduced Birnbaum-Saunders ACD model as an alternative to the existing ACD models which allow a unimodal hazard function. A recent review of the literature on the ACD models and their applications to finance can be found in Pacurar (2008).

Now, let us describe the specification of the Weibul ACD, GG-ACD, Burr-ACD and BS-ACD models. We begin with the WACD model. The Weibull probability density function with shape parameter θ and scale parameter σ is given by

$$f\left(x;\theta,\sigma\right)=\frac{\theta}{\sigma}\left(\frac{x}{\sigma}\right)^{\theta-1}\exp\left[-\left(\frac{x}{\sigma}\right)^{\theta}\right],\quad x>0;\ \theta,\sigma>0. \tag{6.4}$$

The mean of a Weibull(θ,σ) random variable is $E\left(X\right)=\sigma\,\Gamma\left(1+\theta^{-1}\right)$. Applying the change of variable $\varepsilon_i = X_i/\sigma\,\Gamma\left(1+\theta^{-1}\right)$ to (6.4); we obtain the unit mean

Weibull probability density function

$$f_{\varepsilon}\left(\varepsilon_i\right) = \frac{\theta}{\Gamma\left(1+\theta^{-1}\right)^{-1}}\left(\frac{\varepsilon_i}{\Gamma\left(1+\theta^{-1}\right)^{-1}}\right)^{\theta-1}\exp\left[-\left(\frac{\varepsilon_i}{\Gamma\left(1+\theta^{-1}\right)^{-1}}\right)^{\theta}\right], \varepsilon_i > 0, \theta > 0 \tag{6.5}$$

A final transformation must be applied to obtain the distribution of X_i parametrised in terms of the conditional mean ψ_i. Applying the transformation $\varepsilon_i = X_i/\psi_i$ yields the conditional probability density function

$$f_{X_i|\psi_i}\left(X_i\right) = \frac{\theta}{\psi_i/\Gamma\left(1+\theta^{-1}\right)}\left(\frac{x_i}{\psi_i/\Gamma\left(1+\theta^{-1}\right)}\right)^{\theta-1}\exp\left[-\left(\frac{x_i}{\psi_i/\Gamma\left(1+\theta^{-1}\right)}\right)^{\theta}\right]. \tag{6.6}$$

The specification of the ACD model is completed by specifying the dynamic structure for the conditional mean ψ_i.

Next we consider the construction of the GG-ACD model. The probability density function of the GG distribution is very similar to the probability density function of the Weibull distribution

$$f\left(x;\kappa,\sigma,\theta\right) = \frac{\theta}{\sigma\,\Gamma\left(\kappa\right)}\left(\frac{x}{\sigma}\right)^{\kappa\theta-1}\exp\left[-\left(\frac{x}{\sigma}\right)^{\theta}\right],\ x>0;\ \kappa,\sigma,\theta>0, \tag{6.7}$$

where $\Gamma\left(\kappa\right)$ is the usual gamma function defined by

$$\Gamma\left(\kappa\right) = \int_0^{\infty} x^{\kappa-1}\exp(-x)\,dx.$$

Following the same steps as we did for the WACD model, we apply the transformation $\varepsilon_i = X_i\phi(\kappa,\theta)/\sigma$ where $\phi(\kappa,\theta) = \Gamma(\kappa)/\Gamma(\kappa+\theta^{-1})$ to the probability density function (6.7) to obtain the unit mean GG probability density function

$$f_{\varepsilon}\left(\varepsilon_i\right) = \frac{\theta}{\phi(\kappa,\theta)\Gamma(\kappa)}\left(\frac{\varepsilon_i}{\phi(\kappa,\theta)}\right)^{\theta-1}\exp\left[-\left(\frac{\varepsilon_i}{\phi(\kappa,\theta)}\right)^{\theta}\right]. \tag{6.8}$$

Applying the second transformation $\varepsilon_i = X_i/\psi_i$ yields the conditional likelihood function for X_i given ψ_i

$$f_{X_i|\psi_i}\left(X_i\right) = \frac{\theta}{\phi(\kappa,\theta)\,\psi_i\,\Gamma(\kappa)}\left(\frac{x_i}{\phi(\kappa,\theta)\,\psi_i}\right)^{\theta-1}\exp\left[-\left(\frac{x_i}{\phi(\kappa,\theta)\,\psi_i}\right)^{\theta}\right]. \tag{6.9}$$

Grammig and Maurer (2000) proposed a more flexible specification based on the Burr distribution with probability density function

$$f\left(x;\,\mu,\kappa,\sigma^2\right) = \frac{\mu\kappa x^{\kappa-1}}{(1+\sigma^2\mu x^{\kappa})^{\frac{1}{\sigma^2}+1}},\;\; x>0\,;\;\; \mu,\kappa,\sigma^2>0. \tag{6.10}$$

Lancaster (1992) shows that the Burr distribution can be derived as a Gamma mixture of Weibull distributions. Exponential, Weibull and Log-Logistic are limiting cases. Unlike Weibull and Exponential, the Burr distribution is less frequently used in duration analysis.

Bhatti (2010) introduced BS-ACD model by specifying the time-varying model dynamics in terms of the conditional median duration, instead of the conditional mean duration. The probability density function of the BS(κ,σ) distribution is given by

$$f\left(x;\,\kappa,\sigma\right) = \frac{1}{2\kappa\sigma\sqrt{2\pi}}\left[\left(\frac{\sigma}{x}\right)^{\frac{1}{2}}+\left(\frac{\sigma}{x}\right)^{\frac{3}{2}}\right]\exp\left(-\frac{1}{2\kappa^2}\left[\frac{x}{\sigma}+\frac{\sigma}{x}-2\right]\right),\;\; x>0\,;\;\; \kappa,\sigma>0. \tag{6.11}$$

The conditional probability density function of X_i given σ_i is given by

$$f_{X_i|\sigma_i}(X_i) = \frac{1}{2\kappa\sigma_i\sqrt{2\pi}}\left[\left(\frac{\sigma_i}{x_i}\right)^{\frac{1}{2}} + \left(\frac{\sigma_i}{x_i}\right)^{\frac{3}{2}}\right]\exp\left(-\frac{1}{2\kappa^2}\left[\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right]\right) \quad (6.12)$$

where σ_i is the time-varying conditional median duration.

The other classes of ACD models are defined by different choices of functional form of conditional mean ψ_i. Bauwens et al. (2000) propose a logarithmic ACD (LACD) model that allows the introduction of additional variables without sign restrictions on their coefficients, as the LACD ensures the non-negativity of durations. Fernandes and Grammig (2006) develop a family of augmented ACD (AACD) models that encompasses the standard ACD model, the Log-ACD model and other ACD models inspired by the GARCH literature. Some extended ACD models allow for regime-dependence of the conditional mean function. Zhang et al. (2001) propose a threshold ACD (TACD) model to allow the expected duration to depend nonlinearly on past information variables. Unlike the TACD model, where the transition between states follows a jump process, Meitz and Teräsvirta (2006) introduce a smooth transition ACD (STACD) model. Based on the strong persistence of the trading duration, some long memory ACD models have been introduced. Based on the Ding and Granger (1996) two-component model for volatility, Engle (2000) applies the two-component model for duration. This allows for a slower decay autocorrelation function compared to the corresponding standard model. Jasiak (1998) introduces a fractionally integrated ACD (FIACD) model which is based on a fractionally integrated process for the expected duration. The FIACD model is closely linked with the fractionally integrated GARCH model proposed by Baillie et al. (1996). The

FIACD model is not covariance stationary and implies infinite first and second unconditional moments of the duration. Karanasos (2001) provides an alternative long memory ACD model which is analogous to the long-memory GARCH introduced by Robinson and Henry (1999). Drost and Werker (2004) develop a semiparametric ACD model that can relax the assumption of independently, identically distributed innovations of the standard ACD model. Like the similarity between the ACD and GARCH models, and based on the idea of SV model, Bauwens and Veredas (2004) propose the stochastic conditional duration model for duration. The SCD model is based on the assumption that the durations are generated by a dynamic stochastic latent variable.

Many physical phenomena exhibit hazard functions that are non-monotonic. Grammig and Maurer (2000) provide the motivation to deal with non-monotonic hazard functions when modelling financial duration processes. In the following section we propose a more flexible model for conditional durations based on the inverse Gaussian distribution. One of the important distributions studied in the context of modelling the sequence of durations is the inverse Gaussian distribution. For example, Lancaster (1972) used this distribution to model the intervals between events, such as duration of strikes. In the case of financial series, the sequences of log-returns are assumed to be realizations of Brownian motion or Gaussian process. It is well known that inverse Gaussian distribution arises as a fist passage time distribution in any Gaussian process. Duration between events can be compared with the life times of units in renewal/reliability related studies. The distributions having non-monotonic failure rate are important in such studies and the inverse Gaussian distribution possesses such a property (see Chhikara and Folks (1977)).

6.3 Inverse Gaussian ACD Model

Now let us consider the construction of IG-ACD model. A random variable X 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 is given by

$$f(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\},\ x > 0, \tag{6.13}$$

where μ and λ are assumed to be positive. μ is the mean of the distribution and λ is a shape parameter. This density is unimodal and skewed. The variance for the distribution is μ^3/λ, implying μ is not a location parameter in the usual sense.

Assuming that ε_i follows a unit mean IG distribution with probability density function

$$f_\varepsilon(\varepsilon_i) = \sqrt{\frac{\lambda}{2\pi\varepsilon_i^3}} \exp\left\{-\frac{\lambda(\varepsilon_i-1)^2}{2\varepsilon_i}\right\},\ \varepsilon_i > 0, \tag{6.14}$$

and IG-ACD (1,1) model can be written as

$$X_i = \psi_i\varepsilon_i, \psi_i = \omega + \alpha X_{i-1} + \beta\psi_{i-1}. \tag{6.15}$$

Conditional on F_{i-1}, the probability density function of X_i can be expressed as

$$f(x_i|F_{i-1}) = \frac{1}{\psi_i} f_{\varepsilon_i}\left(\frac{x_i}{\psi_i}\right) = \sqrt{\frac{\lambda\psi_i}{2\pi x_i^3}} \exp\left\{-\frac{\lambda\psi_i(x_i-\psi_i)^2}{2\psi_i^2 x_i}\right\}. \tag{6.16}$$

That is, the conditional distribution of X_i given the past information is $IG(\psi_i,\, \lambda\psi_i)$.

6.3.1 Properties of IG-ACD Model

Conditional on $(X_{i-1}, X_{i-2}, ...,)$ the mean and variance of X_i are given by $E(X_i|F_{i-1}) = \psi_i$ and $Var(X_i|F_{i-1}) = \psi_i^2 Var(\varepsilon_i) = \psi_i^2/\lambda$.

Further the model (6.15) implies that the unconditional mean and variance of the stationary distribution of $\{X_i\}$ can be respectively obtained as

$$\mu_x \equiv E(X_i) = E(\psi_i) = \frac{\omega}{1-\alpha-\beta}, \quad Var(X_i) = \frac{\mu_x^2(1-\beta^2-2\alpha\beta)}{\lambda\left[1-\left(1+\frac{1}{\lambda}\right)\alpha^2-\beta^2-2\alpha\beta\right]}.$$

Consequently, for weak stationarity of $\{X_i\}$ we need the condition $0 \leq \alpha+\beta < 1$.

Autocorrelation function:

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

$$\gamma_k = Cov(X_i,\ X_{i-k}) = Cov(\psi_i,\ X_{i-k})$$

$$= Cov(\omega + \alpha X_{i-1} + \beta\psi_{i-1},\ X_{i-k})$$

$$= \alpha Cov(X_{i-1},\ X_{i-k}) + \beta Cov(\psi_{i-1},\ X_{i-k})$$

$$\gamma_k = (\alpha+\beta)\gamma_{k-1} \tag{6.17}$$

The first order auto-covariance function of X_i is

$$\gamma_1 = Cov(X_i,\ X_{i-1}) = Cov(\psi_i,\ X_{i-1})$$

$$= Cov(\omega + \alpha X_{i-1} + \beta\psi_{i-1},\ X_{i-1})$$

$$\gamma_1 = \alpha\gamma_0 + \beta Var(\psi_{i-1})$$

where $Var(\psi_{i-1}) = \frac{\alpha^2\mu^2}{1-2\alpha^2-\beta^2-2\alpha\beta}$.

Finally, k^{th} order ACF of $\{X_i\}$ is derived as

$$\rho_k = (\alpha+\beta)\rho_{k-1}, \quad k > 1 \tag{6.18}$$

with

$$\rho_1 = \frac{\alpha(1-\beta^2-\alpha\beta)}{1-\beta^2-2\alpha\beta}.$$

Forecasts from an IG-ACD model can be obtained using a procedure similar to that of a GARCH model (cf, Pacurar (2008)).

Intensity function or Hazard function:

Let us denote by T the duration of stay in the state of interest and recall the definition of the hazard function as the instantaneous rate of leaving the interval between $T = t$ and $T = t + \Delta t$, given that it stayed up to time t,

$$h(t) = \lim_{\Delta t \to 0} \frac{p(t \le T < t + \Delta t | T \ge t)}{\Delta t}.$$

Then the hazard function implied by the IG-ACD model may now be written as

$$h(x_i) = \frac{\sqrt{\frac{\lambda\psi_i}{2\pi x_i^3}} \exp\left\{-\frac{\lambda\psi_i(x_i-\psi_i)^2}{2\mu^2 x_i}\right\}}{\Phi\left(\sqrt{\frac{\lambda\psi_i}{x_i}}\left(1-\frac{x_i}{\psi_i}\right)\right) - e^{2\lambda}\Phi\left(-\sqrt{\frac{\lambda\psi_i}{x_i}}\left(1+\frac{x_i}{\psi_i}\right)\right)}, \quad x_i > 0, \tag{6.19}$$

where $\Phi(.)$ is the standard normal distribution function. The expression for $h(x_i)$ is rather complicated but it is not difficult to compute for any given values of parameters. Several typical hazard function curves are given in Figure 6.1. Inspection of these curves reveals that the hazard function is non-monotonic for all μ and λ.

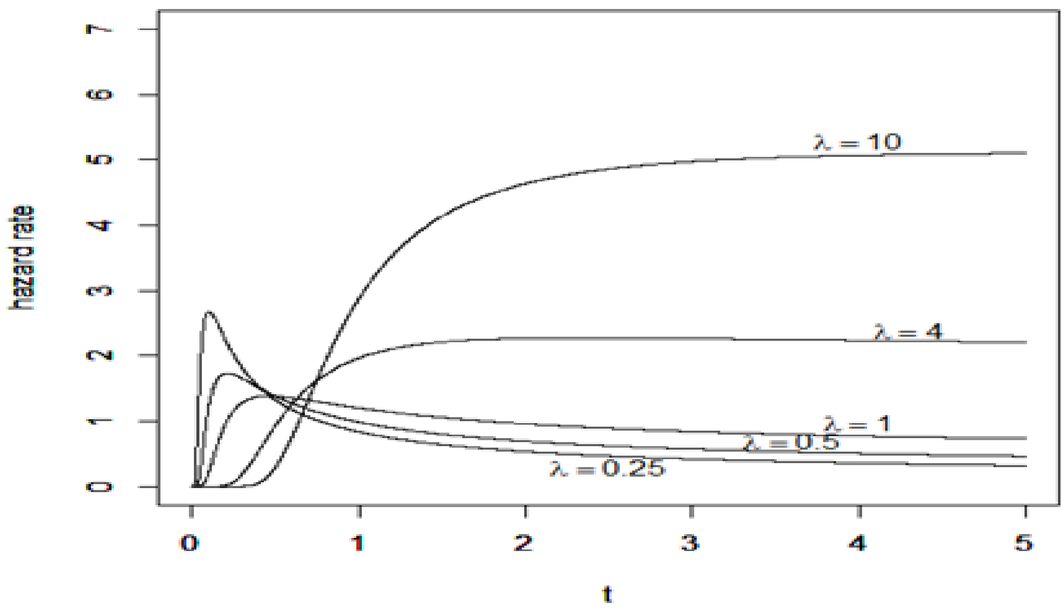

FIGURE 6.1: Hazard rate of inverse Gaussian distribution when $\mu = 1$.

6.4 Estimation of IG-ACD Model

Let $X = (X_1, X_2, ..., X_n)$be a realization from an IG-ACD(1,1) model and the parameter vector to be estimated be $\Theta = (\lambda, \omega, \alpha, \beta)$. The likelihood function of Θ based on X may be expressed as

$$L(\Theta|X) = f(X_1|\Theta) \prod_{i=2}^{n} f(X_i|F_{i-1}; \Theta), \tag{6.20}$$

where $f(X_1;\Theta)$ is the density function of the initial random variable and it does not have a closed form expression. Further its influence on the overall likelihood function diminishes as the sample size n increases and hence we adopt the conditional likelihood method by ignoring the term $f(X_1;\Theta)$. Now using (6.16) in (6.20) the conditional log-likelihood function is given by

$$\log L = \frac{n-1}{2}\log\lambda + \frac{1}{2}\sum_{i=2}^{n}\log\psi_i - \frac{n-1}{2}\log(2\pi) - \frac{3}{2}\sum_{i=2}^{n}\log X_i - \frac{\lambda}{2}\sum_{i=2}^{n}\frac{\psi_i(X_i-\psi_i)^2}{\psi_i^2 X_i}. \tag{6.21}$$

The ML estimator of λ is given by

$$\hat{\lambda} = \left[\frac{1}{n-1}\sum_{i=2}^{n}\frac{\psi_i(X_i-\psi_i)^2}{\psi_i^2 X_i}\right]^{-1}. \tag{6.22}$$

We obtain the ML estimates of the remaining parameters (ω,α,β) by Newton-Raphson iteration method. In order to develop this method let us denote $\Theta = (\omega,\alpha,\beta) = (\theta_1,\theta_2,\theta_3)$.

Taking the first and second order partial derivative of (6.21) with respect to θ_j, we get

$$L'(\theta_j) = \frac{\partial \log L}{\partial\theta_j} = \sum_{i=2}^{n}\left(\frac{1}{2\psi_i} - \frac{\lambda}{X_i} + \frac{\lambda X_i}{2\psi_i^2} + \frac{\lambda}{2X_i}\right)\frac{\partial\psi_i}{\partial\theta_j},$$

$$L''(\theta_j) = \frac{\partial^2 \log L}{\partial\theta_j^2} = -\sum_{i=2}^{n}\left(\frac{1}{2\psi_i^2} + \frac{\lambda X_i}{\psi_i^3}\right)\left(\frac{\partial\psi_i}{\partial\theta_j}\right)^2, \quad j = 1,\ 2,\ 3.$$

Now the iteration formula for estimating θ_j is given by

$$\hat{\theta}_j^{(m+1)} = \hat{\theta}_j^{(m)} - \left[\frac{L'\left(\hat{\theta}_j^{(m)}\right)}{L''\left(\hat{\theta}_j^{(m)}\right)}\right], \qquad m = 1, 2, \ldots, \tag{6.23}$$

where $\hat{\theta}_j^{(m)}$ is the estimate of θ_j obtained at m^{th} iteration. The computation of these estimates based on a simulated sample is illustrated in Section 6.8.

In the next section, we extend the IG-ACD framework to a more flexible specification based on the EGIG distribution. The major advantage of this distribution over the exponential and Weibull distribution, the most common distributions utilized in ACD models, is that these have non-monotonic hazard functions taking bathtub shaped or inverted bathtub shaped forms. In some cases, there is more than one turning point of the hazard rate for EGIG distribution. The shape properties of EGIG hazard are derived in Gupta and Viles (2011).

6.5 Extended Generalized Inverse Gaussian ACD Model

Now let us consider the construction of the EGIG-ACD model. The probability density function of the EGIG distribution is given by

$$f(x; a, b, \lambda, \delta) = \frac{1}{(2/\delta)\,(b/a)^{\lambda/2\delta}\, K_{\frac{\lambda}{\delta}}\left(2\sqrt{ab}\right)} x^{\lambda-1} \exp\left(-ax^{\delta} - bx^{-\delta}\right), \; x > 0, \tag{6.24}$$

where $K_\nu(z)$ is a modified Bessel function of the third kind with index ν and is defined by

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

The domain of variation for the parameters is

$$\lambda \in \mathrm{R}\,, \quad (a, b, \delta) \in \Omega_\lambda,$$

Where

$$\Omega_\lambda = \begin{cases} (a,\, b,\, \delta):\, a > 0,\, b \geq 0,\, \delta > 0 \quad iff\ \lambda > 0 \\ (a,\, b,\, \delta):\, a > 0,\, b > 0,\, \delta > 0 \quad iff\ \lambda = 0 \\ (a,\, b,\, \delta):\, a \geq 0,\, b > 0,\, \delta > 0 \quad iff\ \lambda < 0. \end{cases}$$

For more details on this distribution see Jørgensen (1982). This model includes as special cases the generalized inverse Gaussian distribution for $\delta = 1$, the inverse Gaussian for $\delta = 1$, $\lambda = -1/2$, and the generalized gamma distribution for $\lambda > 0$, $b \to 0$. Now it is straight forward that Exponential, Weibull, gamma distributions are particular cases of generalized gamma distribution for $\lambda = \delta = 1$, $\lambda = \delta$ and $\delta = 1$ respectively.

The mean of the EGIG random variable is

$$E(X) = \left(\sqrt{\frac{b}{a}}\right)^{1/\delta} \frac{K_{(\lambda+1)/\delta}\left(2\sqrt{ab}\right)}{K_{\lambda/\delta}\left(2\sqrt{ab}\right)} = \varphi, \quad (say)\,.$$

Applying the change of variable $\varepsilon_i = \frac{x_i}{\varphi}$, we obtain the unit mean EGIG probability density function

$$f_\varepsilon(\varepsilon_i) = \frac{1}{\left(\frac{2}{\delta}\right)\left(\frac{b\varphi^{-\delta}}{a\varphi^{\delta}}\right)^{\lambda/2\delta} K_{\frac{\lambda}{\delta}}\left(2\sqrt{a\varphi^{\delta} b\varphi^{-\delta}}\right)} \varepsilon_i^{\lambda-1} e^{-(a\varphi^{\delta})\varepsilon_i^{\delta} - (b\varphi^{-\delta})\varepsilon_i^{-\delta}} \tag{6.25}$$

Then the conditional probability density function of x_i given ψ_i is

$$f_{X_i|\psi_i}(X_i) = \left[\left(\frac{2}{\delta}\right)\left(\frac{b\left(\frac{\varphi}{\psi_i}\right)^{-\delta}}{a\left(\frac{\varphi}{\psi_i}\right)^{\delta}}\right)^{\lambda/2\delta} K_{\frac{\lambda}{\delta}}\left(2\sqrt{a\left(\frac{\varphi}{\psi_i}\right)^{\delta} b\left(\frac{\varphi}{\psi_i}\right)^{-\delta}}\right)\right]^{-1} \times x_i^{\lambda-1} e^{-\left[a\left(\frac{\varphi}{\psi_i}\right)^{\delta}\right] x_i^{\delta} - \left[b\left(\frac{\varphi}{\psi_i}\right)^{-\delta}\right] x_i^{-\delta}}. \quad (6.26)$$

That is conditional density of x_i given ψ_i follows $EGIG\left(a\left(\frac{\varphi}{\psi_i}\right)^{\delta}, b\left(\frac{\varphi}{\psi_i}\right)^{-\delta}, \lambda, \delta\right)$.

6.5.1 Special Cases

Accordingly, EGIG distribution includes as special or limiting cases many distributions considered in econometrics and finance. This generalization consists of all the standard ACD models including EACD, WACD, GG-ACD models and IG-ACD model which we described in this chapter. For $\lambda > 0$ and $b \to 0$, EGIG-ACD model reduces to the GG-ACD model proposed by Lunde (1999). The probability density function of a generalized Gamma random variable is given by

$$f(x; a, \lambda, \delta) = \frac{\delta}{\Gamma\left(\frac{\lambda}{\delta}\right)} a^{\frac{\lambda}{\delta}} x^{\lambda-1} e^{-ax^{\delta}} \ ; \ x > 0.$$

Then the conditional density of x_i given ψ_i of GG-ACD is given by

$$f_{X_i|\psi_i}(X_i) = \frac{\delta}{\Gamma\left(\frac{\lambda}{\delta}\right)} \left(a\left(\frac{\varphi}{\psi_i}\right)^{\delta}\right)^{\frac{\lambda}{\delta}} x_i^{\lambda-1} e^{-a\left(\frac{\varphi}{\psi_i}\right)^{\delta} x_i^{\delta}} \ ; \ x_i > 0. \quad (6.27)$$

The generalized gamma family of density functions nests the Weibull distribution and exponential distribution. Both types of distributions have been already successfully applied in ACD framework by Engle and Russell (1998). For $\lambda = \delta = 1$, it reduces to EACD model and for $\lambda = \delta$, it reduces to WACD model. The problem of a flat conditional intensity of EACD was already raised by Engle and Russel as not having a good fit with some semiparametric estimate of the baseline hazard of the data and they therefore propose to extend the EACD model by generalizing the exponential density of the standardized durations to a Weibull density.

For $\delta = 1,\ \lambda = -1/2$, the EGIG-ACD model reduces to IG-ACD model which we have discussed in Section 6.3.

So far we have discussed about the inverse Gaussian ACD model for analysing the financial transaction durations. Next section will discuss about inverse Gaussian SCD model in which the durations are generated by a dynamic stochastic latent variable.

6.6 Inverse Gaussian SCD Model and Properties

Recall the SCD model of order one which we discussed in Chapter 2 as

$$X_i = e^{\psi_i} \varepsilon_i, \qquad \psi_i = \omega + \beta\, \psi_{i-1} + u_i, \tag{6.28}$$

In this section we discuss the model (6.28) when ε_i follows a unit mean IG distribution and $\{u_i\}$ is an independent and identically distributed sequence of $N\left(0, \sigma^2\right)$

random variables. From the definition of the model it follows that $\{\psi_i\}$ is a Gaussian sequence and hence $\{e^{\psi_i}\}$ is a stationary log-normal Markov sequence. Now using the property of log-normal distribution, all the moments of X_i can be computed. In particular the mean and variance are respectively given by

$$E\left(X_i\right) = \exp\left\{\frac{\omega}{1-\beta} + \frac{\sigma^2}{2\left(1-\beta^2\right)}\right\}, \tag{6.29}$$

and

$$Var\left(X_i\right) = \exp\left\{\frac{2\omega}{1-\beta} + \frac{\sigma^2}{1-\beta^2}\right\}\left[\left(1+\frac{1}{\lambda}\right)\exp\left(\frac{\sigma^2}{1-\beta^2}\right) - 1\right]. \tag{6.30}$$

Autocorrelation function:

The k^{th} autocovariance function of X_i is defined as

$$\begin{aligned}\gamma_k &= Cov\left(X_i, X_{i-k}\right)\\ &= E\left(X_i X_{i-k}\right) - E\left(X_i\right)E\left(X_{i-k}\right)\end{aligned}$$

Now we need to compute the expectation of $X_i X_{i-k}$, which is equal to

$$E\left(X_i X_{i-k}\right) = E\left(e^{\psi_i}\varepsilon_i\, e^{\psi_{i-k}}\varepsilon_{i-k}\right)$$

$$= E\left(e^{\psi_i+\psi_{i-k}}\right)E\left(\varepsilon_i\varepsilon_{i-k}\right).$$

From the autoregressive equation of ψ_i, we get

$$\psi_i + \psi_{i-k} = \lambda_{i,k} = 2\omega + \beta\,\lambda_{i-1,k} + u_i + u_{i-k}$$

which is a Gaussian ARMA$(1, k)$ process (with restrictions in the MA polynomial). Unconditionally,

$$e^{\lambda_{i,k}} \sim LN\left(\mu_k, \sigma_k^2\right),$$

where

$$\mu_k = \frac{2\omega}{1-\beta}$$

$$\sigma_k^2 = \frac{2\sigma^2\left(1+\beta^k\right)}{1-\beta^2}.$$

Hence

$$E\left(X_i X_{i-k}\right) = \exp\left\{\frac{2\omega}{1-\beta} + \frac{\sigma^2\left(1+\beta^k\right)}{1-\beta^2}\right\}. \tag{6.31}$$

Therefore, the lag k auto-covariance function of X_i is

$$\gamma_k = \exp\left\{\frac{2\omega}{1-\beta} + \frac{\sigma^2}{1-\beta^2}\right\}\left[\exp\left(\frac{\sigma^2\beta^k}{1-\beta^2}\right) - 1\right]. \tag{6.32}$$

Finally, the k^{th} order autocorrelation function of X_i is $\rho_k = \gamma_k/\gamma_0$ and is given by

$$\rho_k = \frac{\exp\left(\frac{\sigma^2\beta^k}{1-\beta^2}\right) - 1}{\left(1+\frac{1}{\lambda}\right)\left\{\exp\left(\frac{\sigma^2}{1-\beta^2}\right) - 1\right\}} \approx \frac{\sigma^2\beta^k/(1-\beta^2)}{\left(1+\frac{1}{\lambda}\right)\left\{\exp\left(\frac{\sigma^2}{1-\beta^2}\right) - 1\right\}} \approx \beta\,\rho_{k-1}, \tag{6.33}$$

The autocorrelation function ρ_k geometrically decreases at the rate β as the lag k increases.

Hazard Function:

The hazard function $h(.)$ implied by the IG-SCD model can be computed by the formula

$$h\left(x_i\right) = \frac{f\left(x_i\right)}{1 - \int_0^{x_i} f\left(u\right)\, du}, \tag{6.34}$$

where

$$f(x_i) = \frac{\sqrt{\lambda}}{2\pi\sigma\sqrt{x_i^3}} \int_0^\infty \frac{1}{\sqrt{z}} \exp\left\{\frac{-1}{2}\left(\frac{\lambda z\left(\frac{x_i}{z}-1\right)^2}{x_i} + \frac{(\log z - \mu)^2}{\sigma^2}\right)\right\} dz.$$

One way of estimating the hazard function is to replace the parameters in the above expressions by their respective estimates. We do it for the simulated and the real data in Sections 6.8 and 6.9.

6.7 Estimation of IG-SCD Model

A relatively new method for computing the integral needed for evaluating the likelihood function of models with latent variables relies on the efficient importance sampling procedure, recently developed by Richard and Zhang (2007). This method is an extension of the well known importance sampling technique and seems to be particularly well suited for the computation of the multidimensional though relatively well behaved integral needed for evaluation of the SCD likelihood. Given a sequence X of n realizations of the process, with density $g(X|\psi, \theta_1)$ indexed by the parameter vector θ_1, conditional on a vector ψ of a latent variables of the same dimension as X, and given the density $h(\psi|\theta_2)$ indexed by the parameter θ_2, the likelihood function of X can be written as:

$$L(\theta; X) = L(\theta_1, \theta_2; X) = \int g(X|\psi, \theta_1)\; h(\psi|\theta_2)\; d\psi. \qquad (6.35)$$

Actually, the integrand in the previous equation is the joint density $f(X, \psi|\theta)$. Given the assumptions we made, it can be sequentially decomposed as

$$f(X, \psi|\theta) = \prod_{i=1}^{n} d(X_i, \psi_i|X_{i-1}, \psi_{i-1}, \theta) = \prod_{i=1}^{n} p(X_i|\psi_i, \theta_1)\ q(\psi_i|\psi_{i-1}, \theta_2), \quad (6.36)$$

where $p(X_i|\psi_i, \theta_1)$ is obtained from $p(\varepsilon_i)$ (so that θ_1 corresponds to the parameters of inverse Gaussian distribution), and $q(\psi_i|\psi_{i-1}, \theta_2)$ is the Gaussian density $N(\omega + \beta\,\psi_{i-1}, \sigma^2)$ (so that θ_2 includes ω, β and σ^2).

A natural Monte Carlo (MC) estimate of the likelihood function in (6.36) is given by

$$\widetilde{L}(\theta; x) = \frac{1}{S} \sum_{j=1}^{S} \left[\prod_{i=1}^{n} p\left(X_i|\tilde{\psi}_i^{(j)}, \theta_1\right) \right], \quad (6.37)$$

where $\tilde{\psi}_i^{(j)}$ denotes a draw from the density $q\left(\psi_i|\psi_{i-1}^{j}, \theta_2\right)$. This approach bases itself only on the information provided by the distributional assumptions of the model and does not consider the information that comes from the observed sample. It turns out that this estimator is highly inefficient since its sampling variance rapidly increases with the sample size. In any practical case of a duration data set, where the sample size n lies between 500 and 50000 observations, the Monte Carlo sampling size S required to give precise enough estimates of $L(\theta; x)$ would be too high to be affordable and it turns out that this estimator cannot be relied on practically.

EIS tries to make use of the information provided by the observed data in order to come to a reasonably fast and reliable numerical approximation. The principle of EIS is to replace the model-based sampler $\left\{q\left(\psi_i|\psi_{i-1}^{j}, \theta_2\right)\right\}_{i=1}^{n}$ with an optimal auxiliary parametric importance sampler. Let $\{m(\psi_i|\psi_{i-1}, a_i)\}_{i=1}^n$ be a sequence of auxiliary samplers indexed by the set of auxiliary parameter vectors $\{a_i\}_{i=1}^n$. These densities can be defined as a parametric extension of the natural samplers $\left\{q\left(\psi_i|\psi_{i-1}^j, \theta_2\right)\right\}_{i=1}^n$. We rewrite the likelihood function as

$$L(\theta; X) = \int \left[\prod_{i=1}^n \frac{d(X_i, \psi_i|X_{i-1}, \psi_{i-1}, \theta)}{m(\psi_i|\psi_{i-1}, a_i)} \prod_{i=1}^n m(\psi_i|\psi_{i-1}, a_i)\right] d\psi. \tag{6.38}$$

Then, its corresponding IS-MC estimator is given by

$$\tilde{L}(\theta; X, a) = \frac{1}{S}\sum_{j=1}^S \left[\prod_{i=1}^n \frac{d\left(X_i, \tilde{\psi}_i^{(j)}(a_i)|X_{i-1}, \tilde{\psi}_{i-1}^{(j)}(a_{i-1}), \theta\right)}{m\left(\tilde{\psi}_i^{(j)}(a_i)|\tilde{\psi}_{i-1}^{(j)}(a_{i-1}), a_i\right)}\right], \tag{6.39}$$

where $\left\{\left(\tilde{\psi}_i^{(j)}(a_i)\right)\right\}_{i=1}^n$ are trajectories drawn from the auxiliary samplers.

The optimality criterion for choosing the auxiliary samplers is the minimization of the MC variance of (6.39). Relying on the factorized expression of the likelihood, the MC variance minimization problem can be decomposed in a sequence of sub-problems for each element i of the sequence of observations, provided that the elements depending on the lagged values ψ_{i-1} are transferred back to the $(i-1)^{th}$ minimization sub-problem. More precisely, if we decompose m in the product of a function of ψ_i and ψ_{i-1} and one of ψ_{i-1}only, such that

$$m(\psi_i|\psi_{i-1}, a_i) = \frac{k(\psi_i, a_i)}{\chi(\psi_{i-1}, a_i)} = \frac{k(\psi_i, a_i)}{\int k(\psi_i, a_i)\, d\psi_i},$$

we can set up the following minimization problem:

$$\hat{a}_i(\theta) = \arg\min_{a_i} \sum_{j=1}^{S} \left\{ \ln \left[d(X_i, \tilde{\psi}_i^{(j)} | \tilde{\psi}_{i-1}^{(j)}, X_{i-1}, \theta) \chi(\tilde{\psi}_i^{(j)}, \hat{a}_{i+1}) \right] - c_i - \ln(k(\tilde{\psi}_i^{(j)}, a_i)) \right\}^2 \tag{6.40}$$

where c_i is constant that must be estimated along with a_i. If the density kernel $k(\psi_i, a_i)$ belongs to the exponential family of distributions, the problem becomes linear in a_i, and this greatly improves the speed of the algorithm, as a least squares formula can be employed instead of an iterative routine.

The estimated $\hat{a}_i$ are then substituted in (6.39) to obtain the EIS estimate of the likelihood. The EIS algorithm can be initialized by direct sampling, as in Eq. (6.37), to obtain a first series of $\hat{\psi}_i^{(j)}$ and then iterated to allow the convergence of the sequences of $\{a_i\}$, which is usually obtained after 3–5 iterations. EIS-ML estimates are finally obtained by maximizing $\tilde{L}\,(\theta; X, a)$ with respect to θ. Here we adopt a inverse Gaussian distribution for ε_i with parameter λ and a $N(0, \sigma^2)$ for u_i, we come up with the following expressions:

$$p\,(X_i|\psi_{i-1}, \lambda) = \sqrt{\frac{\lambda e^{\psi_i}}{2\pi X_i^3}} \exp\left\{ -\frac{\lambda e^{\psi_i}(X_i - e^{\psi_i})^2}{2(e^{\psi_i})^2 X_i} \right\} \tag{6.41}$$

and

$$q(\psi_i|\psi_{i-1}, \theta_2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2}(\psi_i - \omega - \beta\psi_{i-1})^2 \right\}. \tag{6.42}$$

A convenient choice for the auxiliary sampler $m(\psi_i, a_i)$ is a parametric extension of the natural sampler $q(\psi_i|\psi_{i-1}, \theta_2)$, in order to obtain a good approximation of the integrand without too heavy a cost in terms of analytical complexity. Following Liesenfeld and Richard (2003), we can start by the following specification of the function $k(\psi_i, a_i)$:

$$k(\psi_i, a_i) = q(\psi_i|\psi_{i-1}, \theta_2)\zeta(\psi_i, a_i), \tag{6.43}$$

where $\zeta(\psi_i, a_i) = \exp\{a_{1,i}\psi_i + a_{2,i}\psi_i^2\}$ and $a_i = (a_{1,i}\ a_{2,i})$. This specification is rather straightforward and has two advantages. Firstly, as $q(\psi_i|\psi_{i-1}, \theta_2)$ is present in a multiplicative form, it cancels out in the objective function in (6.40), which becomes a least squares problem with $\ln \zeta(\psi_i, a_i)$ that serves to approximate $\ln p(x_i|\psi_i, \theta_1) + \ln \chi(\psi_i, a_i)$. Secondly, such a functional form for k leads to a distribution of the auxiliary sampler $m(\psi_i, a_i)$ that remains Gaussian, as stated in the following theorem, whose proof is given in Bauwens and Galli (2009).

Theorem 6.1. *If the functional form for $q(\psi_i|\psi_{i-1}, \theta_2)$ and $k(\psi_i, a_i)$ are as in equations (6.42) and (6.43) respectively, then the auxiliary $m(\psi_i|\psi_{i-1}, a_i) = \frac{k(\psi_i, a_i)}{\chi(\psi_{i-1}, a_i)}$ is Gaussian, with conditional mean and variance respectively given by:*

$\mu_i = \upsilon_i^2 \left(\frac{\omega + \beta\psi_{i-1}}{\sigma^2} + a_{1,i}\right)$ *and* $\upsilon_i^2 = \frac{\sigma^2}{1 - 2\sigma^2 a_{2,i}}$,

and the function $\chi(\psi_{i-1}, a_i)$ is given by

$$\frac{1}{\sqrt{1 - 2\sigma^2 a_{2,i}}} \exp\left\{\frac{\sigma^2}{2(1 - 2\sigma^2 a_{2,i})}\left(\frac{\omega + \beta\psi_{i-1}}{\sigma^2} + a_{1,i}\right)^2 - \frac{1}{2}\left(\frac{\omega + \beta\psi_{i-1}}{\sigma}\right)^2\right\}. \tag{6.44}$$

By applying these results, it is possible to compute the likelihood function of the IG-SCD model for a given value of θ, based upon the following steps:

Step 1: Use the natural sampler $q(\psi_i|\psi_{i-1}, \theta_2)$ to draw S trajectories of the latent variable $\{\tilde{\psi}_i^{(j)}\}_{i=1}^n$ as in (6.37).

Step 2: The draws obtained in step 1 are used to solve for each i (in the order from n to 1) the least squares problems described in (6.40), which takes the form of the auxiliary linear regression:

$$\frac{1}{2}\log\lambda - \log(2\pi) + \frac{1}{2}\log e^{\psi_i} - \frac{3}{2}\log X_i - \frac{\lambda e^{\psi_i}(X_i - e^{\psi_i})^2}{2(e^{\psi_i})^2 X_i} + \ln\chi(\tilde{\psi}_i^{(j)}, \hat{a}_{i+1})$$

$$= a_{0,i} + a_{1,i}\tilde{\psi}_i^{(j)} + a_{2,i}(\tilde{\psi}_i^{(j)})^2 + \varepsilon_i^{(i)}, \quad j = 1, 2, ...S,$$

where $\varepsilon_i^{(i)}$ is the error term, $a_{0,i}$ is the constant term, and $\chi(\tilde{\psi}_i^{(j)}, \hat{a}_{t+1})$ is set equal to 1 for $i = n$ and defined by (6.44) for $i < n$. The reverse ordering from n to 1 is due to the fact that for determining $\hat{a}_i$, $\hat{a}_{i+1}$ is required, see (6.40).

Step 3: Use the estimated auxiliary parameters $\hat{a}_i$ to obtain S trajectories $\{\tilde{\psi}_i^{(j)}(\hat{a}_i)\}_{i=1}^N$ from the auxiliary sampler $m(\psi_i|\psi_{i-1}, \hat{a}_i)$, applying the result of Theorem.

Step 4: Return to step 2, this time using the draws obtained with the auxiliary sampler. Steps 2, 3 and 4 are usually iterated a small number of times (from 3 to 5), until a reasonable convergence of the parameters $\hat{a}_i$ is obtained.

Once the auxiliary trajectories have attained a reasonable degree of convergence, the simulated samples can be plugged in formula (6.39) to obtain an EIS estimate of the likelihood. This procedure is embedded in a numerical maximization algorithm that converges to a maximum of the likelihood function. Throughout the EIS steps described above and their iterations, we employed a single set of simulated random numbers to obtain the draws from the auxiliary sampler. This technique, known as common random numbers, is motivated in Richard and Zhang (2007). The same random numbers were also employed for each of the likelihood evaluations required by the maximization algorithm. The number of draws used (S in Eq. (6.39)) for all estimations in this article is equal to 100.

6.8 Simulation Study

A simulation study is carried out here in order to evaluate the performance of the estimation methods proposed for ACD and SCD models with inverse Gaussian innovations.

6.8.1 IG-ACD Model

For the IG-ACD (1,1) model (6.15), we performed the simulation experiment for different sample sizes and for different values of (ω, α, β), fixing $\lambda = 1$. Based on the simulated samples of size n= 1000, 2000, 3000 and 4000, we obtained the ML estimates of λ, ω, α and β. We repeated this computation 100 times and took the average value as the final estimate. These estimates are presented in Table 6.1 with corresponding mean square error in the parentheses.

6.8.2 IG-SCD Model

In this sub-section, we carry out a simulation study to evaluate the performance of the EIS-ML estimation method described in Section 6.7 for the IG-SCD model.

n	True Values $(\lambda, \omega, \alpha, \beta)$	$\hat{\lambda}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
1000	(1.00,0.10,0.10,0.80)	1.0096	0.1231	0.0931	0.7851
		(0.0595)	(0.0805)	(0.0297)	(0.0960)
	(1.00,0.50,0.20,0.60)	1.0088	0.5390	0.1988	0.5848
		(0.0556)	(0.1934)	(0.0374)	(0.0933)
	(1.00,1.00,0.30,0.50)	1.0095	1.0348	0.2961	0.4930
		(0.0527)	(0.3029)	(0.0406)	(0.0775)
	(1.00,1.50,0.50,0.40)	1.0058	1.5288	0.4930	0.4038
		(0.0556)	(0.3539)	(0.0482)	(0.0568)
	(1.00,2.00,0.70,0.10)	1.0119	2.0301	0.6905	0.0997
		(0.0558)	(0.2508)	(0.0452)	(0.0408)
2000	(1.00,0.10,0.10,0.80)	1.0046	0.1085	0.0988	0.7896
		(0.0411)	(0.0723)	(0.0223)	(0.0918)
	(1.00,0.50,0.20,0.60)	1.0040	0.5223	0.1982	0.5924
		(0.0389)	(0.1566)	(0.0226)	(0.0733)
	(1.00,1.00,0.30,0.50)	1.0084	1.0160	0.3018	0.4928
		(0.0390)	(0.2266)	(0.0271)	(0.0581)
	(1.00,1.50,0.50,0.40)	1.0044	1.5169	0.5001	0.3973
		(0.0388)	(0.2553)	(0.0323)	(0.0353)
	(1.00,2.00,0.70,0.10)	1.0030	1.9837	0.6982	0.1018
		(0.0414)	(0.1694)	(0.0355)	(0.0292)
3000	(1.00,0.10,0.10,0.80)	0.9993	0.1028	0.0978	0.7990
		(0.0316)	(0.0647)	(0.0171)	(0.0721)
	(1.00,0.50,0.20,0.60)	1.0059	0.5132	0.1990	0.5939
		(0.0305)	(0.1309)	(0.0220)	(0.0610)
	(1.00,1.00,0.30,0.50)	1.0023	1.0006	0.2993	0.5004
		(0.0343)	(0.1790)	(0.0242)	(0.0463)
	(1.00,1.50,0.50,0.40)	1.0038	1.5086	0.5019	0.3969
		(0.0307)	(0.2140)	(0.0281)	(0.0310)
	(1.00,2.00,0.70,0.10)	1.0039	2.0128	0.6986	0.0982
		(0.0339)	(0.1390)	(0.0303)	(0.0251)
4000	(1.00,0.10,0.10,0.80)	1.0011	0.1066	0.1006	0.7934
		(0.0290)	(0.0627)	(0.0162)	(0.0701)
	(1.00,0.50,0.20,0.60)	1.0032	0.5150	0.1995	0.5925
		(0.0272)	(0.1170)	(0.0177)	(0.0525)
	(1.00,1.00,0.30,0.50)	1.0038	1.0091	0.2987	0.4981
		(0.0273)	(0.1427)	(0.0193)	(0.0373)
	(1.00,1.50,0.50,0.40)	1.0024	1.5090	0.5018	0.3986
		(0.0274)	(0.1819)	(0.0225)	(0.0230)
	(1.00,2.00,0.70,0.10)	1.0003	2.0099	0.6997	0.0987
		(0.0289)	(0.1218)	(0.0270)	(0.0205)

TABLE 6.1: The average ML estimates and the corresponding mean square error for IG-ACD model

n	True values $(\omega, \beta, \sigma, \lambda)$	$\hat{\omega}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\lambda}$
1000	(0.00,0.80,0.50,2.00)	0.0089	0.7814	0.4982	2.0151
		(0.0502)	(0.0567)	(0.0571)	(0.0585)
	(0.00,0.70,0.30,1.50)	0.0088	0.6633	0.2951	1.5175
		(0.0346)	(0.0597)	(0.0660)	(0.0752)
	(0.00,0.50,0.20,1.00)	0.0079	0.4755	0.1850	1.0164
		(0.0382)	(0.0793)	(0.0805)	(0.0756)
	(0.00,0.30,0.05,0.50)	0.0040	0.2977	0.0650	0.5132
		(0.0484)	(0.0802)	(0.0816)	(0.0603)
2000	(0.00,0.80,0.50,2.00)	0.0098	0.7794	0.5010	2.0098
		(0.0507)	(0.0523)	(0.0554)	(0.0606)
	(0.00,0.70,0.30,1.50)	0.0073	0.6678	0.2881	1.4998
		(0.0352)	(0.0539)	(0.0622)	(0.0786)
	(0.00,0.50,0.20,1.00)	0.0044	0.4772	0.1892	1.0075
		(0.0385)	(0.0760)	(0.0767)	(0.0757)
	(0.00,0.30,0.05,0.50)	0.0035	0.3008	0.0626	0.5120
		(0.0504)	(0.0778)	(0.0802)	(0.0597)
3000	(0.00,0.80,0.50,2.00)	0.0118	0.7860	0.5100	2.0133
		(0.0498)	(0.0505)	(0.0565)	(0.0602)
	(0.00,0.70,0.30,1.50)	0.0065	0.6749	0.2974	1.5006
		(0.0348)	(0.0540)	(0.0649)	(0.0811)
	(0.00,0.50,0.20,1.00)	0.0052	0.4834	0.1830	0.9962
		(0.0389)	(0.0744)	(0.0780)	(0.0783)
	(0.00,0.30,0.05,0.50)	0.0021	0.3100	0.0695	0.5147
		(0.0468)	(0.0819)	(0.0817)	(0.0562)
4000	(0.00,0.80,0.50,2.00)	0.0030	0.7825	0.4956	2.0123
		(0.0438)	(0.0496)	(0.0518)	(0.0636)
	(0.00,0.70,0.30,1.50)	0.0029	0.6840	0.2965	1.5007
		(0.0377)	(0.0554)	(0.0625)	(0.0856)
	(0.00,0.50,0.20,1.00)	0.0051	0.4864	0.1908	0.9981
		(0.0312)	(0.0733)	(0.0747)	(0.0828)
	(0.00,0.30,0.05,0.50)	0.0031	0.3034	0.0695	0.5054
		(0.0476)	(0.0818)	(0.0780)	(0.0564)

TABLE 6.2: The average estimates and the corresponding mean square error for the EIS ML estimates

We conducted several repeated simulation experiments with different values of (β, σ, λ), fixing $\omega = 0$. The trajectories of 1000, 2000, 3000 and 4000 observations from a SCD data generating process were simulated 100 times and the model parameters were estimated. These estimates are presented in Table 6.2 with corresponding mean square error in parentheses.

From the table we can see that the EIS-ML method provides estimates which in mean closer to the true parameter values and mean square error of estimates are always remarkably small. The details of computation are given in Appendix E.

The above estimates can be used to evaluate the estimated hazard function through the relation (6.34) to compare with the unconditional empirical hazard function. To obtain the unconditional empirical hazard function we use the relation $\hat{h}(t) = \frac{\hat{f}(t)}{1-\hat{F}(t)}$, where $\hat{f}(t)$ is the kernel based estimator of the marginal probability density function of the IG-SCD sequence and $\hat{F}(t)$ is the empirical distribution function obtained using the relation $\hat{F}(t) = \int_0^t \hat{f}(u)\, du$. We have used formula $\hat{f}(t) = \frac{1}{n\,\Delta t}\sum_{i=1}^{n} \xi\left(\frac{t-T_i}{\Delta t}\right)$ to compute the density estimate in which $\xi(.)$ is the Epanechnikov kernel. See (Silverman (1986),pp 11-13) for details. We computed $\hat{h}(t)$ for the simulated series of IG-SCD sequences for different parameter combinations and compared with the estimated hazard function. Figure 6.2 gives one such graphical comparison, where the dotted line and dashed line represents the unconditional hazard function of the IG-SCD model for different parameter specifications $\omega = 0.10$, $\beta = 0.80$, $\lambda = 2.00$ and its estimated values $\hat{\omega} = 0.1003$, $\hat{\beta} = 0.7895$, $\hat{\lambda} = 2.0145$ respectively. The solid line is the kernel based empirical hazard function of one of the simulated series.

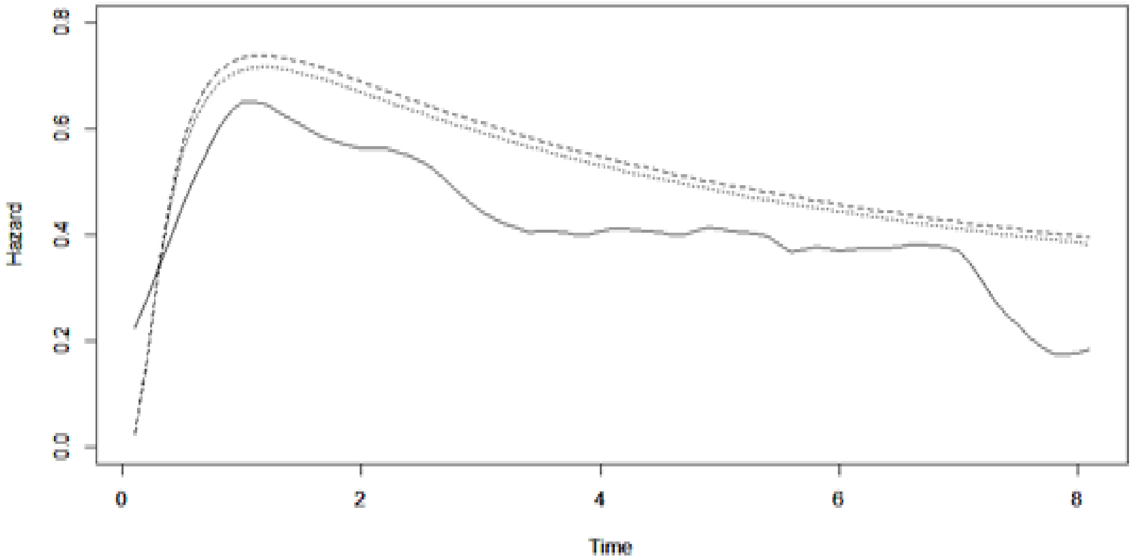

FIGURE 6.2: Unconditional hazard functions of IG-SCD model for simulated data and its empirical hazard rate.

The figure clearly indicates that the empirical hazard function behaves similar to the true as well as the simulation based estimated hazard function.

6.9 Data Analysis

In the present section we apply the inverse Gaussian duration models for analysing the real data sets. We consider the data of intraday foreign exchange rates of Australian Dollar vs Canadian Dollar and US Dollar vs Singapore Dollar for the day 25^{th} April, 2012 and 2^{nd} May, 2012 respectively. The data sets are downloaded from the website of Swiss Forex bank. This is the trade book data (tick-by-tick) corresponding to exchange rate of different currencies traded in Swiss Forex bank.

From the traded data, we took the trade entered time (HH:MM:SS) and find the time duration between the consecutive trades in seconds. The zero durations are excluded. There is a strong seasonality in the durations and we adjusted the data to take care of this diurnal pattern of intraday durations using the method described in Tsay (2005).

Let $f(t_i)$ be the mean value of the diurnal pattern at time t_i. Then define

$$X_i^* = \frac{X_i}{f(t_i)},$$

be the adjusted duration and X_i be the the observed duration i^{th} and $i-1^{th}$ transactions. We construct $f(t_i)$ using two simple time functions.

Define

$$O(t_i) = \begin{cases} t_i - 34200 \; if \; t_i < 43200 \\ 0 \qquad\qquad otherwise, \end{cases}$$

and

$$C(t_i) = \begin{cases} 57600 - t_i \; if \; t_i \geq 43200 \\ 0 \qquad\qquad otherwise, \end{cases}$$

where t_i is the time of i^{th} transaction measured in seconds from midnight and 34200, 43200 and 57600 denote, respectively, market opening, noon and market closing times measured in seconds. Consider the multiple linear regression,

$$\ln(X_i) = \beta_0 + \beta_1 o(t_i) + \beta_2 c(t_i) + e_i,$$

Where $o(t_i) = O(t_i)/10000$ and $c(t_i) = C(t_i)/10000$. Let $\hat{\beta}_i$ be the ordinary least squares estimates of above linear regression. The residual is then given by

$$\hat{e}_i = \ln(X_i) - \hat{\beta}_0 - \hat{\beta}_1 o(t_i) - \hat{\beta}_2 c(t_i).$$

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 series plot of adjusted durations.

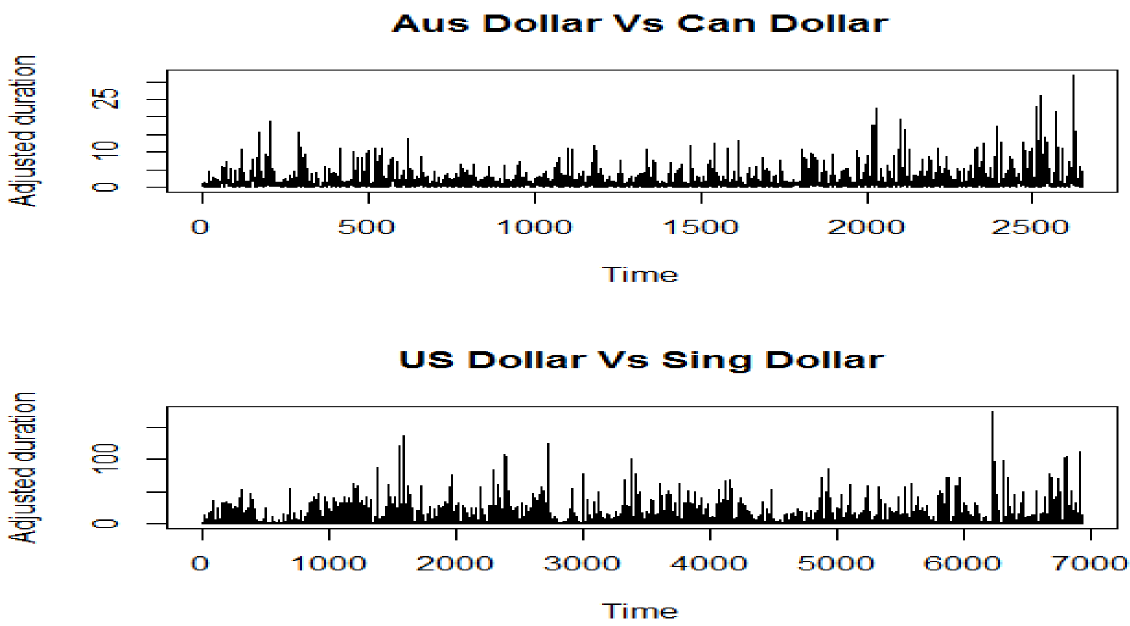

FIGURE 6.3: Time series plot of adjusted durations.

In Table 6.3 we report the summary statistics for the data sets, where $Q(10)$ denotes the Ljung-Box statistic of order 10.

The parameters estimated under IG-ACD(1,1) and IG-SCD models using the methods developed in Sections 6.4 and 6.7 respectively are summarized in Table 6.4.

Statistics	Australian Dollar vs Canadian Dollar	US Dollar vs Singapore Dollar
Sample Size	2650	6933
Minimum	0.1765	0.0546
Maximum	31.9400	175.50
Mean	1.7290	4.5300
Median	1.0210	0.7748
$Q(10)$	36.7616	70.2127

TABLE 6.3: Descriptive statistics for the data

	Aus Dollar vs Can Dollar		US Dollar vs Sing Dollar	
Parameters	IG-ACD	IG-SCD	IG-ACD	IG-SCD
ω	1.4734	0.1890	0.4960	0.2910
α	0.2520	—	0.2527	—
β	0.1574	0.4102	0.2773	0.4507
σ	—	0.4312	—	0.6846
λ	0.6019	0.9011	0.2819	0.9801

TABLE 6.4: Parameter estimates for the data

We now perform a diagnostic check of the models based on the residuals. For IG-ACD model, the standardized residual is defined as $\hat{\varepsilon}_i = x_i/\hat{\psi}_i$. If the fitted ACD model is adequate, then $\{\hat{\varepsilon}_i\}$ should behave as an independent and identically distributed sequence of random variables with the assumed distribution. Since the model assumes that the residuals are independent; any dependence in either the standardized residuals or their squares indicates misspecification of the model. In particular, if the fitted model is adequate, both series $\{\hat{\varepsilon}_i\}$ and $\{\hat{\varepsilon}_i^2\}$ should have no autocorrelations.

Regarding the IG-SCD model, once the estimates of the parameters are obtained, the unobservable component ψ_i is estimated using Kalman filtering. We define the standardized residuals of the IG-SCD model as $\hat{\varepsilon}_i = x_i/e^{\hat{\psi}_i}$, where $\hat{\psi}_i$ is the estimator of ψ_i provided by the Kalman filter at the SML estimate.

Figure 6.4 shows the time series plot of standardized innovations and Figure 6.5 gives the sample autocorrelation function of the standardized innovations for the fitted IG-ACD(1,1) and IG-SCD models respectively. From the figures, the innovations appear to be random and their ACFs fail to indicate any significant serial dependence.

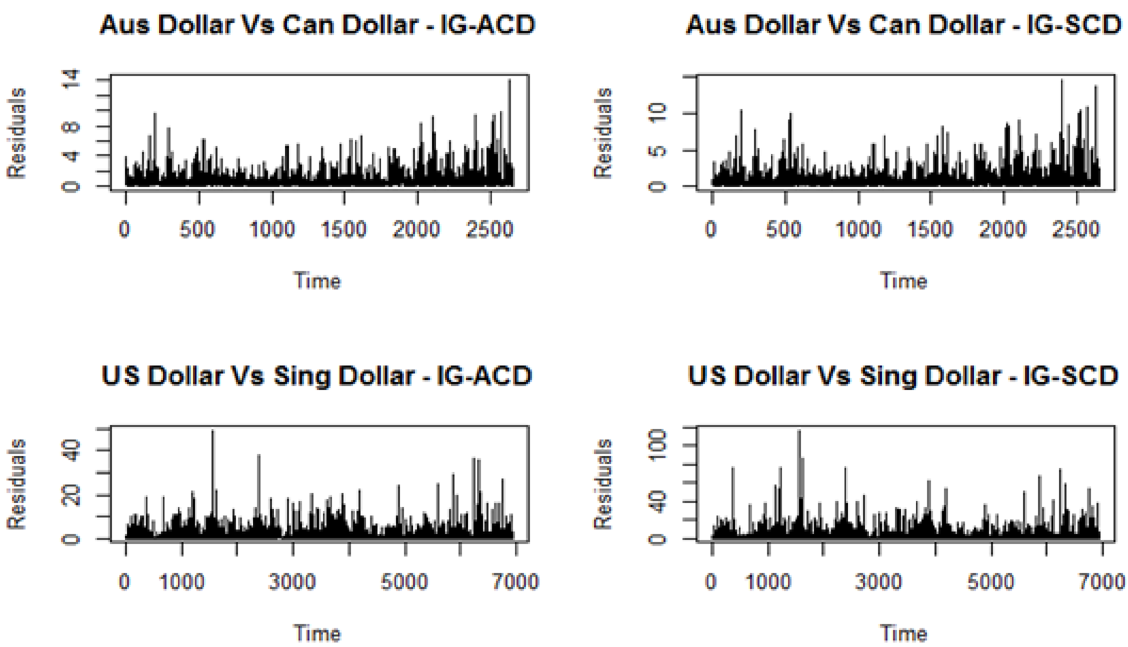

FIGURE 6.4: Time plot of standardized innovation series of IG-ACD(1,1) and IG-SCD models.

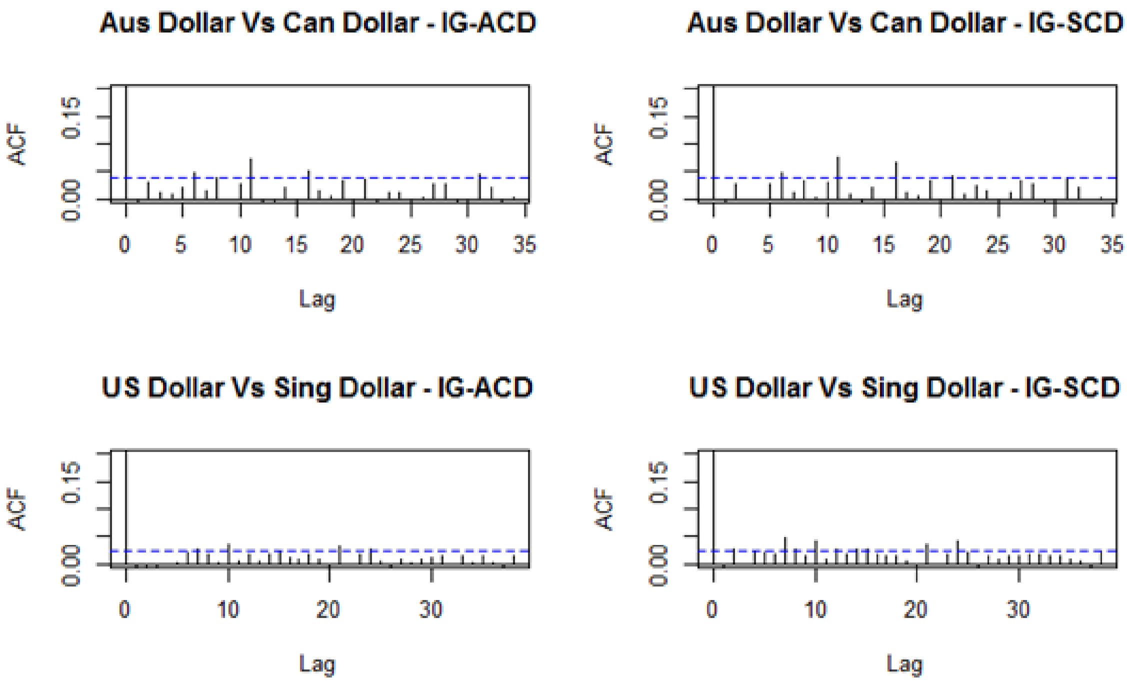

FIGURE 6.5: ACF of the standardized residual series of IG-ACD(1,1) and IG-SCD models.

Figure 6.6 is the histogram of standardized residuals superimposed by the unit mean inverse Gaussian density curve for IG-ACD(1,1) and IG-SCD models. The figures show that the inverse Gaussian distribution is a good approximation for the standardized residuals. Yet, we have to check formally the serial correlations of the series $\{\hat{\varepsilon}_i\}$ and $\{\hat{\varepsilon}_i^2\}$. We adopt Ljung- Box statistics for checking the serial correlations of these two series. The Ljung-Box statistics for the standardized innovations $(Q(.))$ and for the squared innovations $(Q^*(.))$ are calculated and given in Table 6.5. The corresponding Chi-square table values are given in the parenthesis. From the Table 6.5 Ljung-Box statistics for the standardized innovations and the squared innovations are insignificant for both the data sets, so that the fitted models are adequate in describing the dynamic dependence of the adjusted durations.

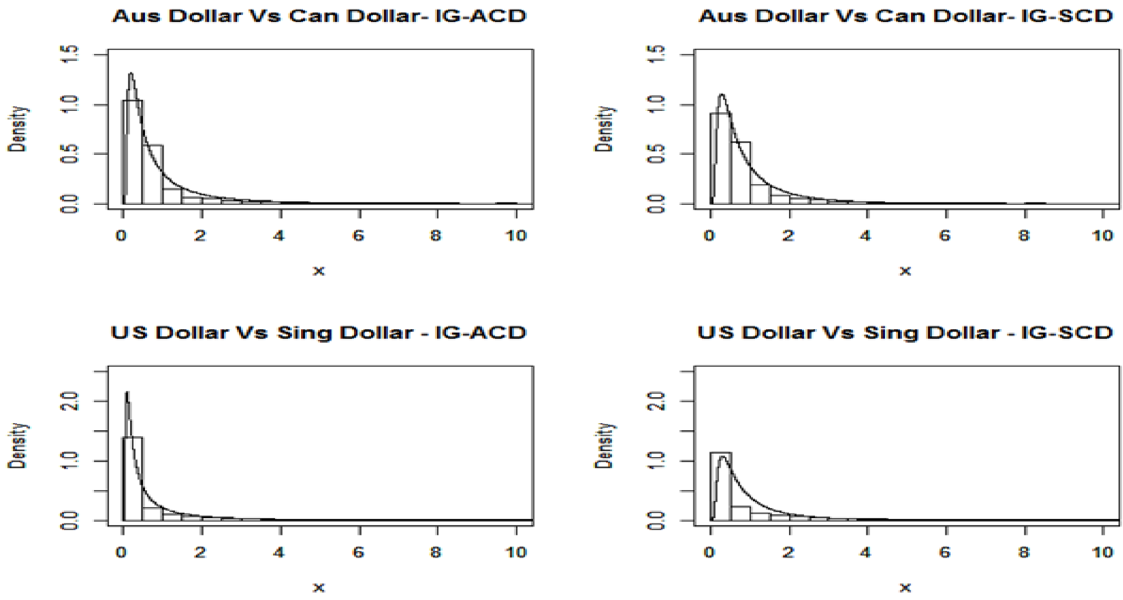

Figure 6.6: Histogram of standardized residuals superimposed by inverse Gaussian density for IG-ACD model and IG-SCD model.

Data	Ljung-Box statistic	IG-ACD	IG-SCD
Aus Dollar vs Can Dollar	$Q(10)$	0.0017(3.3250)	0.0206(3.3250)
	$Q(20)$	3.0725(10.117)	2.7180(10.117)
	$Q^*(10)$	0.1124(3.3250)	0.0906(3.3250)
	$Q^*(20)$	1.3588(10.117)	0.6945(10.117)
US Dollar vs Sing Dollar	$Q(10)$	0.0314(3.3250)	1.5963(3.3250)
	$Q(20)$	0.4805(10.117)	0.2446(10.117)
	$Q^*(10)$	1.2544(3.3250)	1.0017(3.3250)
	$Q^*(20)$	0.0614(10.117)	0.1172(10.117)

Table 6.5: Ljung-Box Statistics for standardized residual series and its squared process with lags 10 and 20.

Finally, we have demonstrated the significance of allowing non-monotonic hazard functions for modelling the conditional durations in financial time series. In Figure 6.7, we have plotted the empirical hazard function of the data and compared it with the unconditional hazard function of IG-SCD model plotted for SML estimates for the data. Thus the use of inverse Gaussian distribution is motivated by the fact that it allows for a non-monotonic hazard function which has been found empirically relevant.

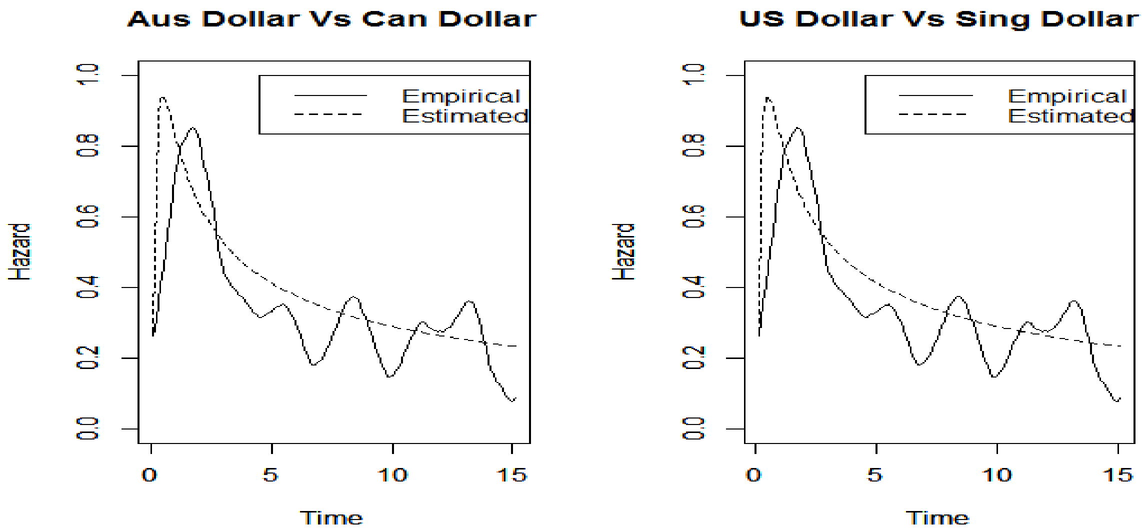

FIGURE 6.7: Empirical and estimated hazard functios of adjusted durations of exchange rates data.

A part of this chapter is published in Balakrishna and Rahul (2014).

Chapter 7

Conclusions and Future Research

The thesis has covered various aspects of modelling and analysis of stochastic volatility and conditional durations in finance. The main objective of analysing financial time series is to model the volatility and forecast its future values. Time series analysis based on Box and Jenkins methods are the most popular approaches when the models are linear and errors are Gaussian. This is considered to be unrealistic in many areas of economics and finance as the distribution of conditional variances are non-Gaussian. As a result, several models have been introduced in the literature to study the behaviour of financial time series. One of the requirements for suggesting new models for stochastic volatility is the existence of class of models for generating non-negative sequences of dependent random variables for generating volatilities. In this thesis, we mainly studied the properties of some non-Gaussian time series models and examined their suitability for modelling stochastic volatility and conditional durations in finance.

With an idea to introduce SV models induced by non-Gaussian volatility sequences, we have proposed a stationary sequence of non-negative random variables with Birnbaum-Saunders marginal distributions. The properties of the model and its estimation procedures have been discussed, and a simulation study has been carried out to evaluate the efficiency of the proposed estimation method. To illustrate the application of the proposed model and the associated estimation method, we have analysed two real data sets. The proposed model provides an additional choice for analysing non-negative time series data. We could establish the asymptotic properties of the estimators for BS-AR(1) model using its Markovian behaviour. Establishing such properties for BS-ARMA models require the development of related theory for non-linear time series models. Construction of BS-ARMA models of higher order and a study of their statistical properties will be of great interest and we hope to consider this for our future work.

The BS distribution is typically applied to positive data with varying degrees of asymmetry and kurtosis and can be used as an alternative to the log-normal and log-skew-normal models. In the fourth chapter, we proposed a SV model generated by first order BS Markov process as an alternative to normal-lognormal SV model. The model parameters are estimated using the method of moments and efficient importance sampling method as the likelihood function is intractable. A simulation experiment is conducted to check the performance of the estimators. The model is used to analyse two sets of data and found that it captures the stylized factors of the financial time series. The problems related to model performance comparison with the existing SV models in terms of in sample fit and out-of-sample forecasting need a separate study. We plan to take this up as our future research.

In Chapter 5, the asymmetric Laplace distribution is used for modelling financial data, which exhibits asymmetry, sharp peaks and heavier tails than normal distribution. We proposed a stationary first order asymmetric Laplace autoregressive model to generate log-volatilities instead of Gaussian AR(1) model. Then we considered stochastic volatility model when the marginal distribution of log-volatility process have an asymmetric Laplace distribution, rather than the Gaussian distribution usually employed. The properties of the AL-SV model and its estimation procedures based on method of moments have been discussed along with a simulation study to evaluate the efficiency of the proposed estimation method. In the empirical study, we adopt AL-SV model to fit the daily returns of exchange rate and future price data.The data analysis illustrates that the AL-SV model is able to capture the skewness and excess kurtosis we observe in financial return series. But we have to come up with more efficient method of estimation and diagnosis procedures for effective use of this model. One practical approach in this context is to develop Bayesian inference procedures. Numerical estimation methods such as Gibbs sampler and Markov chain Monte Carlo procedures will be suitable here. The model presented in this chapter can be extended to analyse multivariate time series. The family of multivariate AL laws can be obtained as a limiting case of the generalized hyperbolic distributions, introduced by Barndorff-Nielsen (1977) and seems to be suitable for modelling heavy tailed asymmetric multivariate data. However, the implementation of multivariate AL-SV model is not straight forward and work is currently under progress in this direction.

Finally, we have proposed two conditional duration models based on inverse Gaussian distribution and studied their properties. We proposed a new specification of the disturbance in the ACD and SCD model by assuming that the standardized durations allow non-monotonic hazard function. The parameters of the IG-ACD model are estimated by maximum likelihood method. A Monte Carlo based efficient importance sampling method is proposed for estimating the parameters of IG-SCD model. The simulation experiments show that the estimates and the resulting estimates of the hazard function perform reasonably well. To illustrate the application of the proposed models, we have analysed two data sets on trade durations of the Australian Dollar/Canadian Dollar and US Dollar/Singapore Dollar exchange and displayed that the proposed models provide a good fit. These results indicate that the inverse Gaussian conditional duration models provide an additional choice for analysing transaction durations in financial point process. Possible extensions to the ACD/SCD models include the use of a wider range of distributional assumptions for conditional durations, in particular those that cater for non-monotonic hazard function of conditional durations. It is important to relax the independent and identically distributed assumption for the innovation, to model higher order conditional moments, and to allow possible regime shifts in price durations.

We conclude this thesis with a note that we have several new problems. Some of these problems can be solved under the Bayesian frame work. The problems related to volatility forecasting and model selection are yet to be discussed. The non-parametric and semi-parametric approaches are potential alternatives to the already established parametric approaches to deal with financial time series. These methods will work better when there are no closed form expressions for likelihood functions. Even though we have focused on discrete time space in our studies so far, the events such as changes in price, temperature, etc. take place continuously. So it is more appropriate to study such problems in continuous time space, which requires the knowledge of stochastic calculus. We would like to tackle these important and interesting, but challenging, problems in future research.

Appendix A

Matlab code for estimation of parameters of BS-ARMA Model

```
ML Estimation (MATLAB Code)
---------------------------------------------
clear;
n1=600;
alpha=input('Enter alpha: ');
beta=input('Enter beta: ');
rho=input('Enter rho: ');
theta=input('Enter theta: ');
sig=(1-rho^2)/(1+theta*theta+2*theta*rho);
bhaa(50)=0;rhaa(50)=0;ahaa(50)=0;thaa(50)=0;
for z=1:50
```

```
tic
ab=normrnd(0,sig,n1,1);
x1(n1)=0;
x1(1)=ab(1);
for t=2:n1
x1(t)=rho*x1(t-1)+theta*ab(t-1)+ab(t);
end
x=x1(101:n1);
y1=(alpha.*x)./2;
y=beta.*power((y1+sqrt(1+power(y1,2))),2);
n=length(y);
%-----------------------%
s1=mean(y);r1=power(mean(1./y),-1);
aa=power((2*(power(s1/r1,0.5)-1)),0.5);
bb=power((s1*r1),0.5);
wt=sqrt(y./bb)-sqrt(bb./y);
tt1(10)=0;
tt1(1)=theta;
for i10=1:10
Rth(n)=0;
R1h(n)=0;
R2h(n)=0;
Rth(1)=wt(1);
for T6=2:n
```

```
su8=0;su9=0;su10=0;su11=0;su12=0;
for k6=1:(T6-1)
su8=su8+(power(-tt1(i10),(k6-1))*wt(T6-k6));
su9=su9+(power(-tt1(i10),(k6-2))*(k6-1)*wt(T6-k6));
su10=su10+(power(-tt1(i10),(k6-1))*k6*wt(T6-k6));
su11=su11+(power(-tt1(i10),(k6-2))*k6*(k6-1)*wt(T6-k6));
su12=su12+(power(-tt1(i10),(k6-3))*(k6-2)*(k6-1)*wt(T6-k6));
end
Rth(T6)=wt(T6)-((rho+tt1(i10))*su8);
R1h(T6)=(rho*su9)-su10;
R2h(T6)=su11-rho*su12;
end
su13=0;su14=0;
for j6=2:n
su13=su13+(Rth(j6)*R1h(j6));
su14=su14+((Rth(j6)*R2h(j6))+power(R1h(j6),2));
end
fth=su13;
fdth=su14;
tt1(i10+1)=tt1(i10)-(fth/fdth);
if tt1(i10+1)-tt1(i10)<0.0001
tt=tt1(i10+1);
end
end
```

```
Rr1(n)=0;
for T4=2:n
su16=0;
for k5=1:(T4-1)
su16=su16+(power(-tt,k5-1)*wt(T4-k5));
end
Rr1(T4)=su16;
end
su17=0;su18=0;
for j5=2:n
su17=su17+(wt(j5)*Rr1(j5));
su18=su18+power(Rr1(j5),2);
end
rr=(su17-tt*su18)/su18;
%------------------------------------%
m=20;
bh(m)=0;th(m)=0;ah(m)=0;rh(m)=0;
bh(1)=bb;
th(1)=tt;
ah(1)=aa;
rh(1)=rr;
for r=1:m
%-----------beta-------------------%
b(10)=0;
```

```
b(1)=bb;
for i=1:10
w=sqrt(y./b(i))-sqrt(b(i)./y);
v=sqrt(y./b(i))+sqrt(b(i)./y);
S(n)=0;
U(n)=0;
S(1)=v(1);
U(1)=w(1);
for T=2:n
sum1=0;
sum2=0;
for k=1:(T-1)
sum1=sum1+(power(-th(r),k-1)*w(T-k));
sum2=sum2+(power(-th(r),k-1)*v(T-k));
end
S(T)=w(T)-(rh(r)+th(r))*sum1;
U(T)=v(T)-(rh(r)+th(r))*sum2;
end
sum3=0;sum4=0;sum5=0;sum6=0;
for j=2:n
sum3=sum3+((b(i)-y(j))/(b(i)+y(j)));
sum4=sum4+(U(j)*S(j));
sum5=sum5+(y(j)/(power(b(i)+y(j),2)));
sum6=sum6+(power(U(j),2)+power(S(j),2));
```

```
end
fbeta=sum3+(((1+th(r)*th(r)+2*th(r)*rh(r))/(ah(r)*ah(r)
       *(1-rh(r)*rh(r))))*sum4);
fdbeta=2*sum5-(((1+th(r)*th(r)+2*th(r)*rh(r))/(2*b(i)
       *ah(r)*ah(r)*(1-rh(r)*rh(r))))*sum6);
b(i+1)=b(i)-(fbeta/fdbeta);
if b(i+1)-b(i)<0.0001
bh1=b(i+1);
end
end
%--------------------------------%
t1(10)=0;
t1(1)=tt;
for i1=1:10
w1=sqrt(y./bh1)-sqrt(bh1./y);
Rt(n)=0;
R1(n)=0;
R2(n)=0;
Rt(1)=w1(1);
for T1=2:n
sum8=0;sum9=0;sum10=0;sum11=0;sum12=0;
for k1=1:(T1-1)
sum8=sum8+(power(-t1(i1),(k1-1))*w1(T1-k1));
sum9=sum9+(power(-t1(i1),(k1-2))*(k1-1)*w1(T1-k1));
```

```
sum10=sum10+(power(-t1(i1),(k1-1))*k1*w1(T1-k1));

sum11=sum11+(power(-t1(i1),(k1-2))*k1*(k1-1)*w1(T1-k1));

sum12=sum12+(power(-t1(i1),(k1-3))*(k1-2)*(k1-1)*w1(T1-k1));

end

Rt(T1)=w1(T1)-((rh(r)+t1(i1))*sum8);

R1(T1)=(rh(r)*sum9)-sum10;

R2(T1)=sum11-rh(r)*sum12;

end

sum13=0;sum14=0;

for j1=2:n

sum13=sum13+(Rt(j1)*R1(j1));

sum14=sum14+((Rt(j1)*R2(j1))+power(R1(j1),2));

end

ftheta=sum13;

fdtheta=sum14;

t1(i1+1)=t1(i1)-(ftheta/fdtheta);

if t1(i1+1)-t1(i1)<0.0001

th1=t1(i1+1);

end

end

%---------------------------------%

w2=sqrt(y./bh1)-sqrt(bh1./y);

Rr(n)=0;

for T2=2:n
```

```
sum16=0;
for k2=1:(T2-1)
sum16=sum16+(power(-th1,k2-1)*w2(T2-k2));
end
Rr(T2)=sum16;
end
sum17=0;sum18=0;
for j2=2:n
sum17=sum17+(w2(j2)*Rr(j2));
sum18=sum18+power(Rr(j2),2);
end
rh1=(sum17-th1*sum18)/sum18;
%-----------------------------%
w3=sqrt(y./bh1)-sqrt(bh1./y);
Ra(n)=0;
for T3=2:n
sum19=0;
for k3=1:(T3-1)
sum19=sum19+(power(-th1,k3-1)*w3(T3-k3));
end
Ra(T3)=(w3(T3))-((rh1+th1)*sum19);
end
sum20=0;
for j3=2:n
```

```
sum20=sum20+power(Ra(j3),2);
end
ah1=sqrt(((1+th1*th1+2*rh1*th1)/((n-1)*(1-rh1*rh1)))*sum20);
%----------------------------%
ah(r+1)=ah1;
bh(r+1)=bh1;
th(r+1)=th1;
rh(r+1)=rh1;
if ah(r+1)-ah(r)<0.0001
ahh=ah(r+1);
end
if bh(r+1)-bh(r)<0.0001
bhh=bh(r+1);
end
if th(r+1)-th(r)<0.0001
thh=th(r+1);
end
if rh(r+1)-rh(r)<0.0001
rhh=rh(r+1);
end
end
ahaa(z)=ahh;
bhaa(z)=bhh;
thaa(z)=thh;
```

```
rhaa(z)=rhh;

toc

end

ahat=mean(ahaa)

ahatvar=sqrt(var(ahaa))

bhat=mean(bhaa)

bhatvar=sqrt(var(bhaa))

rhat=mean(rhaa)

rhatvar=sqrt(var(rhaa))

that=mean(thaa)

thatvar=sqrt(var(thaa))


---------------------------------------------------------
```

Appendix B

R code for computation of I_1

```
f12=0
I1=0
acf=0
r=-0.8
a=2
x1=seq(-6,6,0.01)
n1=length(x1)
x2=seq(-6,6,0.01)
n2=length(x2)
R=20
acf=rep(1,R)
I1=rep(1,R)
for(h in 1:R)
{
```

```
xr=(r^h)*x1

f12=rep(1,n1)

for(i in 1:n1)

{

f12[i]=0.01*sum(x2*sqrt((1+(0.5*a*x2)^2))*dnorm(x2,xr[i],1-r^(2*h)))

}

I1[h]=0.01*sum(x1*sqrt((1+(0.5*a*x1)^2))*dnorm(x1,0,1)*f12)

}

I1

--------------------------------------------------------------------------------
```

Appendix C

R code for estimation of parameters of BS-SV Model

```
1. MM estimation(R Code)
---------------------------
T=600;
alpha=2;
beta=1;
rho=0.9;
sigma=sqrt(1-rho^2);
%--------------------------------------------------%
ala=c();bta=c();rha=c();
for(q in 1:50){
ui=rnorm(T,0,sigma);
```

```
h1=c();
h1[1]=ui[1];
for(t in 2:T){
h1[t]=rho*h1[t-1]+ui[t];
}
h=h1[101:T]
n=length(T)
ab=rnorm(n,0,1);
y=beta*(((alpha*h)/2)+sqrt(1+((alpha*h)/2)^2))^2;
x=sqrt(y)*ab;
m2=mean(x[2:n]*x[2:n])
m4=mean(x[2:n]*x[2:n]*x[2:n]*x[2:n])
m22=mean(x[2:n]*x[2:n]*x[1:(n-1)]*x[1:(n-1)])
#########################
aa=c()
aa[1]=2
for(k in 1:10){
fa=(((1+(0.5*aa[k]*aa[k]))^2)/(3*(1+(2*aa[k]*aa[k])
   +(1.5*aa[k]*aa[k]*aa[k]*aa[k]))))-((m2*m2)/m4)
fda=((6*aa[k]*(1+(2*aa[k]*aa[k])+(1.5*aa[k]*aa[k]*aa[k]*aa[k]))
  *(1+(0.5*aa[k]*aa[k])))-(3*((1+(0.5*aa[k]*aa[k]))^2)*(4*aa[k]
     +6*aa[k]^3)))/(9*(1+(2*aa[k]*aa[k])+(1.5*aa[k]*aa[k]
    *aa[k]*aa[k]))^2)
aa[k+1]=aa[k]-(fa/fda)
```

```
if (aa[k+1]-aa[k])>0.001 ahat=aa[k+1]

}

bhat=m2/(1+(0.5*ahat*ahat))

r=c()

r[1]=0.1

for(j in 1:10){

f12=0

a=ahat

x1=seq(-6,6,0.01)

n1=length(x1)

x2=seq(-6,6,0.01)

n2=length(x2)

xr=r[j]*x1

f12=rep(1,n1)

for(i in 1:n1)

{

f12[i]=0.01*sum(x2*sqrt((1+(0.5*a*x2)^2))
      *dnorm(x2,xr[i],1-r[j]^(2)))

}

I1=0.01*sum(x1*sqrt((1+(0.5*a*x1)^2))*dnorm(x1,0,1)*f12)

while ((m22-((bhat*bhat)*(1+a*a+((1/4)*aa*aa*aa*aa*(1+(2*r[j]*r[j])))
      +(aa*aa*I1))))>0.001) r=r+0.001

r[j+1]=r[j]

if ((r[j+1]-r[j])>0.001) rhat=r[j+1]
```

```
}
ala[q]=ahat
bta[q]=bhat
rha[q]=rhat
}
mean(ala); sqrt(var(ala))
mean(bta); sqrt(var(bta))
mean(rha); sqrt(var(rha))

-------------------------------------------------
```

2. EIS ML Estimation(MATLAB Code)

```
---------------------------
clear;
T=300;
N=100;
alpha=input('enter alpha:');
beta=input('enter beta:');
rho=input('enter rho:');
sigma=sqrt(1-power(rho,2));
rh=rho-0.1:0.001:rho+0.09;
ap=alpha-0.2:0.001:alpha+0.2;
bt=beta-0.2:0.001:beta+0.2;
%-------------------------------------------------%
ala(50)=0;bta(50)=0;rha1(50)=0;
for rahul1=1:50
tic
ui=normrnd(0,sigma,1,T);
sh(T)=0;
sh(1)=ui(1);
for k=2:T
sh(k)=rho*sh(k-1)+ui(k);
end
ab=normrnd(0,1,1,T);
yy=beta.*power(((alpha.*sh)/2)+sqrt(1+power(((alpha.*sh)/2),2)),2);
```

```
x=sqrt(yy).*ab;
%-------------------initial values -------------------------------%
rt=mean(x.*x.*x.*x);
rt1=mean(x.*x);
rt2=rt/(rt1*rt1);
al1=0.1;
while rt2-rahul(al1)>0.001
al1=al1+0.0001;
end
be(1)=rt1/(1+((al1*al1)/2));
al(1)=al1;
c=autocorr(sh);
rh1(1)=c(2);
%----------------------------------------------------------------%
for n=1:10
%------------------- rho -----------------------%
r1=length(rh);
L(r1)=0;
for r=1:r1
u=normrnd(0,1,T,N);
si(T,N)=0;
for j=1:N
si(1,j)=u(1,j);
for i=2:T
```

```
sig=(1-power(rh(r),2));
mu=rh(r)*si(i-1,j);
si(i,j)=mu+sqrt(sig)*u(i,j);
end
end
for q=1:6
a1(T,2)=0;
ch(T,N)=0;
ch(T,:)=1;
for h=T:-1:1
y=-0.5*log(2*pi*be(n))-log(((al(n).*si(h,:)')/2)
    +sqrt(1+power(((al(n).*si(h,:)')/2),2)))-0.5*((x(h)*x(h))
    ./( be(n).*power(((al(n).* si(h,:)')/2)+sqrt(1+power(((al(n).*
     si(h,:)')/2),2)),2)))+(log(ch(h,:)))';
y1=si(h,:)';
p=polyfit(y1,y,2);
a1(h,:)=[p(2),p(1)];
if(h>1)
ch(h-1,:)=(1/sqrt(1-(2*(1-power(rh(r),2))*p(1))))
         .*exp((((1-power(rh(r),2))/(2*(1-(2*(1-power(rh(r),2))
         *p(1))))).*power(((rh(r)*si(h-1,:)')./(1-power(rh(r),2)))
         +p(2),2))-(0.5*power((rh(r)*si(h-1,:)')./sqrt(1
         -power(rh(r),2)),2)));
end
```

```
end
s1(T,N)=0;
for l=1:N
s1(1,l)=u(1,l);
for k=2:T
sig1=(1-power(rh(r),2))/(1-(2*(1-power(rh(r),2))*a1(k,2)));
mu1=sig1*(((rh(r)*s1(k-1,l))/(1-power(rh(r),2)))+a1(k,1));
s1(k,l)=mu1+sqrt(sig1)*u(k,l);
end
end
si=s1;
a(:,:,q)=a1;
end
b=a(:,:,6);
z=s1;
p1(T,N)=0;
q1(T,N)=0;
m1(T,N)=0;
p1(1,:)=(1./(sqrt(2*pi*be(n)).*(((al(n).*z(1,:))/2)
       +sqrt(1+power(((al(n).*z(1,:))/2),2))))).*exp(-0.5*((x(1)*x(1))
       ./(be(n).*power(((al(n).*z(1,:))/2)+sqrt(1+power(((al(n)
       .*z(1,:))/2),2)),2))));
q1(1,:)=(1/(sqrt(1-power(rh(r),2))*sqrt(2*pi)))
        .*(exp(-((1/(2*(1-power(rh(r),2))))
```

```
        .*power((z(1,:)-rh(r)*u(1,:)),2))));
m1(1,:)=((sqrt(1-(2*(1-power(rh(r),2))*b(1,2))))
        /(sqrt(2*pi*(1-power(rh(r),2)))))
        *exp(-((((1-(2*(1-power(rh(r),2))*b(1,2)))
        /(2*(1-power(rh(r),2))))*power((z(1,:)
        -(((1-power(rh(r),2))/(1-(2*(1-power(rh(r),2))
        *b(1,2))))*(((rh(r)*u(1,:))/(1-power(rh(r),2)))+b(1,1)))),2)));
for w=2:T
p1(w,:)=(1./(sqrt(2*pi*be(n)).*(((al(n).*z(w,:))/2)
       +sqrt(1+power(((al(n).*z(w,:))/2),2))))).*exp(-0.5
       *((x(w)*x(w))./(be(n).*power(((al(n).*
        z(w,:))/2)+sqrt(1+power(((al(n).*z(w,:))/2),2)),2))));
q1(w,:)=(1/(sqrt(1-power(rh(r),2))*sqrt(2*pi)))
      .*(exp(-((1/(2*(1-power(rh(r),2)))).*power((z(w,:)
      -rh(r)*z(w-1,:)),2))));
m1(w,:)=(sqrt(1-(2*(1-power(rh(r),2))*b(w,2)))
        /(sqrt(2*pi*(1-power(rh(r),2)))))*exp(-(((1-(2
        *(1-power(rh(r),2))*b(w,2)))/(2*(1-power(rh(r),2))))
        *power((z(w,:)-(((1-power(rh(r),2))/(1-(2*(1
        -power(rh(r),2))*b(w,2))))*(((rh(r)*z(w-1,:))/
        (1-power(rh(r),2)))+b(w,1)))),2)));
end
d=(p1.*q1)./m1;
L(r)=sum(prod(d));
```

```
end
Lr=max(L);
e=1;
while(Lr~=L(e))
e=e+1;
end
rhh=rh(e);
%---------------------------end estimation of rho--------------------%
%-------------------alpha-----------------------%
r2=length(ap);
L1(r2)=0;
for rr=1:r2
u1=normrnd(0,1,T,N);
si1(T,N)=0;
for j=1:N
si1(1,j)=u1(1,j);
for i=2:T
sig2=(1-power(rhh,2));
mu2=rhh*si1(i-1,j);
si1(i,j)=mu2+sqrt(sig2)*u1(i,j);
end
end
for q=1:6
a2(T,2)=0;
```

```
ch1(T,N)=0;
ch1(T,:)=1;
for h=T:-1:1
y2=-0.5*log(2*pi*be(n))-log(((ap(rr).*si1(h,:)')/2)
    +sqrt(1+power(((ap(rr).*si1(h,:)')/2),2)))-0.5*((x(h)*x(h))
    ./( be(n).*power(((ap(rr).* si1(h,:)')/2)+sqrt(1+power(((ap(rr).*
     si1(h,:)')/2),2)),2)))+(log(ch1(h,:)))';
y3=si1(h,:)';
pa=polyfit(y3,y2,2);
a2(h,:)=[pa(2),pa(1)];
if(h>1)
ch1(h-1,:)=(1/sqrt(1-(2*(1-power(rhh,2))*pa(1))))
         .*exp((((1-power(rhh,2))/(2*(1-(2*(1-power(rhh,2))*pa(1)))))
       .*power(((rhh*si1(h-1,:)')./(1-power(rhh,2)))+pa(2),2))
        -(0.5*power((rhh*si1(h-1,:)')./sqrt(1-power(rhh,2)),2)));
end
end
s11(T,N)=0;
for l=1:N
s11(1,l)=u1(1,l);
for k=2:T
sig3=(1-power(rhh,2))/(1-(2*(1-power(rhh,2))*a2(k,2)));
mu3=sig3*(((rhh*s11(k-1,l))/(1-power(rhh,2)))+a2(k,1));
s11(k,l)=mu3+sqrt(sig3)*u1(k,l);
```

```
end
end
si1=s11;
aa(:,:,q)=a2;
end
b1=aa(:,:,6);
z1=s11;
p2(T,N)=0;
q2(T,N)=0;
m2(T,N)=0;
p2(1,:)=(1./(sqrt(2*pi*be(n)).*(((ap(rr).*z1(1,:))/2)
        +sqrt(1+power(((ap(rr).*z1(1,:))/2),2)))))
        .*exp(-0.5*((x(1)*x(1))./(be(n).*power(((ap(rr).*
         z1(1,:))/2)+sqrt(1+power(((ap(rr).*z1(1,:))/2),2)),2))));
q2(1,:)=(1/(sqrt(1-power(rhh,2))*sqrt(2*pi)))
        .*(exp(-((1/(2*(1-power(rhh,2)))).*power((z1(1,:)
        -rhh*u1(1,:)),2))));
m2(1,:)=((sqrt(1-(2*(1-power(rhh,2))*b1(1,2))))/
       (sqrt(2*pi*(1-power(rhh,2)))))*exp(-(((1-(2*(1-power(rhh,2))
       *b1(1,2)))/(2*(1-power(rhh,2))))*power((z1(1,:)
       -(((1-power(rhh,2))/(1-(2*(1-power(rhh,2))*b1(1,2))))
       *(((rhh*u1(1,:))/(1-power(rhh,2)))+b1(1,1)))),2)));
for w=2:T
p2(w,:)=(1./(sqrt(2*pi*be(n)).*(((ap(rr).*z1(w,:))/2)
```

```
         +sqrt(1+power(((ap(rr).*z1(w,:))/2),2))))).*exp(-0.5
         *((x(w)*x(w))./(be(n).*power(((ap(rr).*
          z1(w,:))/2)+sqrt(1+power(((ap(rr).*z1(w,:))/2),2)),2))));
q2(w,:)=(1/(sqrt(1-power(rhh,2))*sqrt(2*pi))).*(exp(-((1
        /(2*(1-power(rhh,2)))).*power((z1(w,:)-rhh*z1(w-1,:)),2))));
m2(w,:)=(sqrt(1-(2*(1-power(rhh,2))*b1(w,2)))/
        (sqrt(2*pi*(1-power(rhh,2)))))*exp(-(((1-(2*(1-power(rhh,2))
        *b1(w,2)))/(2*(1-power(rhh,2))))*power((z1(w,:)
        -(((1-power(rhh,2))/(1-(2*(1-power(rhh,2))*b1(w,2))))
        *(((rhh*z1(w-1,:))/ (1-power(rhh,2)))+b1(w,1)))),2)));
end
d1=(p2.*q2)./m2;
L1(rr)=sum(prod(d1));
end
La=max(L1);
e1=1;
while(La~=L1(e1))
e1=e1+1;
end
ahh=ap(e1);
%-----------------------------end estimation of alpha --------------------%
%------------------------------------beta-------------------------%
r3=length(bt);
L2(r3)=0;
```

```
for rr1=1:r3
u2=normrnd(0,1,T,N);
si2(T,N)=0;
for j=1:N
si2(1,j)=u2(1,j);
for i=2:T
sig4=(1-power(rhh,2));
mu4=rhh*si2(i-1,j);
si2(i,j)=mu4+sqrt(sig4)*u2(i,j);
end
end
for q=1:6
a3(T,2)=0;
ch2(T,N)=0;
ch2(T,:)=1;
for h=T:-1:1
y4=-0.5*log(2*pi*bt(rr1))-log(((ahh.*si2(h,:)')/2)
    +sqrt(1+power(((ahh.*si2(h,:)')/2),2)))-0.5*((x(h)*x(h))
    ./( bt(rr1).*power(((ahh.* si2(h,:)')/2)+sqrt(1+power(((ahh.*
    si2(h,:)')/2),2)),2)))+(log(ch2(h,:)))';
y5=si2(h,:)';
pb=polyfit(y5,y4,2);
a3(h,:)=[pb(2),pb(1)];
if(h>1)
```

```
ch2(h-1,:)=(1/sqrt(1-(2*(1-power(rhh,2))*pb(1))))
          .*exp((((1-power(rhh,2))/(2*(1-(2*(1-power(rhh,2))*pb(1)))))
          .*power(((rhh*si2(h-1,:)')./(1-power(rhh,2)))+pb(2),2))
          -(0.5*power((rhh*si2(h-1,:)')./sqrt(1-power(rhh,2)),2)));
end
end
s12(T,N)=0;
for l=1:N
s12(1,l)=u2(1,l);
for k=2:T
sig5=(1-power(rhh,2))/(1-(2*(1-power(rhh,2))*a3(k,2)));
mu5=sig5*(((rhh*s12(k-1,l))/(1-power(rhh,2)))+a3(k,1));
s12(k,l)=mu5+sqrt(sig5)*u2(k,l);
end
end
si2=s12;
aa1(:,:,q)=a3;
end
b2=aa1(:,:,6);
z2=s12;
p3(T,N)=0;
q3(T,N)=0;
m3(T,N)=0;
p3(1,:)=(1./(sqrt(2*pi*bt(rr1)).*(((ahh.*z2(1,:))/2)
```

```
        +sqrt(1+power(((ahh.*z2(1,:))/2),2))))).*exp(-0.5*((x(1)*x(1))
        ./(bt(rr1).*power(((ahh.*z2(1,:))/2)+sqrt(1+power(((ahh.*z2(1,:))
        /2),2)),2))));
q3(1,:)=(1/(sqrt(1-power(rhh,2))*sqrt(2*pi))).*(exp(-((1/(2*
       (1-power(rhh,2)))).*power((z2(1,:)-rhh*u2(1,:)),2))));
m3(1,:)=((sqrt(1-(2*(1-power(rhh,2))*b2(1,2))))/(sqrt(2*pi
         *(1-power(rhh,2)))))*exp(-(((1-(2*(1-power(rhh,2))*b2(1,2)))
         /(2*(1-power(rhh,2))))*power((z2(1,:)-(((1-power(rhh,2))
         /(1-(2*(1-power(rhh,2))*b2(1,2))))*(((rhh*u2(1,:))
         /(1-power(rhh,2)))+b2(1,1)))),2)));
for w=2:T
p3(w,:)=(1./(sqrt(2*pi*bt(rr1)).*(((ahh.*z2(w,:))/2)
        +sqrt(1+power(((ahh.*z2(w,:))/2),2))))).*exp(-0.5
        *((x(w)*x(w))./(bt(rr1).*power(((ahh.*z2(w,:))/2)
        +sqrt(1+power(((ahh
        .*z2(w,:))/2),2)),2))));
q3(w,:)=(1/(sqrt(1-power(rhh,2))*sqrt(2*pi)))
        .*(exp(-((1/(2*(1-power(rhh,2)))).*power((z2(w,:)
        -rhh*z2(w-1,:)),2))));
m3(w,:)=(sqrt(1-(2*(1-power(rhh,2))*b2(w,2)))/(sqrt(2*pi
          *(1-power(rhh,2)))))*exp(-(((1-(2*(1-power(rhh,2))*b2(w,2)))
          /(2*(1-power(rhh,2))))*power((z2(w,:)-(((1-power(rhh,2))
          /(1-(2*(1-power(rhh,2))*b2(w,2))))*(((rhh*z2(w-1,:))/
           (1-power(rhh,2)))+b2(w,1)))),2)));
```

```
end
d2=(p3.*q3)./m3;
L2(rr1)=sum(prod(d2));
end
Lb=max(L2);
e2=1;
while(Lb~=L2(e2))
e2=e2+1;
end
bhh=bt(e2);
%----------------------------end estimation of beta ------------%
al(n+1)=ahh;
if al(n+1)-al(n)<0.001
aha=al(n+1);
end
be(n+1)=bhh;
if be(n+1)-be(n)<0.001
bha=be(n+1);
end
rh1(n+1)=rhh;
if rh1(n+1)-rh1(n)<0.001
rha=rh1(n+1);
end
end
```

```
ala(rahul1)=aha;

bta(rahul1)=bha;

rha1(rahul1)=rha;

toc

end

est=[mean(ala) mean(bta) mean(rha1)]

stdrr=[sqrt(var(ala)) sqrt(var(bta)) sqrt(var(rha1))]

--------------------------------------------------------------
```

Appendix D

R code for estimation of parameters of AL-SV Model

```
MM Estimation (R Code)
-----------------------
n1=500
mu=2
sigma=2
theta=0.5
Rho=0.9
kappa=((sqrt((mu^2)+(2*sigma^2)))-mu)/(sigma*sqrt(2));
kkk=c();sss=c();rrr=c();
for(p in 1:50){
###############
```

```
rahul.pdf=function(x,y,rho,K)
{
if(x==0 && y==0){prob=rho^2}
if(x==1 && y==1){prob=0}
if(x==0 && y==1){prob=(1-rho)*(rho+(1-rho)*K^2/(1+K^2))}
if(x==1 && y==0){prob=(1-rho)*(rho+(1-rho)/(1+K^2))}

return(prob)
}

sampleGen=function(n,rho,K)
{
s=data.frame(x=numeric(n),y=numeric(n))
i=0
c1=rahul.pdf(0,0,rho,K)
c2=rahul.pdf(0,1,rho,K)
c3=rahul.pdf(1,0,rho,K)
c4=rahul.pdf(1,1,rho,K)
while(i<n){
rnd=runif(1)
 if(rnd<c1){
xval=0
yval=0
s[i,]=c(xval,yval)
```

```
i=i+1
}
else if(rnd<c1+c2){
xval=0
yval=1
s[i,]=c(xval,yval)
i=i+1
}
else  if(rnd<c1+c2+c3){
xval=1
yval=0
s[i,]=c(xval,yval)
i=i+1
}
else{
xval=1
yval=1
s[i,]=c(xval,yval)
i=i+1
}

}
return(s)
}
```

```
d=sampleGen(n1,Rho,kappa)

delta1=d[1]

delta2=d[2]

w1=rexp(n1,1)

w2=rexp(n1,1)

eta1=((1-Rho)*theta)+((sigma/sqrt(2))*(((delta1*w1)/kappa)

     -(kappa*delta2*w2)))

eta=eta1$x

h=c()

h[1]=eta[1]

for(t in 2:n1){

h[t]=Rho*h[t-1]+eta[t]

}

e1=rexp(n1,1)

e2=rexp(n1,1)

absl=(1/sqrt(2))*(e1-e2)

rt=(exp(h/2))*absl

m1=mean(rt[2:n1]*rt[2:n1])

m2=mean(rt[2:n1]*rt[2:n1]*rt[2:n1]*rt[2:n1])

m3=mean(rt[2:n1]*rt[2:n1]*rt[2:n1]*rt[2:n1]*rt[2:n1]*rt[2:n1])

m4=mean(rt[2:n1]*rt[2:n1]*rt[1:(n1-1)]*rt[1:(n1-1)])


########################

ka=c();sa=c()
```

```
ka[1]=kappa
sa[1]=sigma
for(q in 1:10)
{
ss=c()
ss[1]=1
for(i in 1:10){
fs=(1/m1)-1+(0.5*ss[i]*ss[i])+((ss[i]/sqrt(2))*((1/ka[q])-ka[q]))
fds=ss[i]+((1/sqrt(2))*((1/ka[q])-ka[q]))
ss[i+1]=(ss[i]-(fs/fds))
if ((ss[i+1]-ss[i])<0.001) sighat1=ss[i+1]
}
kk=c()
kk[1]=1
for(j in 1:10){
fk=(6/m2)-1+(2*sighat1*sighat1)+(sqrt(2)*sighat1*((1/kk[j])-kk[j]))
fdk=-(sqrt(2)*sighat1*(1+(1/(kk[j]*kk[j]))))
kk[j+1]=(kk[j]-(fk/fdk))
if ((kk[j+1]-kk[j])<0.001) khat1=kk[j+1]
}
ka[q+1]=khat1
sa[q+1]=sighat1
if ((ka[q+1]-ka[q])<0.001) khat=ka[q+1]
if ((sa[q+1]-sa[q])<0.001) shat=sa[q+1]
```

```
}
g1=(1-(0.5*shat*shat)+(0.5*shat*((1/khat)-khat)*sqrt(2)))
g2=(shat/sqrt(2))*((1/khat)-khat)
g3=shat*shat/2
rh=c()
rh[1]=0.5
for(i in 1:20){
f1=m3-((1-(g3*rh[i]*rh[i])+(g2*rh[i]))/(g1*(1-(g3*(1+rh[i])^2)
  +(g2*(1+rh[i])))))
f2=-(1/(g1*(1-(g3*(1+rh[i])^2)+(g2*(1+rh[i])))^2))*(((1
   -(g3*(1+rh[i])^2)+(g2*(1+rh[i])))*(g2-(shat*shat*rh[i])))
   -((1-(g3*(rh[i])^2)+(g2*(rh[i])))*(g2-(shat*shat*(1+rh[i])))))
rh[i+1]=rh[i]-(f1/f2)
if ((rh[i+1]-rh[i])<0.001) rhat=rh[i+1]
}
kkk[p]=khat
sss[p]=shat
rrr[p]=rhat
}
mean(kkk);  sqrt(var(kkk))
mean(sss);  sqrt(var(sss))
mean(rrr);  sqrt(var(rrr))
-----------------------------------------------------------
```

Appendix E

R code for estimation of parameters of IG duration models

```
1. ML Estimation (R Code)
--------------------------
library(statmod)
l=1
o=2
a=0.7
b=0.1
n=4000
x=c();si=c();omhat=c();alhat=c();behat=c();lahat=c()
x[1]=0.5;si[1]=0.1
for(j in 1:100){
```

```
abs=rinvgauss(n, 1, l)
alpha=c();beta=c();omega=c();lamda=c()
alpha[1]=a
beta[1]=b
omega[1]=o
lamda[1]=l
for(t in 2:n){
si[t]=o+a*x[t-1]+b*si[t-1]
x[t]=si[t]*abs[t]
}
m=10
for(k in 1:100){
om=c();
om[1]=0.2
for(i in 1:m){
L1=(0.5*sum(1/(om[i]+alpha[k]*x[1:n-1]+beta[k]*si[1:n-1])))
  -((lamda[k]/2)*sum(((((om[i]+alpha[k]*x[1:n-1]+beta[k]
  *si[1:n-1]))^2)-(x[2:n]^2))/((((om[i]+alpha[k]*x[1:n-1]
  +beta[k]*si[1:n-1]))^2)*x[2:n])))
L2=-0.5*sum(1/((om[i]+alpha[k]*x[1:n-1]+beta[k]*si[1:n-1])^2))
   -(lamda[k]*sum(x[2:n]/((om[i]+alpha[k]*x[1:n-1]
   +beta[k]*si[1:n-1])^3)))
om[i+1]=om[i]-(L1/L2)
if(om[i+1]-om[i]<0.001) omg=om[i+1]
```

```
}
omega[k+1]=omg
#alpha
al=c();
al[1]=0.2
for(i in 1:m){
L3=(0.5*sum(x[1:n-1]/(omg+al[i]*x[1:n-1]+beta[k]*si[1:n-1])))
  -(lamda[k]/2)*sum(((((omg+al[i]*x[1:n-1]+beta[k]*si[1:n-1])^2)
  -(x[2:n]^2))/(((omg+al[i]*x[1:n-1]+beta[k]*si[1:n-1])^2)
  *x[2:n]))*x[1:n-1])
L4=-0.5*sum(((x[1:n-1])^2)/((omg+al[i]*x[1:n-1]+beta[k]*si[1:n-1])^2))
   -(lamda[k]*sum((x[2:n]*(x[1:n-1]^2))/((omg+al[i]*x[1:n-1]
   +beta[k]*si[1:n-1])^3)))
al[i+1]=al[i]-(L3/L4)
if(al[i+1]-al[i]<0.001) alp=al[i+1]
}
alpha[k+1]=alp
#beta
be=c();
be[1]=0.2
for(i in 1:m){
L5=(0.5*sum(si[1:n-1]/(omg+alp*x[1:n-1]+be[i]*si[1:n-1])))
   -(lamda[k]/2)*sum(((((omg+alp*x[1:n-1]+be[i]*si[1:n-1])^2)
   -(x[2:n]^2))/(((omg+alp*x[1:n-1]+be[i]*si[1:n-1])^2)
```

```
   *x[2:n]))*si[1:n-1])
L6=-0.5*sum(((si[1:n-1])^2)/((omg+alp*x[1:n-1]+be[i]
   *si[1:n-1])^2))-(lamda[k]*sum((x[2:n]*(si[1:n-1]^2))
   /((omg+alp*x[1:n-1]+be[i]*si[1:n-1])^3)))
be[i+1]=be[i]-(L5/L6)
if(be[i+1]-be[i]<0.001) bet=be[i+1]
}
beta[k+1]=bet
lamda[k+1]=n*(sum(((x[2:n]-(omg+alp*x[1:n-1]+bet*si[1:n-1]))^2)
            /((omg+alp*x[1:n-1]+bet*si[1:n-1])*x[2:n])))^(-1)
if(omega[k+1]-omega[k]<0.001) omeg=omega[k+1]
if(alpha[k+1]-alpha[k]<0.001) alph=alpha[k+1]
if(beta[k+1]-beta[k]<0.001) beta1=beta[k+1]
if(lamda[k+1]-lamda[k]<0.001) lam=lamda[k+1]
}
omhat[j]=omeg
alhat[j]=alph
behat[j]=beta1
lahat[j]=lam
}
omegahat=mean(omhat)
varomeg=sqrt(var(omhat))
alphahat=mean(alhat)
varalph=sqrt(var(alhat))
```

```
betahat=mean(behat)

varbet=sqrt(var(behat))

lamdahat=mean(lahat)

varlam=sqrt(var(lahat))

#omega

omegahat

varomeg

#alpha

alphahat

varalph

#beta

betahat

varbet

#lamda

lamdahat

varlam
---------------------------------------------------------
```

2. EIS ML Estimation

```
clear;
T=1000;
N=100;
omega=input('enter omega:');
beta=input('enter beta:');
sigma=input('enter sigma:');
lamda=input('enter lamda:');
om=-0.1:0.01:0.1;
bt=0.6:0.01:0.8;
sm=0.2:0.01:0.5;
lm=1.4:0.01:1.6;
sma(100)=0;lma(100)=0;oma(100)=0;bta(100)=0;
for rahul=1:100
tic
ui=normrnd(0,sigma,1,T);
sh(T)=0;
sh(1)=ui(1);
for k=2:T
sh(k)=omega+beta*sh(k-1)+ui(k);
end
ig=normrnd(0,1,1,T);
ab(T)=0;
```

```
for ra=1:T
ig1=ig(ra)*ig(ra);
ab1=1+(ig1/(2*lamda))-((1/(2*lamda))*sqrt((4*lamda*ig1)+(ig1*ig1)));
ig2=unifrnd(0,1,1,1);
ig3=1/(1+ab1);
if(ig2<=ig3)
ab(ra)=ab1;
else
ab(ra)=1/ab1;
end
end
x=exp(sh).*ab;
%--------------------sigma------------------------%
r1=length(sm);
L(r1)=0;
for r=1:r1
u=normrnd(0,sm(r),T,N);
si(T,N)=0;
for j=1:N
si(1,j)=u(1,j);
for i=2:T
sig=sm(r)*sm(r);
mu=omega+beta*si(i-1,j);
si(i,j)=mu+sqrt(sig)*u(i,j);
```

```
end
end
for q=1:4
a1(T,2)=0;
ch(T,N)=0;
ch(T,:)=1;
for h=T:-1:1
y=(0.5*log(lamda))+0.5*(si(h,:)')-0.5*log(2*pi)-1.5*log(x(h))
  -((lamda*power((x(h)-exp(si(h,:)')),2))./(2*exp(si(h,:)')*x(h)))
  +(log(ch(h,:)))';
y1=si(h,:)';
p=polyfit(y1,y,2);
a1(h,:)=[p(2),p(1)];
if(h>1)
ch(h-1,:)=(1/sqrt(1-(2*sm(r)*sm(r)*p(1)))).*exp((((sm(r)*sm(r))
         /(2*(1-(2*sm(r)*sm(r)*p(1))))).*power(((omega+beta*si(h-1,:)')
         ./(sm(r)*sm(r)))+p(2),2))-(0.5*power((omega+beta*si(h-1,:)')
         ./sm(r),2)));
end
end
s1(T,N)=0;
for l=1:N
s1(1,l)=u(1,l);
for k=2:T
```

```
sig1=(sm(r)*sm(r))/(1-(2*sm(r)*sm(r)*a1(k,2)));

mu1=sig1*((omega+beta*s1(k-1,l))+a1(k,1));

s1(k,l)=mu1+sqrt(sig1)*u(k,l);

end

end

si=s1;

a(:,:,q)=a1;

end

b=a(:,:,4);

z=s1;

p1(T,N)=0;

q1(T,N)=0;

m1(T,N)=0;

p1(1,:)=sqrt((lamda*exp(z(1,:)))/(2*pi*x(1)^3))
       *exp(-((lamda*power((x(1)-exp(z(1,:))),2))
       /(2*x(1)*exp(z(1,:)))));

q1(1,:)=(1/(sm(r)*sqrt(2*pi))).*(exp(-((1/(2*sm(r)*sm(r)))
       .*power((z(1,:)-omega-beta*u(1,:)),2))));

m1(1,:)=(sqrt(1-(2*sm(r)*sm(r)*b(1,2)))/(sm(r)*sqrt(2*pi)))
        *exp(-(((1-(2*sm(r)*sm(r)*b(1,2)))/(2*sm(r)*sm(r)))
        *power((z(1,:)-(((sm(r)*sm(r))/(1-(2*sm(r)*sm(r)*b(1,2))))
        *(((omega+beta*u(1,:))/(sm(r)*sm(r)))+b(1,1)))),2)));

for w=2:T

p1(w,:)=sqrt((lamda*exp(z(w,:)))/(2*pi*x(w)^3))
```

```
       *exp(-((lamda*power((x(w)-exp(z(w,:))),2))
       /(2*x(w)*exp(z(w,:)))));
q1(w,:)=(1/(sm(r)*sqrt(2*pi))).*(exp(-((1/(2*sm(r)*sm(r)))
        .*power((z(w,:)-omega-beta*z(w-1,:)),2))));
m1(w,:)=(sqrt(1-(2*sm(r)*sm(r)*b(w,2)))/(sm(r)*sqrt(2*pi)))
        *exp(-(((1-(2*sm(r)*sm(r)*b(w,2)))/(2*sm(r)*sm(r)))
        *power((z(w,:)-(((sm(r)*sm(r))/(1-(2*sm(r)*sm(r)*b(w,2))))
        *(((omega+beta*z(w-1,:))/(sm(r)*sm(r)))+b(w,1)))),2)));
end
d=(p1.*q1)./m1;
L(r)=sum(prod(d));
end
Ls=max(L);
e=1;
while(Ls~=L(e))
e=e+1;
end
smh=sm(e);
%----------------------------end estimation of sigma--------------------%
%--------------------lamda-----------------------%
r2=length(lm);
L1(r2)=0;
for rg=1:r2
u1=normrnd(0,smh,T,N);
```

```
si1(T,N)=0;
for j=1:N
si1(1,j)=u1(1,j);
for i=2:T
sig2=smh*smh;
mu2=omega+beta*si1(i-1,j);
si1(i,j)=mu2+sqrt(sig2)*u1(i,j);
end
end
for q=1:4
a2(T,2)=0;
ch1(T,N)=0;
ch1(T,:)=1;
for h=T:-1:1
y2=(0.5*log(lm(rg)))+0.5*(si1(h,:)')-0.5*log(2*pi)-1.5*log(x(h))
  -((lm(rg)*power((x(h)-exp(si1(h,:)')),2))./(2*exp(si1(h,:)')*x(h)))
  +(log(ch1(h,:)))';
y3=si1(h,:)';
pr=polyfit(y3,y2,2);
a2(h,:)=[pr(2),pr(1)];
if(h>1)
ch1(h-1,:)=(1/sqrt(1-(2*smh*smh*pr(1)))).*exp((((smh*smh)
        /(2*(1-(2*smh*smh*pr(1))))).*power(((omega+beta*si1(h-1,:)')
        ./(smh*smh))+pr(2),2))-(0.5
```

```
        *power((omega+beta*si1(h-1,:)')./smh,2)));
end
end
s2(T,N)=0;
for l=1:N
s2(1,l)=u1(1,l);
for k=2:T
sig3=(smh*smh)/(1-(2*smh*smh*a2(k,2)));
mu3=sig3*((omega+beta*s2(k-1,l))+a2(k,1));
s2(k,l)=mu3+sqrt(sig3)*u1(k,l);
end
end
si1=s2;
aa(:,:,q)=a2;
end
b1=aa(:,:,4);
z1=s2;
p11(T,N)=0;
q11(T,N)=0;
m11(T,N)=0;
p11(1,:)=sqrt((lm(rg)*exp(z1(1,:)))/(2*pi*x(1)^3))
         *exp(-((lm(rg)*power((x(1)-exp(z1(1,:))),2))
         /(2*x(1)*exp(z1(1,:)))));
q11(1,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh))
```

```
         .*power((z1(1,:)-omega-beta*u1(1,:)),2))));
m11(1,:)=(sqrt(1-(2*smh*smh*b1(1,2)))/(smh*sqrt(2*pi)))
         *exp(-(((1-(2*smh*smh*b1(1,2)))/(2*smh*smh))*power((z1(1,:)
         -(((smh*smh)/(1-(2*smh*smh*b1(1,2))))*(((omega+beta*u1(1,:))
         /(smh*smh))+b1(1,1)))),2)));
for w=2:T
p11(w,:)=sqrt((lm(rg)*exp(z1(w,:)))/(2*pi*x(w)^3))
         *exp(-((lm(rg)*power((x(w)-exp(z1(w,:))),2))
         /(2*x(w)*exp(z1(w,:)))));
q11(w,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh))
          .*power((z1(w,:)-omega-beta*z1(w-1,:)),2))));
m11(w,:)=(sqrt(1-(2*smh*smh*b1(w,2)))/(smh*sqrt(2*pi)))
        *exp(-(((1-(2*smh*smh*b1(w,2)))/(2*smh*smh))*power((z1(w,:)
        -(((smh*smh)/(1-(2*smh*smh*b1(w,2))))*(((omega+beta*z1(w-1,:))
        /(smh*smh))+b1(w,1)))),2)));
end
d1=(p11.*q11)./m11;
L1(rg)=sum(prod(d1));
end
Lg=max(L1);
e1=1;
while(Lg~=L1(e1))
e1=e1+1;
end
```

```
lmh=lm(e1);
%--------------------------------end-----------------------------------%
%-------------------------omega-------------------------%
r3=length(om);
L2(r3)=0;
for ro=1:r3
u2=normrnd(0,smh,T,N);
si2(T,N)=0;
for j=1:N
si2(1,j)=u2(1,j);
for i=2:T
sig4=smh*smh;
mu4=om(ro)+beta*si2(i-1,j);
si2(i,j)=mu4+sqrt(sig4)*u2(i,j);
end
end
for q=1:4
a3(T,2)=0;
ch2(T,N)=0;
ch2(T,:)=1;
for h=T:-1:1
y4=(0.5*log(lmh))+0.5*(si2(h,:)')-0.5*log(2*pi)-1.5*log(x(h))
   -((lmh*power((x(h)-exp(si2(h,:)')),2))./(2*exp(si2(h,:)')*x(h)))
   +(log(ch2(h,:)))';
```

```
y5=si2(h,:)';
pr1=polyfit(y5,y4,2);
a3(h,:)=[pr1(2),pr1(1)];
if(h>1)
ch2(h-1,:)=(1/sqrt(1-(2*smh*smh*pr1(1)))).*exp((((smh*smh)
          /(2*(1-(2*smh*smh*pr1(1))))).*power(((om(ro)+beta*si2(h-1,:)')
          ./(smh*smh))+pr1(2),2))-(0.5*power((om(ro)
          +beta*si2(h-1,:)')./smh,2)));
end
end
s3(T,N)=0;
for l=1:N
s3(1,l)=u2(1,l);
for k=2:T
sig5=(smh*smh)/(1-(2*smh*smh*a3(k,2)));
mu5=sig5*((om(ro)+beta*s3(k-1,l))+a3(k,1));
s3(k,l)=mu5+sqrt(sig5)*u2(k,l);
end
end
si2=s3;
aa1(:,:,q)=a3;
end
b2=aa1(:,:,4);
z2=s3;
```

```
p22(T,N)=0;
q22(T,N)=0;
m22(T,N)=0;
p22(1,:)=sqrt((lmh*exp(z2(1,:)))/(2*pi*x(1)^3))
         *exp(-((lmh*power((x(1)-exp(z2(1,:))),2))
         /(2*x(1)*exp(z2(1,:)))));
q22(1,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh)).*power((z2(1,:)
        -om(ro)-beta*u2(1,:)),2))));
m22(1,:)=(sqrt(1-(2*smh*smh*b2(1,2)))/(smh*sqrt(2*pi)))
        *exp(-(((1-(2*smh*smh*b2(1,2)))/(2*smh*smh))*power((z2(1,:)
        -(((smh*smh)/(1-(2*smh*smh*b2(1,2))))*(((om(ro)
        +beta*u2(1,:))/(smh*smh))+b2(1,1)))),2)));
for w=2:T
p22(w,:)=sqrt((lmh*exp(z2(w,:)))/(2*pi*x(w)^3))
         *exp(-((lmh*power((x(w)-exp(z2(w,:))),2))
         /(2*x(w)*exp(z2(w,:)))));
q22(w,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh))
         .*power((z2(w,:)-om(ro)-beta*z2(w-1,:)),2))));
m22(w,:)=(sqrt(1-(2*smh*smh*b2(w,2)))/(smh*sqrt(2*pi)))
         *exp(-(((1-(2*smh*smh*b2(w,2)))/(2*smh*smh))
         *power((z2(w,:)-(((smh*smh)/(1-(2*smh*smh*b2(w,2))))
         *(((om(ro)+beta*z2(w-1,:))/(smh*smh))+b2(w,1)))),2)));
end
d2=(p22.*q22)./m22;
```

```
L2(ro)=sum(prod(d2));
end
Lo=max(L2);
e2=1;
while(Lo~=L2(e2))
e2=e2+1;
end
omh=om(e2);
%-----------------------end---------------------------%
%-------------------------beta------------------------%
r4=length(bt);
L3(r4)=0;
for rb=1:r4
u3=normrnd(0,smh,T,N);
si3(T,N)=0;
for j=1:N
si3(1,j)=u3(1,j);
for i=2:T
sig6=smh*smh;
mu6=omh+bt(rb)*si3(i-1,j);
si3(i,j)=mu6+sqrt(sig6)*u3(i,j);
end
end
for q=1:4
```

```
a4(T,2)=0;

ch3(T,N)=0;

ch3(T,:)=1;

for h=T:-1:1

y6=(0.5*log(lmh))+0.5*(si3(h,:)')-0.5*log(2*pi)-1.5*log(x(h))

   -((lmh*power((x(h)-exp(si3(h,:)')),2))./(2*exp(si3(h,:)')*x(h)))

    +(log(ch3(h,:)))';

y7=si3(h,:)';

pr2=polyfit(y7,y6,2);

a4(h,:)=[pr2(2),pr2(1)];

if(h>1)

ch3(h-1,:)=(1/sqrt(1-(2*smh*smh*pr2(1)))).*exp((((smh*smh)/(2

           *(1-(2*smh*smh*pr2(1))))).*power(((omh+bt(rb)*si3(h-1,:)')

           ./(smh*smh))+pr2(2),2))-(0.5*power((omh

           +bt(rb)*si3(h-1,:)')./smh,2)));

end

end

s4(T,N)=0;

for l=1:N

s4(1,l)=u3(1,l);

for k=2:T

sig7=(smh*smh)/(1-(2*smh*smh*a4(k,2)));

mu7=sig7*((omh+bt(rb)*s4(k-1,l))+a4(k,1));

s4(k,l)=mu7+sqrt(sig7)*u3(k,l);
```

```
end
end
si3=s4;
aa2(:,:,q)=a4;
end
b3=aa2(:,:,4);
z3=s4;
p33(T,N)=0;
q33(T,N)=0;
m33(T,N)=0;
p33(1,:)=sqrt((lmh*exp(z3(1,:)))/(2*pi*x(1)^3))*exp(-((lmh*power((x(1)
         -exp(z3(1,:))),2))/(2*x(1)*exp(z3(1,:)))));
q33(1,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh)).*power((z3(1,:)
         -omh-bt(rb)*u3(1,:)),2))));
m33(1,:)=(sqrt(1-(2*smh*smh*b3(1,2)))/(smh*sqrt(2*pi)))
          *exp(-(((1-(2*smh*smh*b3(1,2)))/(2*smh*smh))*power((z3(1,:)
          -(((smh*smh)/(1-(2*smh*smh*b3(1,2))))*(((omh+bt(rb)
          *u3(1,:))/(smh*smh))+b3(1,1)))),2)));
for w=2:T
p33(w,:)=sqrt((lmh*exp(z3(w,:)))/(2*pi*x(w)^3))*exp(-((lmh*power((x(w)
           -exp(z3(w,:))),2))/(2*x(w)*exp(z3(w,:)))));
q33(w,:)=(1/(smh*sqrt(2*pi))).*(exp(-((1/(2*smh*smh)).*power((z3(w,:)
           -omh-bt(rb)*z3(w-1,:)),2))));
m33(w,:)=(sqrt(1-(2*smh*smh*b3(w,2)))/(smh*sqrt(2*pi)))
```

```
            *exp(-(((1-(2*smh*smh*b3(w,2)))/(2*smh*smh))
            *power((z3(w,:)-(((smh*smh)/(1-(2*smh*smh*b3(w,2))))
            *(((omh+bt(rb)*z3(w-1,:))/(smh*smh))+b3(w,1)))),2)));
end
d3=(p33.*q33)./m33;
L3(rb)=sum(prod(d3));
end
Lb=max(L3);
e3=1;
while(Lb~=L3(e3))
e3=e3+1;
end
bh=bt(e3);
sma(rahul)=smh;
lma(rahul)=lmh;
oma(rahul)=omh;
bta(rahul)=bh;
toc
end
est=[mean(oma) mean(bta) mean(sma) mean(lma)]
stvari=[sqrt(var(oma)) sqrt(var(bta)) sqrt(var(sma)) sqrt(var(lma))]
-----------------------------------------------------------------------
```

List of Published/Communicated Papers

1. Balakrishna, N., and Rahul, T. (2014). Inverse Gaussian distribution for modelling conditional durations in Finance. *Communications in Statistics-Simulation and Computation*, 43, 476 – 486.

2. Rahul, T., Balakrishnan, N., and Balakrishna, N. (2017a). Time Series with Birnbaum-Saunders Marginal Distributions. *Applied Stochastic Models in Business and Industry.* (Revised and resubmitted).

3. Balakrishna, N., and Rahul, T. (2017b). *Asymmetric Laplace Stochastic Volatility Model.* (Submitted).

Bibliography

Abraham, B. and Balakrishna, N. (1999). Inverse gaussian autoregressive models. *Journal of time series analysis*, 20(6):605–618.

Abraham, B., Balakrishna, N., and Sivakumar, R. (2006). Gamma stochastic volatility models. *Journal of Forecasting*, 25(3):153–171.

Adke, S. R. and Balakrishna, N. (1992). Markovian chi-square and gamma processes. *Statistics & probability letters*, 15(5):349–356.

Ahmed, S. E., Castro-Kuriss, C., Flores, E., Leiva, V., and Sanhueza, A. (2010). A truncated version of the birnbaum–saunders distribution with an application in financial risk. *Pakistan Journal of Statistics*, 26(1):293–311.

Anderson, D. N. and Arnold, B. C. (1993). Linnik distributions and processes. *Journal of Applied Probability*, 30(02):330–340.

Andersson, J. (2001). On the normal inverse gaussian stochastic volatility model. *Journal of Business & Economic Statistics*, 19(1):44–54.

Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactor volatility models. *Journal of Empirical Finance*, 15(2):332–341.

Asai, M. and McAleer, M. (2011). Alternative asymmetric stochastic volatility models. *Econometric Reviews*, 30(5):548–564.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30.

Balakrishna, N. and Nampoothiri, C. K. (2003). Cauchy autoregressive process and its applications. *Journal of Indian Statistical Association*, 41(2):143–156.

Balakrishna, N. and Rahul, T. (2014). Inverse gaussian distribution for modelling conditional durations in finance. *Communications in Statistics-Simulation and Computation*, 43:476–486.

Balakrishna, N. and Shiji, K. (2014a). Extreme value autoregressive model and its applications. *Journal of Statistical Theory and Practice*, 8(3):460–481.

Balakrishna, N. and Shiji, K. (2014b). Stochastic volatility processes generated by gumbel extreme value autoregressive model. *Journal of Indian Statistical Association*, 52(1):45–64.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 353, pages 401–419. The Royal Society.

Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of statistics*, 24(1):1–13.

Bauwens, L. and Galli, F. (2009). Efficient importance sampling for ml estimation of scd models. *Computational Statistics & Data Analysis*, 53(6):1974–1992.

Bauwens, L., Giot, P., et al. (2000). The logarithmic acd model: An application to the bid-ask quote process of three nyse stocks. *Annals of Economics and Statistics*, (60):117–149.

Bauwens, L., Hafner, C. M., and Laurent, S. (2012). *Handbook of volatility models and their applications*, volume 3. John Wiley & Sons.

Bauwens, L. and 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.

Berg, A., Meyer, R., and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business & Economic Statistics*, 22(1):107–120.

Bhatti, C. R. (2010). The birnbaum–saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation*, 80(10):2062–2078.

Billingsley, P. (1961). *Statistical inference for Markov processes.* University of Chicago Press, Chicago.

Birnbaum, Z. W. and Saunders, S. C. (1969a). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, 6:328–347.

Birnbaum, Z. W. and Saunders, S. C. (1969b). A new family of life distributions. *Journal of Applied probability*, 6:319–327.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, 52(1-2):5–59.

Box, G. E. P. and Jenkins, G. M. (1970). Time series analysis forecasting and control. Technical report, DTIC Document.

Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. San Fransisco: Holden Day.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time series analysis: forecasting and control*. Englewood Cliffs: Prentice Hall.

Box, G. E. P. and 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.

Brockwell, P. J. and Davis, R. A. (1991). *Time series theory and methods*. Springer.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: a survey. *Journal of Economic Surveys*, 18(5):613–649.

Brown, B. G., Katz, R. W., and Murphy, A. H. (1984). Time series models to simulate and forecast wind speed and wind power. *Journal of climate and applied meteorology*, 23(8):1184–1195.

Chhikara, R. S. and Folks, J. L. (1977). The inverse gaussian distribution as a lifetime model. *Technometrics*, 19(4):461–468.

Chib, S., Nardari, F., and Shephard, N. (2002). Markov chain monte carlo methods for stochastic volatility models. *Journal of Econometrics*, 108(2):281–316.

Choy, B., Wan, W. Y., and Chan, C. M. (2008). Bayesian student-t stochastic volatility models via scale mixtures. *Advances in Econometrics, Special issue on Bayesian Econometrics Methods*, 23:595–618.

Danielsson, J. (1994). Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, 64(1):375–400.

Danielsson, J. and Richard, J. F. (1993). Accelerated gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics*, 8(S1):S153–S173.

Delatola, E. I. and Griffin, J. E. (2013). A bayesian semiparametric model for volatility with a leverage effect. *Computational Statistics & Data Analysis*, 60:97–110.

Ding, Z. and Granger, C. W. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of econometrics*, 73(1):185–215.

Drost, F. C. and Werker, B. J. M. (2004). Semiparametric duration models. *Journal of Business & Economic Statistics*, 22(1):40–50.

Durbin, J. and Koopman, S. J. (1997). Monte carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika*, 84(3):669–684.

Engelhardt, M., Bain, L. J., and Wright, F. T. (1981). Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, 23(3):251–256.

Engle, R. and Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66(5):1127–1162.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. (2000). The econometrics of ultra-high-frequency data. *Econometrica*, 68(1):1–22.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1):34–105.

Fernandes, M. and Grammig, J. (2006). A family of autoregressive conditional duration models. *Journal of Econometrics*, 130(1):1–23.

Fox, E., Gavish, B., and Semple, J. H. (2008). A general approximation to the distribution of count data with applications to inventory modeling.

Gallant, A. R., Hsieh, D., and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, 81(1):159–192.

Gaver, D. P. and Lewis, P. A. W. (1980). First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability*, 12(03):727–745.

Ghysels, E., Harvey, A. C., and Renault, E. (1996). Stochastic volatility. *Handbook of statistics*, 14:119–191.

Gibson, J. D. (1983). Data compression of a first order intermittently excited ar process. In *Naval Research Sponsored Workshop on Statistical Image Processing and Graphics*, pages 115–126.

Grammig, J. and Maurer, K. O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*, 3(1):16–38.

Gupta, R. C. and Viles, W. (2011). Roller-coaster failure rates and mean residual life functions with application to the extended generalized inverse gaussian model. *Probability in the Engineering and Informational Sciences*, 25(01):103–118.

Hamilton, J. D. and Jorda, O. (2002). A model of the federal funds rate target. *Journal of Political Economy*, 110(5):1135–1167.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

Harvey, A., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Hautsch, N. (2001). The generalized f acd model. *Unpublished Paper*.

Hautsch, N. (2012). *Econometrics of financial high-frequency data*. Berlin: Springer.

Jacquier, E., Polson, N. G., and Rossi, P. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 12(4):371–89.

Jacquier, E., Polson, N. G., and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

Jasiak, J. (1998). Persistence in intertrade durations. *Finance*, 19:166–195.

Jayakumar, K. and Kuttykrishnan, A. P. (2007). A time-series model using asymmetric laplace distribution. *Statistics & Probability Letters*, 77(16):1636–1640.

Jin, X. and Kawczak, J. (2003). Birnbaum-saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, 4:103–124.

Jørgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*. Springer.

Jose, K. K. and Thomas, M. M. (2011). Generalized laplacian distributions and autoregressive processes. *Communications in Statistics-Theory and Methods*, 40(23):4263–4277.

Jose, K. K., Tomy, L., and Sreekumar, J. (2008). Autoregressive processes with normal-laplace marginals. *Statistics & Probability Letters*, 78(15):2456–2462.

Karanasos, M. (2001). The statistical properties of long-memory and exponential acd models. *Unpublished Paper, University of York*.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous multivariate distributions Vol 1.* Second Edition. John Wiley & Sons, New York.

Kotz, S., Kozubowski, T., and Podgorski, K. (2012). *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance.* Springer Science & Business Media.

Kozubowski, T. J. and Podgorski, K. (2000). A multivariate and asymmetric generalization of laplace distribution. *Computational Statistics*, 15(4):531–540.

Kundu, D., Balakrishnan, N., and Jamalizadeh, A. (2010). Bivariate birnbaum–saunders distribution and associated inference. *Journal of multivariate Analysis*, 101(1):113–125.

Kundu, D., Balakrishnan, N., and Jamalizadeh, A. (2013). Generalized multivariate birnbaum–saunders distributions and related inferential issues. *Journal of Multivariate Analysis*, 116:230–244.

Lancaster, T. (1972). A stochastic model for the duration of a strike. *Journal of the Royal Statistical Society. Series A (General)*, 135(2):257–271.

Lancaster, T. (1992). *The econometric analysis of transition data.* Cambridge university press.

Lawrance, A. J. (1991). Directionality and reversibility in time series. *International Statistical Review*, 59(1):67–79.

Lawrance, A. J. and Kottegoda, N. T. (1977). Stochastic modelling of riverflow time series. *Journal of the Royal Statistical Society. Series A (General)*, pages 1–47.

Lawrance, A. J. and Lewis, P. A. W. (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(2):165–202.

Leiva, V. (2016). *The Birnbaum-Saunders Distribution*. Elsevier, Academic Press.

Leiva, V., Marchant, C., Saulo, H., Aslam, M., and Rojas, F. (2014a). Capability indices for birnbaum–saunders processes applied to electronic and food industries. *Journal of Applied Statistics*, 41(9):1881–1902.

Leiva, V., Rojas, E., Galea, M., and Sanhueza, A. (2014b). Diagnostics in birnbaum–saunders accelerated life models with an application to fatigue data. *Applied Stochastic Models in Business and Industry*, 30(2):115–131.

Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2016). A methodology for stochastic inventory models based on a zero-adjusted birnbaum-saunders distribution. *Applied Stochastic Models in Business and Industry*, 32(1):74–89.

Leiva, V., Saulo, H., Leão, J., and Marchant, C. (2014c). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics & Data Analysis*, 79:175–191.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics*, 15(2):137–160.

Liesenfeld, R. and Richard, J. F. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of empirical finance*, 10(4):505–531.

Liesenfeld, R. and Richard, J. F. (2006). Classical and bayesian analysis of univariate and multivariate stochastic volatility models. *Econometric Reviews*, 25(2-3):335–360.

Lio, Y. L. and Park, C. (2008). A bootstrap control chart for birnbaum–saunders percentiles. *Quality and Reliability Engineering International*, 24(5):585–600.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. *Unpublished Paper*.

Mandebrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial markets*, 8(4):377–399.

Marchant, C., Bertin, K., Leiva, V., and Saulo, H. (2013). Generalized birnbaum–saunders kernel density estimators and an analysis of financial data. *Computational Statistics & Data Analysis*, 63:1–15.

Meitz, M. and Teräsvirta, T. (2006). Evaluating models of autoregressive conditional duration. *Journal of Business & Economic Statistics*, 24(1):104–124.

Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1):239–265.

Nakajima, J. and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics & Data Analysis*, 53(6):2335–2353.

Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using gh skew students t-distribution. *Computational Statistics & Data Analysis*, 56(11):3690–3704.

Nelson, H. L. and Granger, C. W. J. (1979). Experience with using the box-cox transformation when forecasting economic time series. *Journal of Econometrics*, 10(1):57–69.

Ng, H. K. T., Kundu, D., and Balakrishnan, N. (2003). Modified moment estimation for the two-parameter birnbaum–saunders distribution. *Computational Statistics & Data Analysis*, 43(3):283–298.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449.

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.

Paula, G. A., Leiva, V., Barros, M., and Liu, S. (2012). Robust statistical modeling using the birnbaum-saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, 28(1):16–34.

Popović, B. V. (2010). Ar(1)time series with approximated beta marginal. *Publ. Inst. Math.(Belgrade)(NS)*, 80(102):89–98.

Popović, B. V., Pogány, T. K., and Nadarajah, S. (2010). On mixed ar (1) time series model with approximated beta marginal. *Statistics & probability letters*, 80(19):1551–1558.

Richard, J. F. and Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141(2):1385–1411.

Rieck, J. R. (1999). A moment-generating function with application to the birnbaum-saunders distribution. *Communications in Statistics-Theory and Methods*, 28(9):2213–2222.

Robinson, P. M. and Henry, M. (1999). Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. *Econometric theory*, 15(03):299–336.

Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of econometrics*, 63(1):289–306.

Sandmann, G. and Koopman, S. J. (1998). Estimation of stochastic volatility models via monte carlo maximum likelihood. *Journal of Econometrics*, 87(2):271–301.

Shephard, N. (1996). Statistical aspects of arch and stochastic volatility. *Monographs on Statistics and Applied Probability*, 65:1–68.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika*, 84(3):653–667.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.

Sim, C. H. (1990). First-order autoregressive models for gamma and exponential processes. *Journal of Applied Probability*, 27(02):325–332.

Taylor, S. J. (1986). *Modelling financial time series.* John Wiley & Sons, New York.

Taylor, S. J. (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical finance*, 4(2):183–204.

Tong, H. (1995). A personal overview of nonlinear time series analysis from a chaos perspective. *Scandinavian journal of Statistics*, 22:399–421.

Tsay, R. S. (2005). *Analysis of financial time series*, volume 543. John Wiley & Sons.

Tsay, R. S. (2009). Autoregressive conditional duration models. in terence c. mills and kerry patterson (eds), palgrave handbook of econometrics. 2.

Wang, J. J. J. (2012). On asymmetric generalised t stochastic volatility models. *Mathematics and Computers in Simulation*, 82(11):2079–2095.

Wang, J. J. J., Chan, J. S. K., and Choy, S. T. B. (2011). Stochastic volatility models with leverage and heavy-tailed distributions: A bayesian approach using scale mixtures. *Computational Statistics & Data Analysis*, 55(1):852–862.

Wang, J. J. J., Choy, S. T. B., and Chan, J. S. K. (2013). Modelling stochastic volatility using generalized t distribution. *Journal of Statistical Computation and Simulation*, 83(2):340–354.

Wanke, P. and Leiva, V. (2015). Exploring the potential use of the birnbaum-saunders distribution in inventory management. *Mathematical Problems in Engineering*, 1:1–9.

Zhang, M. Y., Russell, J. R., and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, 104(1):179–207.