A Study of Non-Gaussian Modeling for Financial Economics

Yuichi Nagahara

Doctor of Philosophy

Department of Statistical Science
School of Mathematical and Physical Science
The Graduate University for Advanced Studies

1996

Acknowledgement

First of all, I wish to express my deepest gratitude to my supervisor Prof. Genshiro Kitagawa of the Institute of Statistical Mathematics for his helpful suggestions and comments. He encouraged me heartly in my studies and gave me the knowledge of statistics. With great patience, he spent a lot of time discussing my studies with me, reading my research papers, and giving valuable ideas. I learned a lot from his strict attitude.

And I would like to thank my adviser Prof. Tohru Ozaki of the Institute of Statistical Mathematics for giving a lot of valuable advice and lecture. He gave me the knowledge of the stochastic differential equation and many valuable comments.

I wish to thank the Director General of the Institute of Statistical Mathematics, Prof. Ryoichi Shimizu for making lecture about the probability theory and giving me many advices.

I also would like to thank Prof. Nakahiro Yoshida for giving me many helpful advice and seminar, especially on stochastic process.

And I would like to thank Prof. Makio Ishiguro and Prof. Yoshiyasu Tamura of the Institute of Statistical Mathematics for giving me many helpful comments about my studies.

I should be grateful to Prof. Yasuhiro Yonezawa of the University of Tsukuba for giving me many helpful comments about my studies. He gave me the knowledge of the financial ecomomics since my stay in the master course of the University of Tsukuba.

I would like to thank Prof. Takeaki Kariya of the Hitotsubashi University for giving me many helpful comments about my studies and how to write a paper in English.

And I would like to thank Prof. Sadao Naniwa of the Ritsumeikan University for giving me many helpful comments about my studies and the knowledge of the time series analysis of economic data.

My thank are due to the staffs of the Institute of Statistical Mathematics, my colleague students in the Graduate University for Advanced Studies for supporting me during the course of my research.

My thank are also due to The Nikko Securities Co., Ltd. for giving me the chance to study at graduate universities.

Finally, I would like to thank my wife, Rie, for her heartly support, my son, Hiroyuki, for his birth, and my deceased father, Syouji, deceased mother, Kuniko, for their protection.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the financial economics, the management of the market risk and credit risk becomes the most important subject. The market risk is defined to be the loss by the market fluctuation, of financial institutions and individual investors who have the financial assets. The credit risk is defined to be the loss by the default of the counterparties, of those who have some contracts. In recent years, the international financial markets are expanded and deeply linked. There is a possibility that the financial crisis of a certain country causes the financial instability of other countries. Therefore, the financial regulators propose the methods of the risk management and intend to control risks by its methods. The main concept of this method is called VaR (Value at Risk). The VaR is defined to be the maximum loss of an asset return. It is measured by the following process. Firstly, the 99 percents one-sided confidence interval is decided by the financial regulators. Secondly, on the one-sided confidence interval, the VaR is calculated by using actual distribution of returns. For example, if the distribution of asset returns is normal and the expectation is zero, the result of multiplying 2.326 by the standard deviation yields the VaR. In the proposed method, it is asssumed that the distribution is normal.

In this way, most of the financial asset returns traditionally tend to be analyzed on the assumption of the linearity and Gaussianity. However, it can be seen that the actual distribution is not normal. Recently, various methods of analyzing nonlinear and non-Gaussian system are developed. The fact that the financial asset returns have the nonlinearity and non-Gaussianity has been made clear. Furthermore, the number of the bankruptcy is very important information for the credit risk management. The number of the bankruptcy of large companies is very small but its influence is very wide. Therefore, the time series analysis of the bankruptcy is also very important.

The purpose of this study is to develop financial models which represent the nonlinearity and non-Gaussianity of the financial economic data. This study may be directly applied to the risk management.

In this study, we focus on the returns of stocks and the number of the bankruptcy. To begin with, we describe the characteristics of stock returns. Firstly, the asymmetry of price movements in stock markets is observed. The upward trend tends to continue. On the contrary, the downward movement is instantaneous and its amount of movement is occasionally very large. Therefore, the distribution of returns becomes asymmetric. Secondly, the distribution of returns is observed to be heavy-tailed than the normal distribution. Thirdly, the volatility is time-varying. In this study, we adopt the type VII and IV family of Pearson system in order to describe the asymmetry and heavy-tail. For the Pearson system, there is the relationship between the stationary distribution and a stochastic differentical equation.

In Chapter 2, recursive formula for evaluating the normalizing constant for the type IV family of Pearson system is derived. The non-central and heavy-tailed distribution is introduced by adding non-central parameter to numerator of equation of the type VII Pearson System. This distribution belongs to the type IV of Pearson System. Analytic solutions of normalizing constant are derived in order to evaluate numerical integration of normalizing this distribution.

Cauchy distribution and t distribution (with low degree of freedom) have heavy-tails compared with normal distribution. They belong to the type VII family of Pearson System. In this thesis, the non-central and heavy-tailed distribution is introduced by adding non-central parameter, δ, to the numerator of the Pearson System equation (1.1)

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta)}{(\frac{x-\mu}{\tau})^2 + 1}. \tag{1.1}$$

The probability density function is given by

$$p(x|\mu, \tau, \delta, b) = \frac{C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2 + \tau^2\}^b}, \tag{1.2}$$

for $b > 1/2$, $\tau > 0$, $\infty > \delta > -\infty$.

This type of distribution is known as the type IV family of Pearson System (Johnson and Kotz (1970), Pearson (1914), Pearson and Hartrey (1954)). In the past research, this distribution is not introduced in this way, but used as a approximation to non-central t distribution by using moment ratio (Merrington and Pearson (1958), Shenton and Carpenter (1964), Pearson (1963)). And the type IV family has not been used in actual data analysis because of its difficulty in computation. Therfore, in order to use this distribution in actual statistical modeling, it is necessary to develop a practical method for evaluating normalizing constant. In this thesis, a recursive formula for evaluating normalizing constant is given for a family with discrete parameter series, $b = n$ and $b = n + 1/2$, $n = 1, 2, 3....$ Analytic form of normalizing constant is derived by using recursive formula. The recursive formula is also useful for the numerical evaluation for general b. Also, the first four central moments, which specify the distribution of Pearson System, are explicitly calculated.

Price movements in a stock market are often associated with industrywise movements of stock prices. In other words, it is often said that stock prices move industrywise, which is usually referred to as "industry-effect". However, Nagahara (1992) observed that there is an asymmetry in the contributions of industry prices to the movements of a market price. In the upward movements of the market price, it is often the case that a certain industry prices move first prior to the movement of other industry prices. On the other hand, in the downward movements of the market price, most industry prices move together. This asymmetry may be measured by the skewness of the cross-sectional distribution of industry returns, and it will be made a good use of a description or prediction of the market average returns. In fact, in Chapter 3 of this thesis it shows that the skewness of the distribution of industry returns is very effective in describing and predicting the variations of the market average returns. In the stock market, by the empirical studies many researches have observed an industry-effect in the movement of prices or returns, meaning that industry returns (returns based on industry indices) move rather independently. However, Nagahara (1992) observed that there also exists co-movement (correlations) of industry returns and asymmetry of the co-movements. In fact, when the market price is in a downward trend, almost all industry prices tend to move downwards together. On the other hand, when the market is in an upward trend, the following two typical cases are observed: In one case only some specific industry returns move upwards while others do not move so much, and the other case all the industry returns move together and form an upward trend. This implies that the correlations among industry returns and the skewness of the cross-sectional distributions of industry returns will change with time t. In other words, there may be some information in these changes of the co-movements and asymmetry of cross-sectional industry returns over time for predicting future market average returns. These observations motivate our study.

In the literature, much attention has not been paid on this effect, though there are many time series studies on the distributional aspects of market and individual returns. For examples, Fama (1965) and Mandelbrot (1966) suggested the use of such a heavy-tailed distribution as the stable Paretian distribution as a model for daily returns. Kariya, et al. (1995) extensively studied on this feature as well as some nonlinear features in the Japanese market (see also Kariya (1993)), which showed that the distribution of market returns is not only leptokurtic but also skewed. In our time series analysis below, to take into account of this feature, we use the Pearson type IV distribution which is a non-central version of the type VII distribution and forms a broad class of distributions including t-distribution and Cauchy distribution (Pearson, K (1914), Pearson, E.S and Hartrey (1954), Pearson, E.S (1963), Johnson and Kotz (1970)) and skewed version of these distribution.

In this thesis, we mainly study on a predictive power of past values of the cross-sectional skewness variable S for future values of the averaged return R. In predicting the averaged return, we compare with an AR-GARCH (Autoregressive model with Generalized Autoregressive Conditional Heteroskedasticity errors, Bolleslev(1986)) type time series model by using the AIC criterion for model selection, and examine that the skewness of the cross-sectional distribution of industries has more predictive power for future averaged returns R_t's than only the past averaged returns. We also propose some state-dependent models to find that the conditional variances depend on some shape parameters such as skewness, which may be functions of economic fundamentals, in addition to the past conditional variances and past errors.

In Chapter 4, daily returns of stock prices are observed to have heavy-tailed and non-central distribution. In this chapter, we adopt the type VII and IV family of Pearson System to express the daily returns of stock prices. Furthermore, we also consider a stochastic differential equation whose stationary distribution is the type VII or IV of Pearson system (Wong, 1963). This reveals the relationship between the stationary distribution of stock returns and related stochastic differential equation. Wong showed the transition probability density functions of stochastic differential equation for the Type VII Pearson distribution. It can be written more explicitly for t distribution. In this chapter, we derive the transition probability density function for a more general family of distributions, and present a method for estimating the parameters of its stochastic differential equation. Furthermore, in order to estimate the parameters of stochastic differential equation which corresponds to the Pearson IV type (asymmetric distribution and continuous shape parameter), we consider local linearization method (Ozaki (1985a, 1985b, 1989, 1992a, 1992b, 1993), Biscay, Jimenez, Riera and Valdes (1994), Shoji and Ozaki (1994) and Nagahara (1995b)).

In Chapter 5, Non-Gaussian stochastic volatility model is proposed. The model assumes the time series is distributed as the Pearson type VII distribution. The scale parameter of the distribution is stochastic and is described by AR model with a constant term. For estimating the parameters of stochastic volatility model, we apply the non-Gaussian filter. The model can be further generalized to the case when the shape parameter of the Pearson type VII distribution is time-varying. Usefulness of the model is demonstrated by the analysis of stock returns data. In the finance area, the management of the market risk is very important subject. In this area, ARCH model (Engle (1982)), GARCH model (Bollerslev (1986)) has been used because of its easy implementation. The method of estimating a stochastic volatility model was recently developed. On the other hand, the comparison of GARCH and stochastic volatility model are researched only empirically (Heyman, et.al. 1994). According to their result, for stock returns, the stochastic volatility model is better than GARCH and EGARCH (Exponential GARCH, Nelson (1991)). In this thesis, in order to construct the stochastic volatility model for the risk management, we consider the simple form of a stochastic volatility model, namely the observation model has the heavy tailed distribution and the expectation of observation is zero. Recently, stochastic volatility models are extensively researched and applied to actual data analysis. We apply the non-Gaussian filter to a stochastic volatility model (Kitagawa 1987, 1991). This paper focus on Pearson VII type, including t-distribution and Cauchy distribution. The research by Nagahara (1995b) and other research for the daily returns of stock index prices conclude that the distribution of stock daily returns are heavy-tailed like the Pearson type VII or the Paretian distribution compared with Gaussian distribution. Furthermore, according to Nagahara (1995b), the shape parameter of the Pearson type VII tends to be time-varying. Therefore, we develop general state space model of a stochastic volatility model with the time-varying shape parameter. We estimate these parameters by using the method by Kitagawa (1987, 1991).

Finally, we consider the characteristics of the number bankruptcy of the large companies. These data are discrete and very small, occasionally becoming zero. And there is seasonal component in its intensity. In Chapter 6, we develop the non-Gaussian model by using general state space model, in which the observation model is given by the Poisson distribution. We use Monte Carlo filter to estimate the parameters.

In Chapter 7, the conclusions are given.

Chapter 2

Recursive Formula for Evaluating the Normalizing Constant and Cumulative Distribution Function of Type IV Family of Pearson System of Distributions

2.1 Introduction

Financial data, especially daily or weekly rate of returns of stock prices, are observed as heavy-tailed and non-central distribution. In order to fit this distribution, we need non-central heavy-tailed distribution whose parameters vary continuously according to exogeneous variables like economic fundamentals.

Cauchy distribution and t distribution (with low degree of freedom) have heavy-tail compared with normal distribution. They belong to type VII family of Pearson System. In this chapter, heavy-tailed non-central distribution is introduced by adding non-central parameter, δ, to numerator of the Pearson System equation (2.1)

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta)}{(\frac{x-\mu}{\tau})^2 + 1}. \tag{2.1}$$

The probability density function is given by

$$p(x|\mu, \tau, \delta, b) = \frac{C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2 + \tau^2\}^b}, \tag{2.2}$$

for $b > 1/2$, $\tau > 0$, $\infty > \delta > -\infty$.

This type of distribution is known as type IV family of Pearson System (Johnson and Kotz (1970), Pearson (1914), Pearson and Hartrey (1954)). In the past research, this distribution is not introduced in this way, but used as a approximation to non-central t distribution by using moment ratio (Merrington and Pearson (1958), Shenton and Carpenter (1964), Pearson (1963)). And the type IV family is not used for empirical purpose because of its difficulty in computation. Therfore, in order to use this distribution in actual statistical modeling, it is necessary to develop a practical method for evaluating normalizing constant. In this chapter, a recursive formula for evaluating normalizing constant is given. Analytic form of normalizing constant, for $b = n$ and $b = n+1/2, \quad n = 1, 2, 3...$, are derived by using recursive formula. The recursive formula is also useful for the numerical evaluation for general b. Also, the first four moment, which specify the distribution of Pearson System, are calculated. Furthermore, recursive formula is shown to calculate cumulative distribution function.

In section 2.2, normalizing constant is derived by using recursive formula, and the first four moments are given. In section 2.3, cumulative distribution functions for $b = 1, 3/2$ and recursive formula are shown.

2.2 Evaluation of the Normalizing Constant

Let p be probability density function. We assume that the distribution is a solution to the following differential equation

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta)}{(\frac{x-\mu}{\tau})^2 + 1}. \tag{2.3}$$

As it can be seen later, μ and τ specify the center and the dispersion of the distribution. δ is introduced here to specify the non-centrality of the distribution. From (2.3) we have

$$\begin{aligned}
-\int d\log p &= \int \frac{2b(x-\mu-\delta\tau)}{(x-\mu)^2+\tau^2} dx \\
-\log p &= 2b\left[\frac{1}{2}\log\{(x-\mu)^2+\tau^2\} - \delta\arctan(\frac{x-\mu}{\tau}) + C'\right] \\
&= b\log\{(x-\mu)^2+\tau^2\} - 2b\delta\arctan(\frac{x-\mu}{\tau}) + 2bC'.
\end{aligned}$$

Therefore, the density function $p(x)$ is given by

$$p(x) = C\exp\left[-b\log\{(x-\mu)^2+\tau^2\} + 2b\delta\arctan(\frac{x-\mu}{\tau})\right]$$

$$= \frac{C}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}.$$

Obviously, the normalizing constant C should satisfy the following relation

$$\int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx = 1. \tag{2.4}$$

Lemma 2.2.1 *For $b = 1$, C is given by*

$$C = \frac{\delta\tau}{\sinh(\delta\pi)}. \tag{2.5}$$

Proof. By evaluating the left hand side of (2.4), we have

$$\int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}} \exp\{2\delta \arctan(\frac{x-\mu}{\tau})\} dx$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{C}{2\delta} \frac{2\delta}{\{(x-\mu)^2+\tau^2\}} \exp\{2\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \int_{-\infty}^{\infty} \frac{C}{2\delta\tau} [\exp\{2\delta \arctan(\frac{x-\mu}{\tau})\}]' dx \\
&= \frac{C}{2\delta\tau} \left[\exp\{2\delta \arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&= \frac{C}{2\delta\tau} \{\exp(\delta\pi) - \exp(-\delta\pi)\} = 1.
\end{aligned}$$

Therefore, the constant term is given by

$$C = \frac{\delta\tau}{\sinh(\delta\pi)}.$$

∎

Corollary 2.2.1 *For $b = 1$, the density function is given by*

$$p(x) = \frac{\delta\tau}{\sinh(\delta\pi)} \frac{\exp\{2\delta \arctan(\frac{x-\mu}{\tau})\}}{(x-\mu)^2+\tau^2}. \tag{2.6}$$

Lemma 2.2.2 *For $b = \frac{3}{2}$, C is given by*

$$C = \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}. \tag{2.7}$$

Proof. We consider the transformation $\frac{x-\mu}{\tau} = \tan\theta$, $\frac{dx}{d\theta} = \tau\frac{1}{\cos^2\theta} = \tau(1+\tan^2\theta)$.

Then we have

$$\int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}^{\frac{3}{2}}} \exp\left\{2 \times \frac{3}{2} \times \delta \arctan(\frac{x-\mu}{\tau})\right\} dx \tag{2.8}$$

$$\begin{aligned}
&= C\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(3\delta\theta)}{\{\tau^2(1+\tan^2\theta)\}^{\frac{3}{2}}}\frac{dx}{d\theta}d\theta \\
&= C\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(3\delta\theta)}{\tau^3\frac{1}{\cos^3\theta}} \times \tau \times \frac{1}{\cos^2\theta}d\theta \\
&= C\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\cos\theta}{\tau^2}\exp(3\delta\theta)d\theta \\
&= \frac{C}{\tau^2}\left[\frac{\exp(3\delta\theta)}{(3\delta)^2+1^2}(3\delta\cos\theta+\sin\theta)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \frac{C}{\tau^2}\left\{\frac{\exp(\frac{3\delta\pi}{2})+\exp(-\frac{3\delta\pi}{2})}{(3\delta)^2+1^2}\right\} \\
&= C\frac{2\cosh(\frac{3\delta\pi}{2})}{(3\delta\tau)^2+\tau^2} = 1.
\end{aligned}$$

Therefore, the constant term C is given by

$$C = \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}.$$

∎

Corollary 2.2.2 *For $b = \frac{3}{2}$, the density function is given by*

$$p(x) = \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\frac{\exp\{3\delta\arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2+\tau^2\}^{\frac{3}{2}}}. \tag{2.9}$$

Lemma 2.2.3 *Let I_b be defined by*

$$I_b = \int_{-\infty}^{x} \frac{d\xi}{\{(\xi-\mu)^2+\tau^2\}^b},$$

then we have the following recursion formula,

$$I_b = \frac{1}{\tau^2}\left[\frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}} + \frac{2b-3}{2b-2}I_{b-1}\right]. \tag{2.10}$$

Proof. The proof of the lemma is omitted since it is trivial.

Lemma 2.2.4 *Let J(b,c) be defined by*

$$\int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx$$

then we have the following recursive formula

$$\frac{(2b-2)}{(2b-3)} \left\{ 1 + \frac{(c\delta)^2}{(b-1)^2} \right\} \tau^2 J(b,c) = J(b-1,c). \tag{2.11}$$

Proof. According to Lemma 2.2.3, we obtain

$$\begin{aligned}
J(b,c) &= \int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\tau^2} \left[\left\{ \frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}} \right\}' + \frac{2b-3}{2b-2} \frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}} \right] \\
&\quad \times \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \frac{1}{\tau^2} \left[\frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} \right]_{-\infty}^{\infty} \\
&\quad - \frac{1}{\tau^2} \int_{-\infty}^{\infty} \frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}} \frac{2c\delta\tau}{(x-\mu)^2+\tau^2} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&\quad + \frac{1}{\tau^2} \frac{2b-3}{2b-2} \int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= -\frac{1}{\tau^2} \frac{c\delta\tau}{b-1} \int_{-\infty}^{\infty} \frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&\quad + \frac{1}{\tau^2} \frac{2b-3}{2b-2} J(b-1,c). \qquad (2.12)
\end{aligned}$$

Next, for $b \geq \frac{3}{2}$,

$$\begin{aligned}
J(b,c) &= \int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \int_{-\infty}^{\infty} \frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau} \times \frac{2c\delta\tau}{(x-\mu)^2+\tau^2} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \int_{-\infty}^{\infty} \frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau} [\exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\}]' dx \\
&= \left[\frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} \right]_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} \frac{(-b+1)}{2c\delta\tau} \frac{2(x-\mu)}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \frac{b-1}{c\delta\tau} \int_{-\infty}^{\infty} \frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx
\end{aligned}$$

Summarizing the above results, we have

$$\begin{aligned} J(b,c) &= \frac{1}{\tau^2}\frac{2b-3}{2b-2}J(b-1,c) \\ &\quad -\frac{1}{\tau^2}\frac{c\delta\tau}{b-1}\int_{-\infty}^{\infty}\frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx \end{aligned}$$

Secondly, we have

$$J(b,c) = \frac{b-1}{c\delta\tau}\int_{-\infty}^{\infty}\frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx.$$

Substituting the second formula into the second term of the first formula, we get the following formula

$$J(b,c) = \frac{1}{\tau^2}\frac{2b-3}{2b-2}J(b-1,c) - \frac{1}{\tau^2}\frac{c\delta\tau}{b-1}\frac{c\delta\tau}{b-1}J(b,c). \tag{2.13}$$

Finally from (2.12) and (2.13), we have

$$\left\{1+\frac{(c\delta)^2}{(b-1)^2}\right\}J(b,c) = \frac{1}{\tau^2}\frac{2b-3}{2b-2}J(b-1,c).$$

This completes the proof of the recursion formula. ∎

Theorem 2.2.1 (Normalizing Constant) *Normalizing Constant of the distribution (2.2) is given as follows,*

1)For b=n , $n = 1, 2, ...$

$$p(x|\mu,\tau,\delta,b) = \frac{b\delta\tau^{2n-1}}{\sinh(b\delta\pi)}\frac{\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2+\tau^2\}^b}\prod_{k=1}^{n-1}\left(1+\frac{b^2\delta^2}{k^2}\right)\frac{2k}{2k-1}. \tag{2.14}$$

2)For b=n+$\frac{1}{2}$, $n = 1, 2, ...$

$$p(x|\mu,\tau,\delta,b) = \frac{(1+(2b\delta)^2)\tau^{2n}}{2\cosh(b\delta\pi)}\frac{\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2+\tau^2\}^b}\prod_{k=1}^{n-1}\left\{1+\frac{b^2\delta^2}{(k+\frac{1}{2})^2}\right\}\frac{2k+1}{2k}. \tag{2.15}$$

Proof. 1)For $b = 2, 3, ...(b = n,\ n = 2, 3, ...)$

Firstly, for $b = 1$

$$J(1,c) = \int_{-\infty}^{\infty}\frac{1}{\{(x-\mu)^2+\tau^2\}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{1}{2c\delta\tau} \times \frac{2c\delta\tau}{(x-\mu)^2+\tau^2} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{2c\delta\tau} \left[\exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\}\right]' dx \\
&= \frac{1}{2c\delta\tau} \left[\exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&= \frac{1}{2c\delta\tau} \{\exp(c\delta\pi) - \exp(-c\delta\pi)\}.
\end{aligned}$$

Therefore, we obtain

$$1 = \frac{c\delta\tau}{\sinh(c\delta\pi)} J(1,c). \tag{2.16}$$

Secondly, the recursion formula (2.11) is applied to (2.16).

$$\begin{aligned}
1 &= \frac{c\delta\tau}{\sinh(c\delta\pi)} J(1,c) \\
&= \frac{c\delta\tau}{\sinh(c\delta\pi)} \times \frac{(2\times 2-2)}{(2\times 2-3)} \left\{1+\frac{(c\delta)^2}{(2-1)^2}\right\} \tau^2 J(2,c) \\
&= \frac{c\delta\tau}{\sinh(c\delta\pi)} \frac{(2\times 2-2)}{(2\times 2-3)} \left\{1+\frac{(c\delta)^2}{(2-1)^2}\right\} \tau^2 \frac{(2\times 3-2)}{(2\times 3-3)} \left\{1+\frac{(c\delta)^2}{(3-1)^2}\right\} \tau^2 J(3,c) \\
&\vdots \\
&= \frac{c\delta\tau}{\sinh(c\delta\pi)} \times 2\left\{1+\frac{(c\delta)^2}{1^2}\right\}\tau^2 \times \frac{4}{3}\left\{1+\frac{(c\delta)^2}{2^2}\right\}\tau^2 \times \cdots \times \frac{(2c-2)}{(2c-3)}\left\{1+\frac{(c\delta)^2}{(c-1)^2}\right\}\tau^2 J(c,c)
\end{aligned}$$

2)For b=5/2,7/2,... (b=n+$\frac{1}{2}$, $n = 2, 3, ...$)

Firstly, $J(b,c)$ is evaluated for $b = \frac{3}{2}$. By the transformation $\frac{x-\mu}{\tau} = \tan\theta$ $\frac{dx}{d\theta} = \tau\frac{1}{\cos^2\theta} = \tau(1+\tan^2\theta)$, we have

$$J(\frac{3}{2},c) = \int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2+\tau^2\}^{\frac{3}{2}}} \exp\{2c\delta \arctan(\frac{x-\mu}{\tau})\} dx \tag{2.17}$$

$$\begin{aligned}
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(2c\delta\theta)}{\{\tau^2(1+\tan^2\theta)\}^{\frac{3}{2}}} \frac{dx}{d\theta} d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(2c\delta\theta)}{\tau^3\frac{1}{\cos^3\theta}} \times \tau \times \frac{1}{\cos^2\theta} d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp(2c\delta\theta) \frac{\cos\theta}{\tau^2} d\theta \\
&= \frac{1}{\tau^2} \left[\frac{\exp(2c\delta\theta)}{(2c\delta)^2+1^2}(3\delta\cos\theta+\sin\theta)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \frac{1}{\tau^2} \left[\frac{\exp(c\delta\pi)+\exp(-c\delta\pi)}{(2c\delta)^2+1^2}\right]
\end{aligned}$$

$$= \frac{2\cosh(c\delta\pi)}{(2c\delta\tau)^2 + \tau^2}.$$

Therefore, we have

$$1 = \frac{(2c\delta\tau)^2 + \tau^2}{2\cosh(c\delta\pi)} J(\frac{3}{2}, c). \tag{2.18}$$

Secondly, by applying the recursion formula (2.11) to (2.18), we have

$$\begin{aligned}
1 &= \frac{(2c\delta\tau)^2 + \tau^2}{2\cosh(c\delta\pi)} \frac{(2 \times \frac{5}{2} - 2)}{(2 \times \frac{5}{2} - 3)} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(\frac{5}{2} - 1)^2} \right\} J(\frac{5}{2}, c) \\
&= \frac{(2c\delta\tau)^2 + \tau^2}{2\cosh(c\delta\pi)} \frac{(2 \times \frac{5}{2} - 2)}{(2 \times \frac{5}{2} - 3)} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(\frac{5}{2} - 1)^2} \right\} \frac{(2 \times \frac{7}{2} - 2)}{(2 \times \frac{7}{2} - 3)} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(\frac{7}{2} - 1)^2} \right\} J(\frac{7}{2}, c) \\
&\vdots \\
&= \frac{(\tau^2 + \frac{(c\delta\tau)^2}{(3/2-1)^2})}{2\cosh(c\delta\pi)} \times \frac{3}{2} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(\frac{5}{2} - 1)^2} \right\} \times \frac{5}{4} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(\frac{7}{2} - 1)^2} \right\} \times \cdots \\
&\quad \cdots \times \frac{(2c-2)}{(2c-3)} \left\{ \tau^2 + \frac{(c\delta\tau)^2}{(c-1)^2} \right\} J(c, c).
\end{aligned}$$

Incidentally, the following formula are known in the literature (Gröbner and Hofreiter (1958)). The method above prove the following formula indirectly.

$$\int_0^{\pi} \exp(-px) \sin^{2m} x dx = \frac{(2m)!(1 - \exp(-p\pi))}{p \prod_{k=1}^{m} (p^2 + (2k)^2)}$$

$$\int_0^{\pi} \exp(-px) \sin^{2m+1} x dx = \frac{(2m+1)!(1 - \exp(-p\pi))}{\prod_{k=0}^{m} (p^2 + (2k+1)^2)}.$$

By the transformation $\frac{x-\mu}{\tau} = \tan\theta$ $\frac{dx}{d\theta} = \tau \frac{1}{\cos^2\theta} = \tau(1 + \tan^2\theta), b = \frac{r+2}{2}$ we have

$$\int_{-\infty}^{\infty} p(x) dx = C \int_{-\infty}^{\infty} \frac{1}{\{(x-\mu)^2 + \tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \tag{2.19}$$

$$\begin{aligned}
&= C \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(2b\delta\theta)}{\{\tau^2(1 + \tan^2\theta)\}^{\frac{r+2}{2}}} \frac{dx}{d\theta} d\theta \\
&= C \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\exp(2b\delta\theta)}{\tau^{r+2} \frac{1}{\cos^{r+2}\theta}} \times \tau \times \frac{1}{\cos^2\theta} d\theta \\
&= C \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp(2b\delta\theta) \frac{\cos^r\theta}{\tau^{r+1}} d\theta
\end{aligned}$$

By the transformation $\lambda = \theta + \pi/2$

$$\int_{-\infty}^{\infty} p(x) dx = C \int_0^{\pi} \frac{\sin^r \lambda}{\tau^{r+1}} \exp(-b\delta\pi + 2b\delta\lambda) d\lambda \tag{2.20}$$

By the formula above, we get the theorem. And the recursive formula is also introduced by the following formula.

$$\begin{aligned}
I[m,0] &= \int \exp(ax)\sin^m x dx \\
&= \frac{1}{a^2+m^2}\{\exp(ax)\sin^{m-1}x(a\sin x - m\cos x) + m(m-1)I[m-2,0]\} \\
&= \frac{1}{(m+1)(m+2)}\{\exp(-ax)\sin^{m+1}x(a\sin x - (m+2)\cos x) \\
&\quad +(a^2+(m+2)^2)I[m+2,0]\}\ , \qquad m \neq -1,-2.
\end{aligned}$$

∎

Theorem 2.2.2 (The Central Moments, Skewness and Kurtosis) *The expectation, variance, third and fourth central moments of the distribution are given by*

$$E[X] = \frac{b\delta\tau}{b-1} + \mu, \quad (1 < b), \tag{2.21}$$

$$VAR[X] = \frac{\tau^2}{2b-3}\left\{1+\left(\frac{b}{b-1}\delta\right)^2\right\}, \quad (3/2 < b), \tag{2.22}$$

$$\begin{aligned}
E[(X-E[X])^3] &= \frac{2b\delta\tau^3}{(b-1)(b-2)(2b-3)}\left\{1+\left(\frac{b}{b-1}\delta\right)^2\right\} \\
&= \frac{2b\delta\tau}{(b-1)(b-2)}VAR[X], \quad (2 < b),
\end{aligned} \tag{2.23}$$

$$\begin{aligned}
E[(X-E[X])^4] &= \frac{3\tau^4}{(2b-3)(2b-5)}\left\{1+\frac{b+2}{b-2}\left(\frac{b\delta}{b-1}\right)^2\right\}\left\{1+\left(\frac{b}{b-1}\delta\right)^2\right\} \\
&= \frac{3\tau^2}{(2b-5)}\left\{1+\frac{b+2}{b-2}\left(\frac{b\delta}{b-1}\right)^2\right\}VAR[X], \quad (5/2 < b).
\end{aligned} \tag{2.24}$$

$$\begin{aligned}
\text{Skewness} = \frac{E[(X-E[X])^3]}{VAR^{3/2}[X]} &= \frac{2b\delta\tau}{(b-1)(b-2)}VAR^{-1/2}[X], \\
&= \frac{2b\sqrt{2b-3}}{(b-1)(b-2)}\frac{\delta}{\sqrt{1+(\frac{b}{b-1}\delta)^2}}
\end{aligned} \tag{2.25}$$

$$\begin{aligned}
\text{Kurtosis} = \frac{E[(X-E[X])^4]}{VAR^2[X]} &= \frac{3\tau^2}{(2b-5)}\left\{1+\frac{b+2}{b-2}\left(\frac{b\delta}{b-1}\right)^2\right\}VAR^{-1}[X] \\
&= \frac{3(2b-3)}{(2b-5)}\frac{\{1+\frac{b+2}{b-2}(\frac{b\delta}{b-1})^2\}}{\{1+(\frac{b}{b-1}\delta)^2\}}.
\end{aligned} \tag{2.26}$$

Proof.

1) The expectation of this distribution is calculated.

$$E[X] = \int_{-\infty}^{\infty} \frac{Cx}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left[\frac{C(x-\mu)}{\{(x-\mu)^2+\tau^2\}^b} + \frac{C\mu}{((x-\mu)^2+\tau^2)^b}\right] \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx \\
&= C\int_{-\infty}^{\infty} \left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]' \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx + \mu \\
&= C\left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&\quad -C\int_{-\infty}^{\infty} \frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1} 2b\delta \times \frac{1}{1+(\frac{x-\mu}{\tau})^2} \times \frac{1}{\tau} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx + \mu \\
&= \frac{b\delta\tau}{b-1}\int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx + \mu \\
&= \frac{b\delta\tau}{b-1} + \mu.
\end{aligned}$$

2) The variance is caluculated below.

$$Var[X] = E[X^2]-E[X]^2 = E[(X-\mu)X+\mu X]-E[X]^2 = E[(X-\mu)X]+E[X](\mu-E[X]).$$

$$E[(X-\mu)X] = \int_{-\infty}^{\infty} (x-\mu)x\frac{C}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]' Cx \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx \\
&= C\left[\frac{1}{2(1-b)}\frac{x}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&\quad -C\int_{-\infty}^{\infty} \left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]\left[x \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}\right]' dx \\
&= -C\int_{-\infty}^{\infty} \frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx \\
&\quad -C\int_{-\infty}^{\infty} \frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1} x2b\delta\frac{1}{1+(\frac{x-\mu}{\tau})^2}\frac{1}{\tau} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx \\
&= \frac{C}{2(b-1)}J(b-1,b) + \frac{C}{2(b-1)}\int_{-\infty}^{\infty} \frac{2bx\delta\tau}{\{(x-\mu)^2+\tau^2\}^b} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}dx \\
&= \frac{C}{2(b-1)} \times \frac{2b-2}{2b-3} \times \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} J(b,b)
\end{aligned}$$

$$
\begin{aligned}
& +\frac{b\delta\tau}{b-1}\int_{-\infty}^{\infty}\frac{Cx\delta\tau}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
= \; & \frac{1}{2b-3}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-1)^2}\right\}+\frac{b\delta\tau}{b-1}\left(\frac{b\delta\tau}{b-1}+\mu\right) \\
= \; & \frac{\tau^2}{2b-3}+\frac{2(b\delta\tau)^2}{(2b-3)(b-1)}+\mu\frac{b\delta\tau}{b-1}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Var[X] &= E[(X-\mu)X]+E[X](\mu-E[X]) \\
&= \frac{\tau^2}{2b-3}+\frac{2(b\delta\tau)^2}{(2b-3)(b-1)}+\mu\frac{b\delta\tau}{b-1}+\left(\mu+\frac{b\delta\tau}{b-1}\right)\left(\mu-\mu-\frac{b\delta\tau}{b-1}\right) \\
&= \frac{\tau^2}{2b-3}+\frac{2(b\delta\tau)^2}{(2b-3)(b-1)}-\left(\frac{b\delta\tau}{b-1}\right)^2 \\
&= \frac{\tau^2}{2b-3}+\frac{(b\delta\tau)^2\{2(b-1)-(2b-3)\}}{(2b-3)(b-1)^2} \\
&= \frac{1}{2b-3}\left\{\tau^2+(\frac{b}{b-1}\delta\tau)^2\right\}.
\end{aligned}
$$

3) Third central moment is caluculated by the following.

$$E[(X-E[x])^3]=E[X^3]-3E[X^2]E[X]+2E^3[X].$$

Firstly, we caluculate $E[X^3]$.

$$E[X^3]=E[(X-\mu)X^2+\mu X^2]=E[(X-\mu)X^2]+\mu E[X^2].$$

$$E[(X-\mu)X^2]=\int_{-\infty}^{\infty}(x-\mu)x^2\frac{C}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx$$

$$
\begin{aligned}
= \; & \int_{-\infty}^{\infty}\left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]' Cx^2\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
= \; & C\left[\frac{1}{2(1-b)}\frac{x^2}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
& -C\int_{-\infty}^{\infty}\left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right][x^2\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}]'dx \\
= \; & -C\int_{-\infty}^{\infty}\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}2x\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
& -C\int_{-\infty}^{\infty}\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}x^2 2b\delta\frac{1}{1+(\frac{x-\mu}{\tau})^2}\frac{1}{\tau}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx
\end{aligned}
$$

$$
\begin{aligned}
= & \frac{1}{b-1} \int_{-\infty}^{\infty} \frac{Cx}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
& + \frac{b\delta\tau}{b-1} \int_{-\infty}^{\infty} \frac{Cx^2}{\{(x-\mu)^2+\tau^2\}^{b}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx.
\end{aligned}
$$

First term is caluculated by the following.

$$
\begin{aligned}
& \int_{-\infty}^{\infty} \frac{Cx}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
= & \int_{-\infty}^{\infty} \frac{C\{(x-\mu)+\mu\}}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
= & \int_{-\infty}^{\infty} \frac{C(x-\mu)}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx + \mu C J(b-1,b) \\
= & \int_{-\infty}^{\infty} \left[\frac{C}{2(2-b)} \{(x-\mu)^2+\tau^2\}^{-b+2} \right]' C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx + \mu C J(b-1,b) \\
= & \left[\frac{C}{2(2-b)} \{(x-\mu)^2+\tau^2\}^{-b+2}]' C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} \right]_{-\infty}^{\infty} \\
& - C \int_{-\infty}^{\infty} \frac{1}{2(2-b)} \{(x-\mu)^2+\tau^2\}^{-b+2} 2b\delta \frac{1}{1+(\frac{x-\mu}{\tau})^2} \frac{1}{\tau} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
& + \mu C J(b-1,b) \\
= & \frac{b\delta\tau}{b-2} \int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx + \mu C J(b-1,b) \\
= & \left(\frac{b\delta\tau}{b-2} + \mu \right) C J(b-1,b) \\
= & \left(\frac{b\delta\tau}{b-2} + \mu \right) \times \frac{(2b-2)}{(2b-3)} \left\{ \tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2} \right\}.
\end{aligned}
$$

Second term is caluculated by the following.

$$
\begin{aligned}
& \int_{-\infty}^{\infty} \frac{Cx^2}{\{(x-\mu)^2+\tau^2\}^{b}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
= & E[X^2] = VAR[X] + E^2[X] \\
= & \frac{1}{2b-3} \left\{ \tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2} \right\} + \left(\frac{b\delta\tau}{b-1} + \mu \right)^2.
\end{aligned}
$$

Accordingly,

$$
E[(X-\mu)X^2] = 2\left(\frac{b\delta\tau}{b-2} + \mu \right) VAR[X] + \frac{b\delta\tau}{b-1} E[X^2].
$$

And,

$$
E[X^3] = E[(X-\mu)X^2] + \mu E[X^2] = 2\left(\frac{b\delta\tau}{b-2} + \mu \right) VAR[X] + \left(\frac{b\delta\tau}{b-1} + \mu \right) E[X^2].
$$

Consequently,

$$
\begin{aligned}
E[(X-E[x])^3] &= E[X^3]-3E[X^2]E[X]+2E^3[X] \\
&= 2\left(\frac{b\delta\tau}{b-2}+\mu\right)VAR[X]+\left(\frac{b\delta\tau}{b-1}+\mu\right)E[X^2]-3E[X^2]E[X]+2E^3[X] \\
&= \frac{2b\delta\tau}{(b-1)(b-2)}\{E[X^2]-E^2[X]\}=\frac{2b\delta\tau}{(b-1)(b-2)}VAR[X].
\end{aligned}
$$

4) Forth central moment is caluculated by the following.

$$E[(X-E[x])^4]=E[X^4]-4E[X^3]E[X]+6E[X^2]E^2[X]-3E^4[X].$$

Firstly, we caluculate $E[X^4]$.

$$E[X^4]=E[(X-\mu)X^3+\mu X^3]=E[(X-\mu)X^3]+\mu E[X^3].$$

$$E[(X-\mu)X^3]=\int_{-\infty}^{\infty}(x-\mu)x^3\frac{C}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty}\left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]' Cx^3\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= C\left[\frac{1}{2(1-b)}\frac{x^3}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&\quad -C\int_{-\infty}^{\infty}\left[\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}\right]\left[x^3\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}\right]' dx \\
&= -C\int_{-\infty}^{\infty}\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}3x^2\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&\quad -C\int_{-\infty}^{\infty}\frac{1}{2(1-b)}\{(x-\mu)^2+\tau^2\}^{-b+1}x^3 2b\delta\frac{1}{1+(\frac{x-\mu}{\tau})^2}\frac{1}{\tau}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= \frac{3}{2(b-1)}\int_{-\infty}^{\infty}\frac{Cx^2}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&\quad +\frac{b\delta\tau}{b-1}\int_{-\infty}^{\infty}\frac{Cx^3}{\{(x-\mu)^2+\tau^2\}^{b}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx.
\end{aligned}
$$

First item is caluculated by the following.

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\frac{Cx^2}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= \int_{-\infty}^{\infty}\frac{C\{(x-\mu)x+\mu x\}}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= \int_{-\infty}^{\infty}\frac{C(x-\mu)x}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2b\delta\arctan(\frac{x-\mu}{\tau})\}dx
\end{aligned}
$$

$$
\begin{aligned}
&+\mu \int_{-\infty}^{\infty} \frac{Cx}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
= &\int_{-\infty}^{\infty} \left[\frac{Cx}{2(2-b)}\{(x-\mu)^2+\tau^2\}^{-b+2}\right]' C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&+\mu \left(\frac{b\delta\tau}{b-2}+\mu\right) \times \frac{(2b-2)}{(2b-3)} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \\
= &\left[\frac{Cx}{2(2-b)}\{(x-\mu)^2+\tau^2\}^{-b+2} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{\infty} \\
&-C \int_{-\infty}^{\infty} \frac{1}{2(2-b)}\{(x-\mu)^2+\tau^2\}^{-b+2} 2b\delta \frac{1}{1+(\frac{x-\mu}{\tau})^2}\frac{1}{\tau} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&+\mu \left(\frac{b\delta\tau}{b-2}+\mu\right) \times \frac{(2b-2)}{(2b-3)} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \\
= &\frac{1}{2(b-2)} C J(b-2,b) + \frac{b\delta\tau}{b-2} \int_{-\infty}^{\infty} \frac{C}{\{(x-\mu)^2+\tau^2\}^{b-1}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
&+\mu \left(\frac{b\delta\tau}{b-2}+\mu\right) \times \frac{(2b-2)}{(2b-3)} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \\
= &\frac{1}{2(b-2)} \times \frac{2b-2}{2b-3} \times \frac{2b-4}{2b-5} \times \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-2)^2}\right\} \\
&+\frac{b\delta\tau}{b-2} \times \frac{2b-2}{2b-3} \left(\frac{b\delta\tau}{b-2}+\mu\right) \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \\
&+\mu \left(\frac{b\delta\tau}{b-2}+\mu\right) \times \frac{(2b-2)}{(2b-3)} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \\
= &\frac{(2b-2)}{(2b-3)} \left(\frac{b\delta\tau}{b-2}+\mu\right)^2 \left\{\tau^2 + (\frac{b\delta\tau}{b-1})^2\right\} \\
&+\frac{(2b-2)}{(2b-3)} \times \frac{1}{2b-5} \times \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-1)^2}\right\} \left\{\tau^2 + \frac{(b\delta\tau)^2}{(b-2)^2}\right\}.
\end{aligned}
$$

Second item is caluculated by the following.

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \frac{Cx^3}{\{(x-\mu)^2+\tau^2\}^{b}} \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\} dx \\
= &E[X^3] = 2\left(\frac{b\delta\tau}{b-2}+\mu\right) VAR[X] + E[X]E[X^2].
\end{aligned}
$$

Accordingly,

$$
\begin{aligned}
&E[(X-\mu)X^2] \\
= &\frac{3}{2(b-1)} \left[\frac{(2b-2)}{(2b-3)} \left(\frac{b\delta\tau}{b-2}+\mu\right)^2 \left\{\tau^2 + (\frac{b\delta\tau}{b-1})^2\right\}\right.
\end{aligned}
$$

$$+\frac{(2b-2)}{(2b-3)}\frac{1}{2b-5}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-1)^2}\right\}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-2)^2}\right\}\Bigg]$$
$$+\frac{b\delta\tau}{b-1}E[X^3].$$

And,

$$\begin{aligned}
E[X^4] &= E[(X-\mu)X^3]+\mu E[X^3] \\
&= \frac{3}{2(b-1)}\left[\frac{(2b-2)}{(2b-3)}\left(\frac{b\delta\tau}{b-2}+\mu\right)^2\left\{\tau^2+(\frac{b\delta\tau}{b-1})^2\right\}\right. \\
&\quad \left.+\frac{(2b-2)}{(2b-3)}\times\frac{1}{2b-5}\times\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-1)^2}\right\}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-2)^2}\right\}\right]+E[X]E[X^3].
\end{aligned}$$

Consequently,

$$\begin{aligned}
E[(X-E[x])^4] &= = E[X^4]-4E[X^3]E[X]+6E[X^2]E^2[X]-3E^4[X] \\
&= 3VAR[X]\left[(\frac{b\delta\tau}{b-2}+\mu)^2+\frac{1}{2b-5}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-2)^2}\right\}\right] \\
&\quad -3E[X^3]E[X]+6E[X^2]E^2[X]-3E^4[X] \\
&= 3VAR[X]\left[\frac{1}{2b-5}\left\{\tau^2+\frac{(b\delta\tau)^2}{(b-2)^2}\right\}+\left\{\frac{b\delta\tau}{(b-1)(b-2)}\right\}^2\right] \\
&= 3VAR[X]\left\{\frac{1}{2b-5}\tau^2+\frac{b+2}{(2b-5)(b-2)}(\frac{b\delta\tau}{b-1})^2\right\}.
\end{aligned}$$

2.3 Cumulative Distribution Functions

The cumulative distribution functions are evaluated by the following theorem.

Theorem 2.3.1 (Cumulative Distribution Functions) *The cumulative distribution functions are given by*

(1)For $b=1$

$$P(x)=\int_{-\infty}^{y}p(x)dx=\frac{1}{2\sinh(\delta\pi)}[\exp\{2\delta\arctan(\frac{y-\mu}{\tau})\}-\exp(-\delta\pi)]. \quad (2.27)$$

(2)For $b=\frac{3}{2}$

$$P(x)=\int_{-\infty}^{y}p(x)dx=\frac{1}{2\cosh(\frac{3\delta\pi}{2})}\left[\frac{\frac{3}{2}\delta\tau+\frac{y-\mu}{\tau}}{\sqrt{1+(\frac{y-\mu}{\tau})^2}}\exp\{3\delta\arctan(\frac{y-\mu}{\tau})\}+\exp(-\frac{3\delta\pi}{2})\right]. \quad (2.28)$$

(3)Recursive Formula

$$\left\{1+\frac{(c\delta)^2}{(b-1)^2}\right\}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx$$

$$\begin{aligned}
&= \frac{1}{2(b-1)\tau^2}\left(y-\mu+\frac{c\delta\tau}{b-1}\right)\frac{1}{\{(y-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\} \\
&\quad +\frac{1}{\tau^2}\frac{(2b-3)}{(2b-2)}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx. \qquad (2.29)
\end{aligned}$$

Proof. (1)For $b=1$

$$p(x)=\frac{\delta\tau}{\sinh(\delta\pi)}\frac{\exp\{2\delta\arctan(\frac{x-\mu}{\tau})\}}{(x-\mu)^2+\tau^2}.$$

$$\begin{aligned}
\int_{-\infty}^{y}p(x)dx &= \frac{\delta\tau}{\sinh(\delta\pi)}\int_{-\infty}^{y}\frac{\exp\{2\delta\arctan(\frac{x-\mu}{\tau})\}}{(x-\mu)^2+\tau^2}dx \\
&= \frac{\delta\tau}{\sinh(\delta\pi)}\times\frac{1}{2\delta\tau}\left[\exp\{2\delta\arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{y} \\
&= \frac{\delta\tau}{\sinh(\delta\pi)}\frac{\exp\{2\delta\arctan(\frac{x-\mu}{\tau})\}}{(x-\mu)^2+\tau^2}
\end{aligned}$$

(2)For $b=\frac{3}{2}$

$$p(x)=\frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\frac{\exp\{3\delta\arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2+\tau^2\}^{\frac{3}{2}}}. \qquad (2.30)$$

We considers $\frac{x-\mu}{\tau}=\tan\lambda$, $\frac{dx}{d\lambda}=\tau\frac{1}{\cos^2\lambda}=\tau(1+\tan^2\lambda)$

$$\begin{aligned}
\int_{-\infty}^{y}p(x)dx &= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\int_{-\frac{\pi}{2}}^{\theta}\frac{\exp(3\delta\lambda)}{\{\tau^2(1+\tan^2\lambda)\}^{\frac{3}{2}}}\frac{dx}{d\lambda}d\lambda \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}C\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{\exp(3\delta\theta)}{\{\tau^2(1+\tan^2\theta)\}^{\frac{3}{2}}}\frac{dx}{d\theta}d\theta \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\int_{-\frac{\pi}{2}}^{\theta}\frac{\exp(3\delta\lambda)}{\tau^3\frac{1}{\cos^3\lambda}}\times\tau\times\frac{1}{\cos^2\lambda}d\lambda \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\int_{-\frac{\pi}{2}}^{\theta}\frac{\cos\lambda}{\tau^2}\exp(3\delta\lambda)d\lambda \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\frac{1}{\tau^2}\left[\frac{\exp(3\delta\lambda)}{(3\delta)^2+1^2}(3\delta\cos\lambda+\sin\lambda)\right]_{-\frac{\pi}{2}}^{\theta} \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\frac{1}{\tau^2}\left[\frac{\exp(\frac{3\delta\pi}{2})\{3\delta\cos(\arctan\frac{y-\mu}{\tau})+\sin(\arctan\frac{y-\mu}{\tau})\}+\exp(-\frac{3\delta\pi}{2})}{(3\delta)^2+1}\right] \\
&= \frac{\{(3\delta\tau)^2+\tau^2\}}{2\cosh(\frac{3\delta\pi}{2})}\frac{\exp\{3\delta\arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2+\tau^2\}^{\frac{3}{2}}}.
\end{aligned}$$

Because of the following formula,

$$\cos(\arctan\frac{y-\mu}{\tau}) = \cos(\arccos\frac{1}{\sqrt{1+(\frac{y-\mu}{\tau})^2}}),$$

$$\sin(\arctan\frac{y-\mu}{\tau}) = \sin(\arcsin\frac{\frac{y-\mu}{\tau}}{\sqrt{1+(\frac{y-\mu}{\tau})^2}}).$$

(3) Recursive Formula

Firstly, we obtain the following formula.

$$\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx$$

$$\begin{aligned}
=\ & \int_{-\infty}^{y}\frac{1}{\tau^2}\left[\left\{\frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}}\right\}' \right.\\
& \left. +\frac{2b-3}{2b-2}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\right]\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\\
=\ & \frac{1}{\tau^2}\left[\frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}\right]_{-\infty}^{y}\\
& -\frac{1}{\tau^2}\int_{-\infty}^{y}\frac{x-\mu}{(2b-2)\{(x-\mu)^2+\tau^2\}^{b-1}}\times\frac{2c\delta\tau}{(x-\mu)^2+\tau^2}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\\
& +\frac{1}{\tau^2}\frac{2b-3}{2b-2}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\\
=\ & \frac{1}{\tau^2}\left[\frac{y-\mu}{(2b-2)\{(y-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\}\right]\\
& -\frac{1}{\tau^2}\frac{c\delta\tau}{b-1}\int_{-\infty}^{y}\frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^{b}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\\
& +\frac{1}{\tau^2}\frac{2b-3}{2b-2}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx.
\end{aligned}$$

Next, in the case of $b \geq \frac{3}{2}$ the formula below is introduced.

$$\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx.$$

$$\begin{aligned}
=\ & \int_{-\infty}^{y}\frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau}\times\frac{2c\delta\tau}{(x-\mu)^2+\tau^2}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\\
=\ & \int_{-\infty}^{y}\frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau}[\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}]'dx
\end{aligned}$$

$$
\begin{aligned}
&= [\frac{\{(x-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}]_{-\infty}^{y} \\
&\quad -\int_{-\infty}^{y}\frac{(-b+1)}{2c\delta\tau}\frac{2(x-\mu)}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= \left[\frac{\{(y-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\}\right] \\
&\quad +\frac{b-1}{c\delta\tau}\int_{-\infty}^{\infty}\frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx
\end{aligned}
$$

We input second formula to first formula.

$$
\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx
$$

$$
\begin{aligned}
&= \frac{1}{\tau^2}\left[\frac{y-\mu}{(2b-2)\{(y-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\}\right] \\
&\quad -\frac{1}{\tau^2}\frac{c\delta\tau}{b-1}\times\frac{c\delta\tau}{b-1}\left[\int_{-\infty}^{y}\frac{x-\mu}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx\right. \\
&\quad \left.-\frac{\{(y-\mu)^2+\tau^2\}^{-b+1}}{2c\delta\tau}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\}\right] \\
&\quad +\frac{1}{\tau^2}\frac{2b-3}{2b-2}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&= \frac{1}{2\tau^2(b-1)}\left[y-\mu+\frac{c\delta\tau}{b-1}\right]\frac{\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\}}{\{(y-\mu)^2+\tau^2\}^{b-1}} \\
&\quad -\frac{1}{\tau^2}(\frac{c\delta\tau}{b-1})^2\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx \\
&\quad +\frac{1}{\tau^2}(\frac{2b-3}{2b-2})\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx.
\end{aligned}
$$

Consequently, we obtain the following recursive formula.

$$
\left\{1+\frac{(c\delta)^2}{(b-1)^2}\right\}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^b}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx
$$

$$
\begin{aligned}
&= \frac{1}{2(b-1)\tau^2}\left(y-\mu+\frac{c\delta\tau}{b-1}\right)\frac{1}{\{(y-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{y-\mu}{\tau})\} \\
&\quad +\frac{1}{\tau^2}\frac{(2b-3)}{(2b-2)}\int_{-\infty}^{y}\frac{1}{\{(x-\mu)^2+\tau^2\}^{b-1}}\exp\{2c\delta\arctan(\frac{x-\mu}{\tau})\}dx.
\end{aligned}
$$

We can caluculate cumulative distribution functions in the same way of normalizing constant by using this results.

Chapter 3

Cross-Sectional-Skew-Dependent Distribution Models for Industry Returns in Japanese Stock Market

3.1 Introduction

In the stock market, many empirical researches have observed an industry-effect in the movement of prices or returns, meaning that industry returns (returns based on industry indices) move rather independently. However, the author (1992) observed that there are also co-movement (correlations) of industry returns and an asymmetry of the co-movements. In fact, when the market price is in a downward trend, almost all industry prices tend to move downwards together. On the other hand, when the market is in an upward trend, the two typical cases are observed ; the case that only some specific industry returns move upwards while others do not move much, and the case that all the industry returns move together and form an upward trend. This implies that the correlations among industry returns and the skewness of the cross-sectional distributions of industry returns will change with time t. In other words, there may be some information in these changes of the co-movements and asymmetry of cross-sectional industry returns over time for predicting future market average returns. This motivates our study.

In the literature, this part has not much been paid an attention to, though there are many time series studies on the distributional aspects of market and individual returns. For examples, Fama (1965) and Mandelbrot (1966) suggested the use of such a heavy-tailed distribution as stable Paretian distribution as a model for daily returns. Kariya, et al. (1995) extensively studied on this feature as well as some nonlinear features in the Japanese market (see also Kariya (1993)), which showed that the distribution of market returns is not only leptokurtic but also skewed. In our time series analysis below, to take into account this feature, we use the Pearson type IV distribution which is a non-central version of the type VII distribution and forms a broad class of distributions including t-distribution and Cauchy distribution (Pearson, K (1914), Pearson, E.S and Hartrey (1954), Pearson, E.S (1963), Johnson and Kotz (1970)).

To describe our problem more specifically, suppose there are N industries in the stock market, let $r_{j,t}$ denote the j-th weekly industry return at t, and let the averaged return of all the industry returns at t be

$$R_t = \frac{1}{N}\sum_{i=1}^{N} r_{i,t}, \tag{3.1}$$

which we sometimes call market return. It is noted that this averaged return is not equal to the so-called market return based on a market index. The skewness of the cross-sectional distribution of industry returns is defined by

$$S_t = \frac{m_{3,t}}{m_{2,t}^{3/2}} \qquad \text{with} \quad m_{k,t} = \frac{1}{N}\sum_{i=1}^{N}(r_{i,t} - R_t)^k. \tag{3.2}$$

Here $r_{i,t}$ is the weekly rate of return obtained as the average of daily NIKKEI 36 Industry Indices over each week. Hence N equals to 36, and R_t is not such an index return as the NIKKEI 225 Index return.

In this chapter, we mainly study on a predictive power of past values of the cross-sectional skewness variable S for future values of the averaged return R. In predicting the averaged return, we use an AR-GARCH type time series model with the AIC criterion for model selection, and examine that the skewness of the cross-sectional distribution over industries has a predictive power for future averaged returns R_t's. We also present some state-dependent models to find that the conditional variances depend on some shape parameters such as skewness, which may be functions of economic fundamentals, in addition to the past conditional variances and past errors.

In section 3.2, our framework is introduced including Pearson IV distribution. In section 3.3, we make our empirical analysis with various models and draw our conclusion.

3.2 Model

To study on an effect of past cross-sectional skewness S_{t-j}'s on the averaged return R_t, we start with the conditional likelihood function by

$$\prod_{n=K+1}^{T} p(R_n \mid \theta, R_{n-1}, \dots, R_1, S_{n-1}, \dots, S_0). \tag{3.3}$$

This formulation of the conditional likelihood ignores the following term

$$\prod_{n=K+1}^{T} p(S_{n-1} \mid \theta, R_{n-1}, \dots, R_1, S_{n-2}, \dots, S_0) \tag{3.4}$$

in the unconditional likelihood. This point is described by the following.

We consider the joint distribution of $R_T, \dots, R_1, S_{T-1}, \dots, S_0$. This is an ordinary likelihood function. As can be seen below, this joint distribution can be expressed by using conditional probability distributions.

$$\begin{aligned} & p(R_T, \dots, R_1, S_{T-1}, \dots, S_0 \mid \theta) \\ & = p(R_T, S_{T-1} \mid \theta, R_{T-1}, \dots, R_1, S_{T-2}, \dots, S_0) p(R_{T-1}, \dots, R_1, S_{T-2}, \dots, S_0 \mid \theta) \\ & = p(R_K, \dots, R_1, S_{K-1}, \dots, S_0 \mid \theta) \prod_{n=K+1}^{T} p(R_n, S_{n-1} \mid \theta, R_{n-1}, \dots, R_1, S_{n-2}, \dots, S_0). \end{aligned} \tag{3.5}$$

Here the second term in the right hand side of (3.5) can be expressed as

$$p(S_{n-1} \mid \theta, R_{n-1}, \dots, R_1, S_{n-2}, \dots, S_0) p(R_n \mid \theta, R_{n-1}, \dots, R_1, S_{n-1}, \dots, S_0)$$

Then, by taking logarithm, we have

$$\begin{aligned} \log p(R_T, \dots, R_1, S_{T-1}, \dots, S_0 \mid \theta) &= \log p(R_K, \dots, R_1, S_{K-1}, \dots, S_0 \mid \theta) \\ &\quad + \sum_{n=K+1}^{T} \log p(S_{n-1} \mid \theta, R_{n-1}, \dots, R_1, S_{n-2}, \dots, S_0) \\ &\quad + \sum_{n=K+1}^{T} \log p(R_n \mid \theta, R_{n-1}, \dots, R_1, S_{n-1}, \dots, S_0) \end{aligned} \tag{3.6}$$

The second term in the right hand side of (3.6) is ignored for our model. Therefore our log-likelihood function is the third term. Further, for simplicity, we also omit the first term. This is justified since the sample size T is large (in our particular example $T = 511, K = 5$). Thus, we regard the third term as a reasonable approximation to the log-likelihood function.

In addition, we assume the following Markov property in (3.6) ;

$$p(R_n \mid \theta, R_{n-1}, \dots, R_1, S_{n-1}, \dots, S_0) = p(R_n \mid \theta, R_{n-1}, R_{n-2}, S_{n-1}). \tag{3.7}$$

It is noted that our likelihood is used for estimation via the Maximum Likelihood Method and model selection via the AIC (Akaike (1973, 1974), Sakamoto, Ishiguro, and Kitagawa (1986)) ;

$$\begin{aligned}\text{AIC} &= -2 \times (\text{maximum log likelihood of the model}) \\ &\quad +2 \times (\text{number of free parameters of the model}).\end{aligned}$$

The models we use to take into account an effect of skewness on prediction are as follows.

[**Model 1** : AR model with skewness]

$$R_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i R_{t-i} + \sum_{j=1}^{l} \beta_j S_{t-j} + e_t, \tag{3.8}$$

where $e_t \sim N(0, \sigma_1^2)$.

[**Model 2** : AR-ARCH model with skewness]

$$R_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i R_{t-i} + \sum_{j=1}^{l} \beta_j S_{t-j} + u_t, \tag{3.9}$$

where $u_t \mid \Omega_{t-1} \sim N(0, h_{1,t})$ and the conditional variance $h_{1,t}$ is given by

$$h_{1,t} = \gamma_1 + \gamma_2 u_{t-1}^2. \tag{3.10}$$

[**Model 3** : AR-GARCH model with skewness]

$$R_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i R_{t-i} + \sum_{j=1}^{l} \beta_j S_{t-j} + e_t, \tag{3.11}$$

where $e_t \mid \Omega_{t-1} \sim N(0, h_{2,t})$ and the conditional variance $h_{2,t}$ is given by

$$h_{2,t} = \gamma_3 + \gamma_4 e_{t-1}^2 + \gamma_5 h_{2,t-1}. \tag{3.12}$$

These three models are based on the normal assumption for errors. We also use the models with the type IV Pearson family of probability distributions for errors, whose density $p(x)$ is given by the solution to the differential equation

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta)}{(\frac{x-\mu}{\tau})^2 + 1}, \tag{3.13}$$

where δ is called the non-centrality parameter. The density is given by

$$p(x|\mu, \tau, \delta, b) = \frac{C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2 + \tau^2\}^b}, \tag{3.14}$$

where $b > \frac{1}{2}$, $\tau > 0$, $-\infty < \delta < \infty$. Note that, if $\delta = 0$, then the solution defines the type VII distributions of Pearson system.

In the past research, this distribution is not introduced in this way, but used as an approximation to non-central t distribution by using moment ratio (Merington and Pearson (1958)). And this type IV family is not used for actual data analysis because of its difficulty in computation. The analytic solution of the normalizing constant C is obtained by using a recursive formula for $b = n$ and $b = n+1/2$, $n = 1, 2, 3\ldots$ (Chapter 2). The recursive formula as well as the first four moment, which specify the distribution of Pearson System, are given in Chapter 2. We use this distribution to model the conditional distribution $p(R_t \mid \theta, R_{t-1}, R_{t-2}, S_{t-1})$.

[**Model 4** : State-dependent model under type IV distribution]

$$p(x|\mu_t, \tau_t, \delta_t, b_t) = C_t \exp\{2b_t\delta_t \arctan(\frac{x-\mu_t}{\tau_t})\}/\{(x-\mu_t)^2 + \tau_t^2\}^{b_t}. \tag{3.15}$$

The shape of the distribution varies depending on the skewness, S_{t-1}. Therefore, μ, τ, δ, and b depend on t. For simplicity, we use the quadratic formulation of S_{t-1} ;

$$\begin{aligned}
\mu_t &= a_1 + a_2 S_{t-1} + a_3 S_{t-1}^2 + a_{13} R_{t-1} + a_{14} R_{t-2} \\
\tau_t &= a_4 + a_5 S_{t-1} + a_6 S_{t-1}^2 \\
\delta_t &= a_7 + a_8 S_{t-1} + a_9 S_{t-1}^2 \\
b_t &= a_{10} + a_{11} S_{t-1} + a_{12} S_{t-1}^2.
\end{aligned}$$

[**Model 5** : State-dependent model under type IV distribution with τ_t, τ_{t-1} and error term]

$$p(x|\mu_t, \tau_t, \delta_t, b_t) = C_t \exp\left\{2b_t\delta_t \arctan\left(\frac{x-\mu_t}{\tau_t}\right)\right\}/\{(x-\mu_t)^2 + \tau_t^2\}^{b_t}, \tag{3.16}$$

$$\begin{aligned}
\mu_t &= a_1 + a_2 R_{t-1} \\
\tau_t^2 &= a_3 + a_4 (R_{t-1} - \mu_{t-1})^2 + a_5 \tau_{t-1}^2 \\
\delta_t &= a_6 + a_7 S_{t-1} + a_8 S_{t-1}^2 \\
b_t &= a_9 + a_{10} S_{t-1} + a_{11} S_{t-1}^2.
\end{aligned}$$

The difference between Model 4 and 5 lies in the formulation of the scale parameter τ_t.

To estimate the parameters and identify the models, the procedure described in Kitagawa (1993) is used, where pseudo-Newton-Raphson method. Furthermore, We adopt DE (double exponential formula) to calculate the normalizing constant by numerical integration for any b (Mori (1987)). We examine the accuracy of numerical integration by using the analytical solution (Chapter 2).

3.3 Empirical Results

3.3.1 Data

Our data consist of weekly returns from January, 1983 to December, 1992 with 511 observations. In Figure 3.1 and Figure 3.2, the time series of the averaged returns $\{R_t\}$ and skewness $\{S_t\}$ are plotted. According to these figures, in the first half of the time interval, certain industry sectors caused an increase of the market price and the low volatility of averaged returns, which is characterized by wide range of skewness. Contrarily, in the latter half, a co-movement of industry returns formed a downward trend of the market price and the high volatility of average returns, which is charaterized by narrow range of skewness. A period in 1989, which is characterized by narrow range of skewness, low volatility of averaged returns and the increase of market price, is an exception to the asymmetry, which was probably caused by a huge volume of the index funds invested on NIKKEI 225 or TOPIX in this period.

In Figure 3.3, the distribution of R_t is given with Pearson IV distribution fitted, where the parameters are estimated
$\mu = 0.93858$, $\tau = 2.22641$, $\delta = -0.15801$, $b = 1.86462$.

Figure 3.1: The time series of the averaged returns

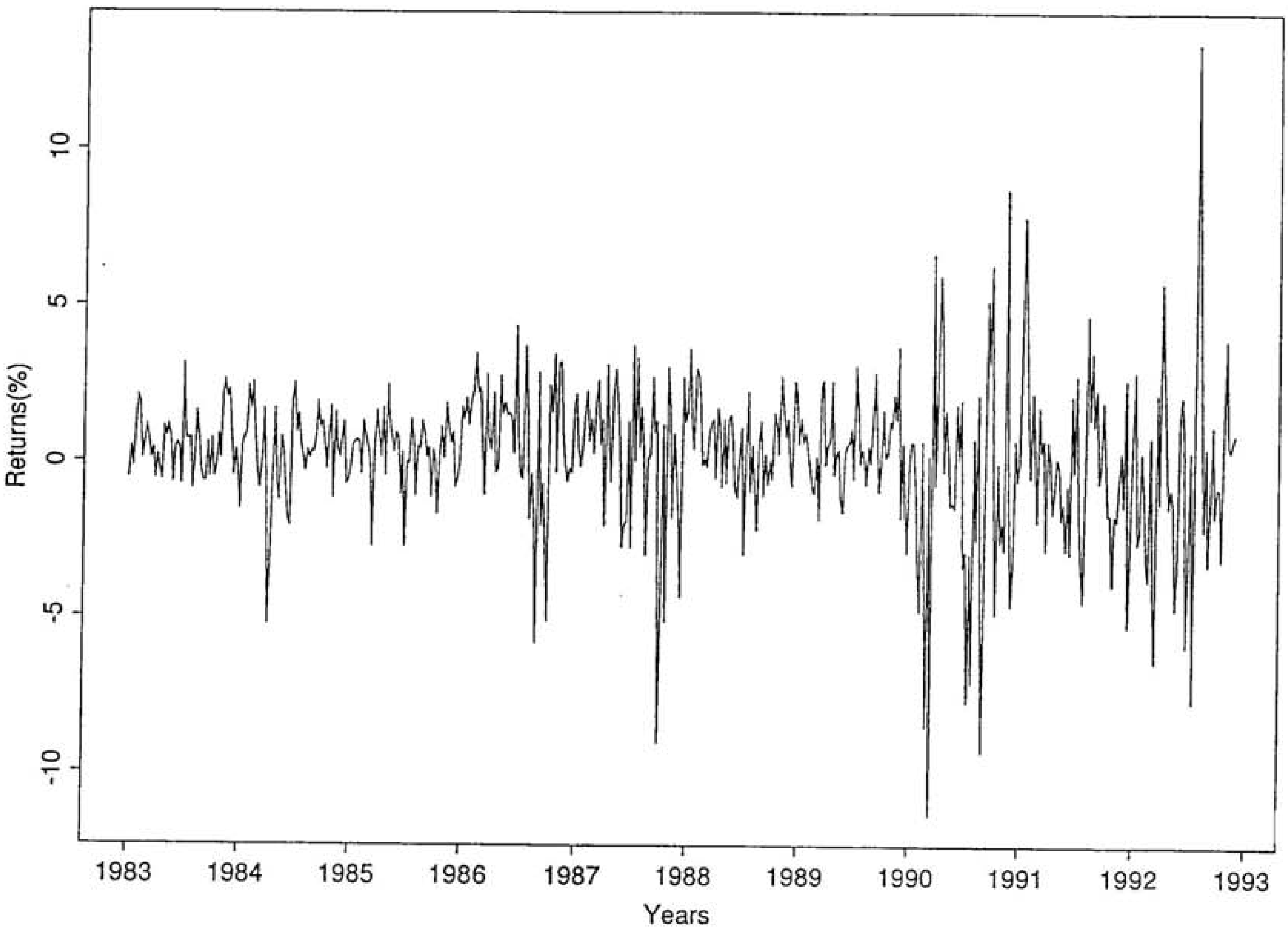


Figure 3.2: The time series of skewness

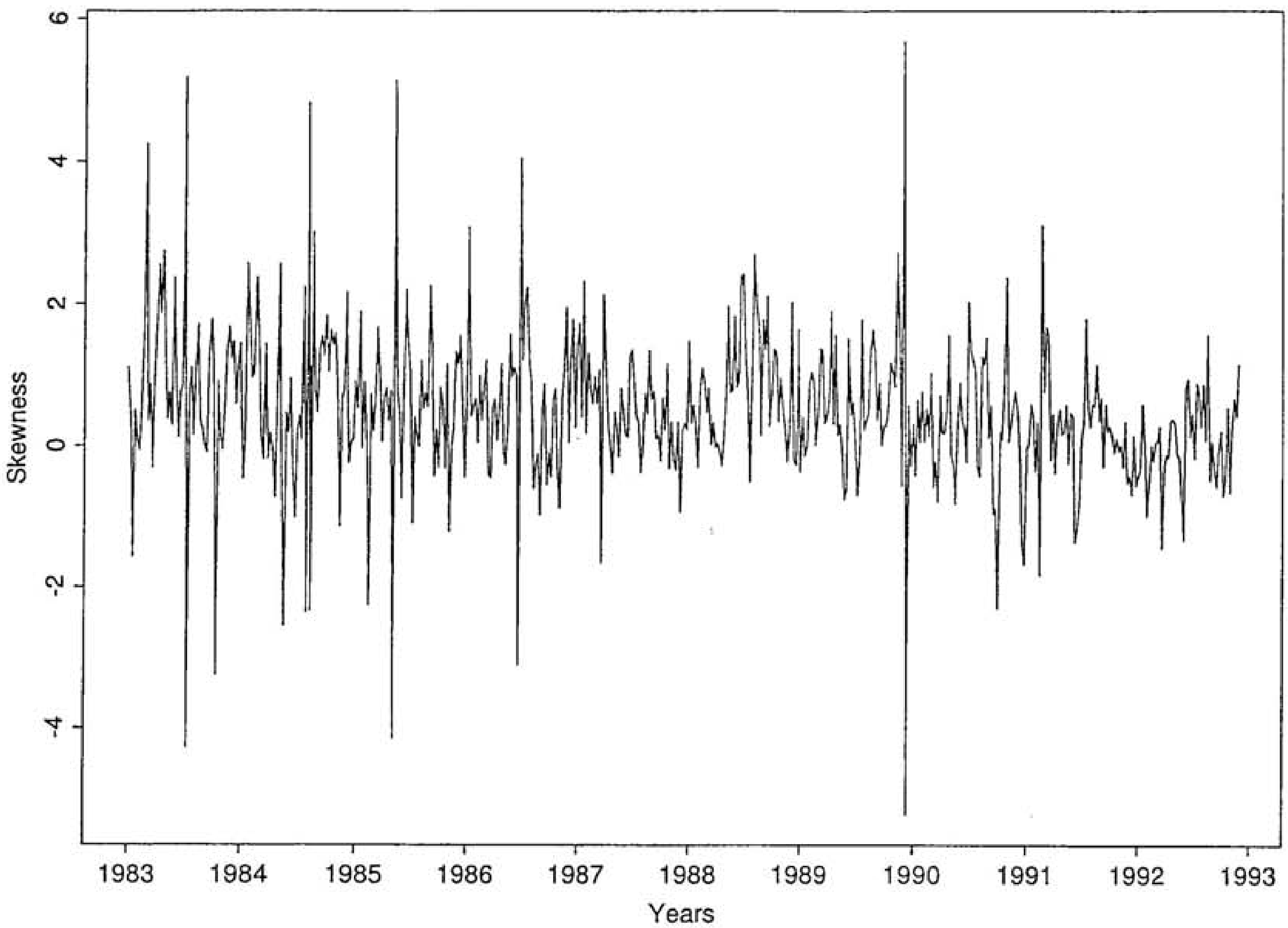
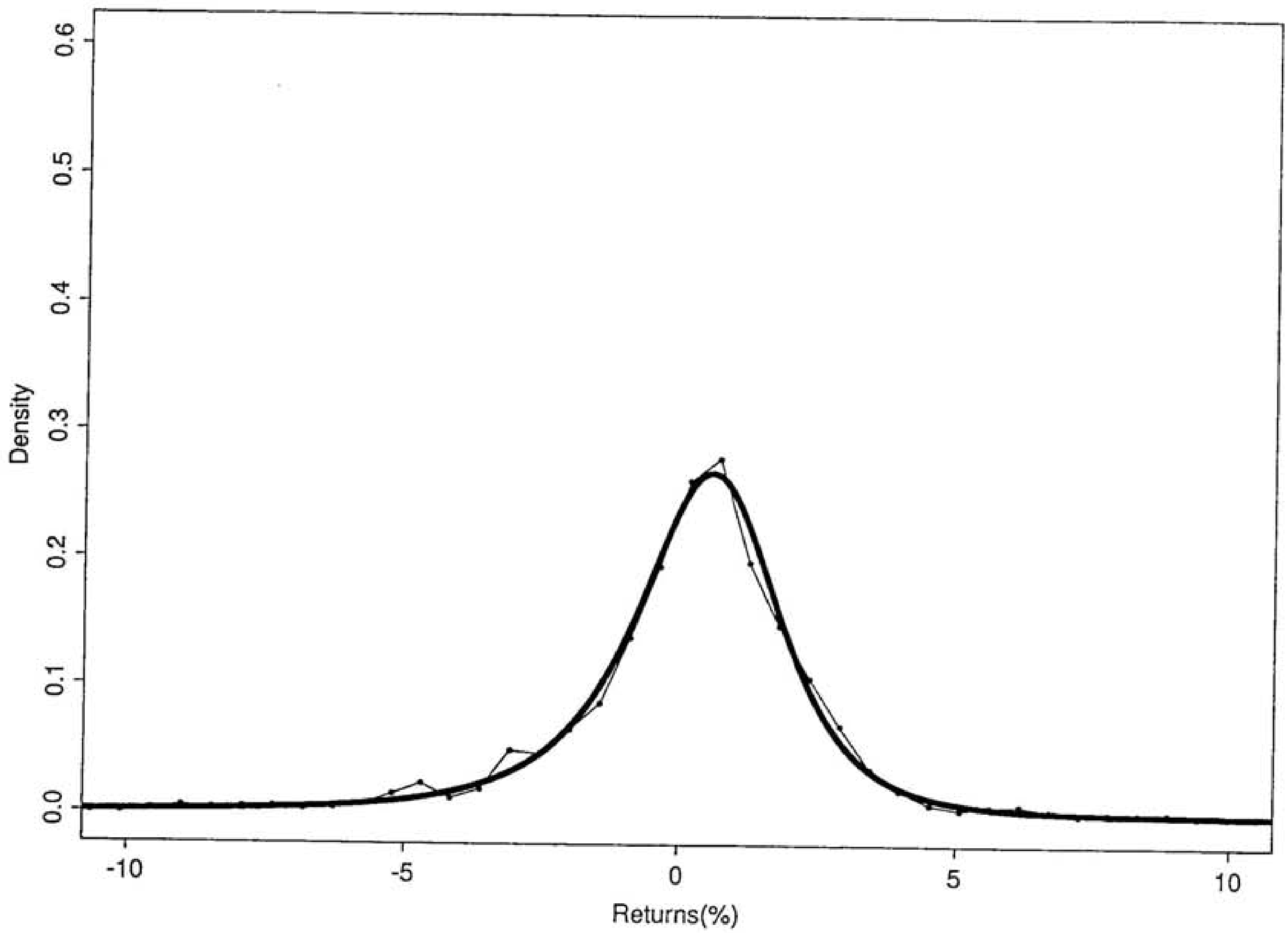


Figure 3.3: The distribution of returns



3.3.2 Comparison of Models

To investigate the predictive power of $\{S_t\}$ for $\{R_t\}$, we fit the models presented in section 3.2.

First, we fit Model 1. With $k = 5$ and $l = 5$ fixed, all possible submodels are considered by the AIC. Table 3.1 summarizes the results.

In Model 1, the model given by,

$$R_t = 0.247520R_{t-1} - 0.187468S_{t-1} + 0.254871 + \varepsilon_t, \quad \varepsilon_t \sim N(0, 5.018) \tag{3.17}$$

attains the minimum of AIC among all possible submodels. On the other hand, when the skewness is not introduced into the model ($l = 0$, i.e. $\beta_j = 0$), the AR(1) with constant term given by,

$$R_t = 0.228235R_{t-1} + 0.165205 + \varepsilon_t, \quad \varepsilon_t \sim N(0, 5.057)$$

attains the minimum of AIC. The table also indicates that under the presence of R_{t-1}, the information of skewness S_{t-i} is more informative than R_{t-i} because both the variance of residuals and the AIC are smaller than the case without the skewness variables. Instead of skewness, we also tried models with cross-sectional standard deviation or kurtosis. The AIC values of AR(1) with s.t.d and kurtosis are 2260.1 and 2261.4 respectively, which will imply that the cross-sectional skewness is better than these variables.

Table 3.1: The results of submodels

		σ^2	AIC
(1)	$\alpha_0, \alpha_1, \beta_1$	5.018	2258.1
(2)	$\alpha_0, \alpha_1, \beta_1, \beta_5$	5.002	2258.5
(3)	$\alpha_0, \alpha_1, \beta_1, \beta_4$	5.005	2258.9
(4)	$\alpha_0, \alpha_1, \alpha_2, \beta_1$	5.008	2259.1
(5)	$\alpha_1, \beta_1, \beta_5$	5.030	2259.4
	$\vdots$	$\vdots$	$\vdots$
	α_0, α_1	5.057	2260.1
	$\alpha_0, \alpha_1, \alpha_2$	5.047	2263.9

Second, we also fit Model 2 (AR-ARCH) and Model 3 (AR-GARCH). The AIC of AR-ARCH(1,1) and AR-ARCH(2,1) without S_{t-1} are 2170.5, 2172.3 respectively, and smaller than the minimum AIC in Model 1.

On the other hand, the AR-ARCH model with S_{t-1},

$$\begin{aligned} R_t &= 0.2121 + 0.4490 R_{t-1} - 0.1518 S_{t-1} + u_t, \\ h_t &= 2.5096 + 0.6541 u_{t-1}^2, \end{aligned}$$

has smaller value of AIC, 2167.8 than these models. The AIC of AR-GARCH(1,1) and AR-GARCH(2,1) are 2072.0, 2073.7 respectively. But the GARCH model with S_{t-1},

$$\begin{aligned} R_t &= 0.4062 + 0.2767 R_{t-1} - 0.1202 S_{t-1} + e_t, \\ h_t &= 0.1340 + 0.1794 e_{t-1}^2 + 0.8083 h_{t-1}, \end{aligned}$$

has further smaller value of AIC, 2069.3.

Consequently the skewness variable S_{t-1} under the presence of R_{t-1} has more information than other lagged values, R_{t-i} in these models and will give a predictive power.

Thirdly, we consider Model 4 with errors following Pearson type IV distribution. In the case that both of a_{13} and a_{14} equal to zero in Model 4, the AIC's of the some of the models are shown in Table 3.2.

Table 3.2: The results of Model 4 without R_{t-1} and R_{t-2}

	Model	AIC
(a)	b is quadratic, δ is linear, others are constant	2143.0
(b)	b is quadratic, others are constant	2144.3
(c)	b and δ are quadratic, others are constant	2144.6
(d)	b, δ, and τ are quadratic, others are constant	2145.5
(e)	All parameters are quadratic function of skewness	2146.3
(f)	τ is quadratic, others are constant	2149.0
(g)	All parameters are constant	2157.5
(h)	δ is quadratic, others are constant	2158.6

The estimated parameters of the AIC best model (a) are

$$\begin{aligned} &\mu_t = 1.09006, \quad \tau_t = 2.39850, \quad \delta_t = -0.17558 - 0.02898 S_{t-1}, \\ &b_t = 1.78561 + 0.26733 S_{t-1} + 0.14339 S_{t-1}^2. \end{aligned}$$

In all of the prefered models by the AIC, the parameter b is quadratic function of S_{t-1}. This suggests that the market tends to be unstable when the absolute value of S_{t-1} is small, and stable when it is large.

When S_{t-1} is around zero, the value of δ is negative and the market is unstable. It will imply that no-preference over industries causes a decrease of the market price. When S_{t-1} is negative due to increases of most industry prices, the return distribution of the next week tends to be symmetric. But, when it is positive, the prices of some industries are increasing with that the market price increasing, or the prices of most industries decline with the market price declining, the return distribution of the next week tends to be asymmetric. In the latter case, since the value of b is high, the standard deviation of the distribution is low (stable).

Next, we fit Model 4 with R_{t-1} and R_{t-2}. The estimated models are summarized in Table 3.3, where AR(1) and AR(2) indicate $\mu_t = a_1 + a_{13}R_{t-1}$ and $\mu_t = a_1 + a_{13}R_{t-1} + a_{14}R_{t-2}$ respectively.

Table 3.3: The results of Model 4 with R_{t-1} and R_{t-2}

	Model	Order	AIC
(a)	b is quardratic, δ is linear, others are constant	AR(1)	2118.0
(b)	b and δ are quadratic, others are constant	AR(1)	2119.8
(c)	b, δ, and τ are quadratic, others are constant	AR(1)	2120.8
(d)	All parameters are quadratic function of skewness	AR(1)	2120.8
(e)	b and δ are quadratic, others are constant	AR(2)	2121.8
(f)	All parameters are quadratic function of skewness	AR(2)	2123.0
(g)	b is quadratic, others are constant	AR(1)	2125.4
(h)	τ is quadratic, others are constant	AR(1)	2127.1
(i)	All parameters are constant	AR(1)	2137.0

The estimated parameters of the AIC best model (a) with AR(1) are

$$\begin{aligned}\mu_t &= 0.88404 + 0.22741R_{t-1}, \quad \tau_t = 2.38417,\\ \delta_t &= -0.13664 - 0.05210S_{t-1},\\ b_t &= 1.81264 + 0.29279S_{t-1} + 0.13521S_{t-1}^2.\end{aligned}$$

For this AIC best model, b is quadratic and δ is linear. The coefficient of $R_{t-1}(0.22741)$ is close to the one of the AR model (0.247520) shown in Model 1. The reduction of the AIC values from the ones in the previous subsection is significant.

Finally, we consider Model 5 whose distribution depends not only on the skewness S_{t-1} and averaged return R_{t-1} but also on the error term and τ_{t-1} as follows:

$$
\begin{aligned}
\mu_t &= a_1 + a_2 R_{t-1} \\
\tau_t^2 &= a_3 + a_4(R_{t-1} - \mu_{t-1})^2 + a_5\tau_{t-1}^2 \\
\delta_t &= a_6 + a_7 S_{t-1} + a_8 S_{t-1}^2 \\
b_t &= a_9 + a_{10} S_{t-1} + a_{11} S_{t-1}^2.
\end{aligned}
$$

Firstly, we adopt the model only with error term like AR-ARCH model. The best AIC value among them is 2086.4 and the estimated parameters are

$$
\begin{aligned}
\mu_t &= 0.66419 + 0.28628 R_{t-1} \\
\tau_t^2 &= 4.53618 + 0.80384(R_{t-1} - \mu_{t-1})^2 \\
\delta_t &= -0.09269 - 0.04701 S_{t-1} \\
b_t &= 2.07579 + 0.25633 S_{t-1} + 0.10831 S_{t-1}^2.
\end{aligned}
$$

We can improve the model with b being quadratic and δ linear and AR(1), by introducing error term to τ_t. The qualitative characterestics of b and δ are the same as the ones in model 4. The coefficient of error term (0.80384) is similar to the AR-ARCH term (0.65).

Next, we adopt the model with both error term and the previous term of τ like AR-GARCH model. The best AIC value among them is 2030.7 and the estimated parameters are

$$
\begin{aligned}
\mu_t &= 0.90539 + 0.24308 R_{t-1} \\
\tau_t^2 &= 0.63235 + 0.66434(R_{t-1} - \mu_{t-1})^2 + 0.77277\tau_{t-1}^2 \\
\delta_t &= -0.12744 - 0.04395 S_{t-1} \\
b_t &= 3.47490.
\end{aligned}
$$

3.3.3 Summary of our Results

Table 3.4 summarizes the various models in this section. SDM stands for state-dependent model based on Pearson IV type distribution. The state-dependent model without AR has better AIC than AR-ARCH model. It will suggest that the variance is explained by some fundamental factors like skewness. The state-dependent model with error term is better than the state-dependent model with AR, but not better than the model with AR-GARCH. Furthermore, the state-dependent model with error term and τ in the formulation of AR-GARCH is better than the AR-GARCH. Consequently, the state-dependent model with error term, τ, b constant and δ a linear function of S_{t-1}, is the best AIC model among our possible models.

Table 3.4: The summary of results

	Model	AIC
(1)	SDM with error and τ(b:const,δ:linear)	2022.7
(2)	SDM with error and τ(b:const,δ:quadratic)	2023.6
(3)	SDM with error and τ(b:quadratic,δ:quadratic)	2024.8
(4)	SDM with error and τ(b:const,δ:const)	2028.6
(5)	SDM with error and τ(b:quadratic,δ:const)	2029.9
(6)	SDM with error and τ(b:const,$\delta = 0$)	2040.6
(7)	AR-GARCH(2,1)(skewness included)	2069.3
(8)	AR-GARCH(1,1)	2072.0
(9)	AR-GARCH(2,1)	2073.7
(10)	SDM with error(b:quadratic,δ:linear)	2086.4
(11)	SDM with error(b:quadratic,δ:quadratic)	2087.8
(12)	SDM with error(b:const,δ:const)	2098.8
(13)	SDM with AR(1)	2118.0
(14)	SDM without AR	2143.0
(15)	AR-ARCH(2,1)(skewness included)	2167.8
(16)	AR-ARCH(1,1)	2170.5
(17)	AR-ARCH(2,1)	2172.3
(18)	AR(skewness included)	2258.1

3.3.4 Examples

Although Model 5 has better AIC than Model 4, Model 4 is more interpretable by some financial factors (such as asymmetric phenomena). Therefore, we adopt three examples by using Model 4.

The first example is the crash of October 17-24, 1987, so-called 'Black Monday'. The previous cross-sectional skewness is close to zero (-0.204845). No specific industries are preferred. The parameters of the state-dependent model are $\mu_t = 1.18018$, $\tau_t^2 = 5.68427$, $\delta_t = -0.12597$, and $b_t = 1.75834$. Figure 3.4 gives the distributions of R_t obtained by the state-dependent model and AR(1) model with S_{t-1} respectively. The distribution of the state-dependent model has a heavier tail than normal distribution. The +sign in the figure denotes the actual return (-9.03956) of R_t. The four dots (0.24891, 2.71310, 0.97135, 1.30221) indicate the returns of four previous weeks. No information is drawn from these four numbers. The expectation and variance of the state-dependent model are 0.48380 and 11.94010 respectively. On the other hand, those of AR(1) with S_{t-1} are 0.61560 and 5.018 respectively.

The second example is for an increase of the market price by a movement of several domestic demand-related industries (Real Estate, Electric Power, Rail and Bus) and financial sectors (Securities, Insurance) in March 8-15, 1986. The cross-sectional skewness of the previous week is 0.98630, implying a heavy-tail to the right side. The parameters of the state-dependent model are $\mu_t = 1.40406$, $\tau_t^2 = 5.68427$, $\delta_t = -0.18803$, $b_t = 2.23294$. Figure 3.5 gives the distributions of the state-dependent model and the AR(1) model with S_{t-1} respectively. The +sign denotes the actual return (3.44620). The four dots (2.02946, 1.18535, 1.89667, 2.28671) indicate the returns of four previous weeks. When the skewness is positive, due to investors' preference over several specific industries, the positive returns tend to continue. But the range of returns is not so wide. This actual return (3.4462) is one of the biggest returns. Therefore, it is reasonable that the density obtained by the state-dependent model has heavier tail than normal density. The expectation and variance of the state-dependent model are 0.59217 and 4.32740 respectively. On the other hand, those of AR(1) with skewness variable are 0.63598 and 5.018 respectively.

The third example is on the increase of market prices for June 23-30, 1984 by the movement of many industry returns. The previous cross-sectional skewness is -2.54400. The parameters of state-dependent model are $\mu_t = 0.60939$, $\tau_t^2 = 5.68427$, $\delta_t = -0.00410$, $b_t = 1.94288$. Figure 3.6 shows the distributions of state-dependent model and AR(1) model with skewness variable respectively. The +sign indicates actual return (0.78510). The four dots (-0.28890, 1.69412, -0.75399, -1.20773) indicate returns of four previous weeks. Contrary to the second example, in this example it does not continue to get positive returns. And the variance is larger than the case of the second example. The expectation and variance of state-dependent model are 0.58925 and 6.41785 respectively. On the other hand, those of AR(1) with skewness variable are 0.43285 and 5.018 respectively.

Finally, Figure 3.7 shows the time series of expected returns and actual returns of Model 5. The solid line and dotted line indicate expected returns and actual returns respectively.

3.3.5 Examples of the Out-of-sample Case

Examples in the previous subsection are based on the in-sample analysis. In this subsection, two typical examples based on the out-of-sample analysis are shown. The first example is the crash of October 17-24, 1987, so-called 'Black Monday'. To begin with, we estimate the parameters of Model 4 by using the time series data from 1983, January. to the previous week of October 17-24, 1987, with 242 observations. The estimated parameters are

$$\mu_t = 0.63095 + 0.22368R_{t-1}, \quad \tau_t = 2.40117,$$
$$\delta_t = -0.03807 - 0.02630S_{t-1},$$
$$b_t = 2.47479 + 0.43378S_{t-1} + 0.17112S_{t-1}^2.$$

Next, compared with the parameters estimated by the in-sample analysis, the absolute value of the constant term of δ (-0.03807) is smaller than that of the in-sample case (-0.13664). The constant term of b (2.47479) is bigger than that of the in-sample case (1.181264). The value of τ is almost the same. Finally, the parameters of the predicted distribution by using this out-of-sample analysis are $\mu_t = 0.9222$, $\tau_t^2 = 5.765617$, $\delta_t = -0.03268$, $b_t = 2.39311$. The δ is bigger than that of the in-sample case. It indicates that the degree of skewness of the out-sample case is smaller than that of the in-sample case. Nevertheless, the tendency of left-skewed distribution is the same. The expectation and variance are 0.78742 and 3.23800 respectively.

The second example is the price down from March, 30 to April, 6, 1990. The previous cross-sectional skewness and return are -0.220891 and 0.2113191 respectively. To begin with, we estimate the parameters of Model 4 by using the time series data from 1983, January, to the previous week of March 30, 1990, with 367 observations. The estimated parameters are

$$\mu_t = 1.09857 + 0.16559R_{t-1}, \quad \tau_t = 2.33814,$$
$$\delta_t = -0.17056 - 0.04292S_{t-1},$$
$$b_t = 2.33637 + 0.31985S_{t-1} + 0.11179S_{t-1}^2.$$

Next, compared with the parameters estimated by the in-sample analysis, the absolute value of the constant term of δ (-0.17056) is bigger than that of the in-sample case (-0.13664). The constant term of b (2.33637) is bigger than that of the in-sample case (1.181264). The value of τ is almost the same. Finally, the parameters of the predicted distribution by using this out-of-sample analysis are $\mu_t = 1.13356$, $\tau_t^2 = 5.46690$, $\delta_t = -0.16108$, $b_t = 2.27117$. The parameters of the predicted distribution by using the in-sample analysis are $\mu_t = 0.93210$, $\tau_t^2 = 5.68427$, $\delta_t = -0.12513$, $b_t = 1.75456$.

The δ is smaller than that of the in-sample case. It indicates that the degree of skewness of the out-sample case is bigger than that of the in-sample case. The tendency of left-skewed distribution is the same. The expectation and variance of the out-of-sample analysis are 0.47510 and 3.52743 respectively. And the expectation and variance of the in-sample analysis are 0.23840 and 12.11010 respectively.

Figure 3.4: Distribution of example 1

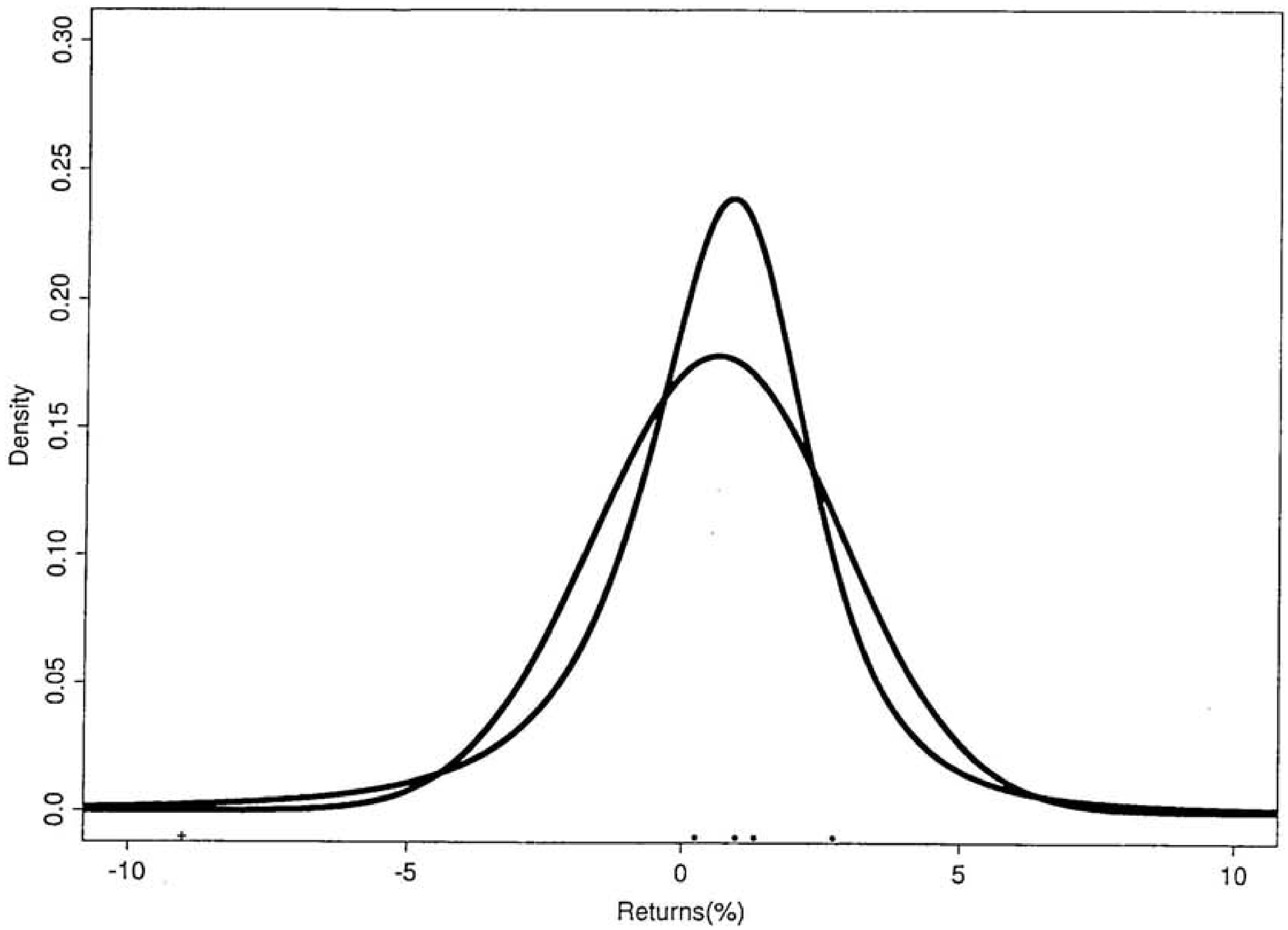


Figure 3.5: Distribution of example 2

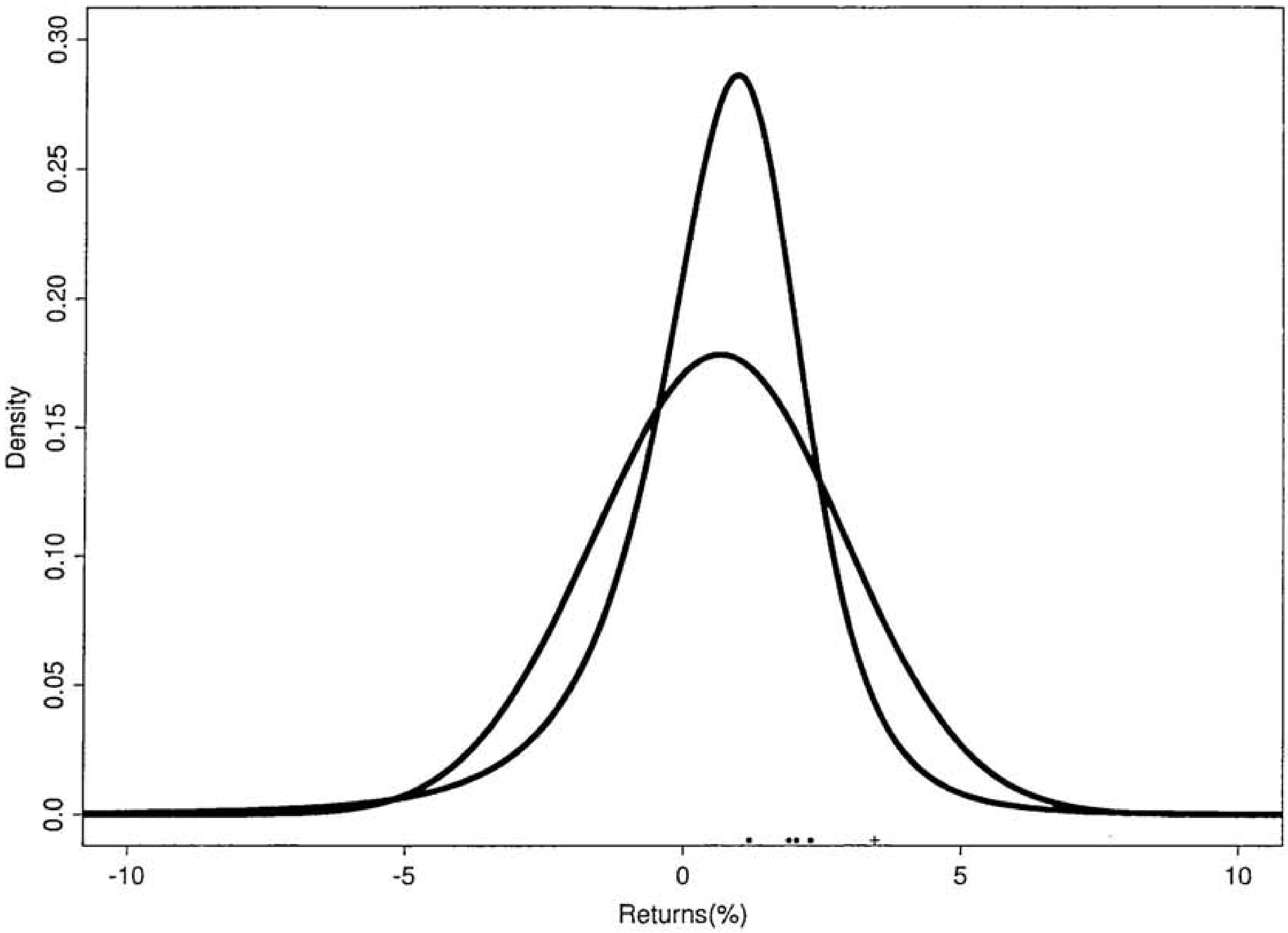


Figure 3.6: Distribution of example 3

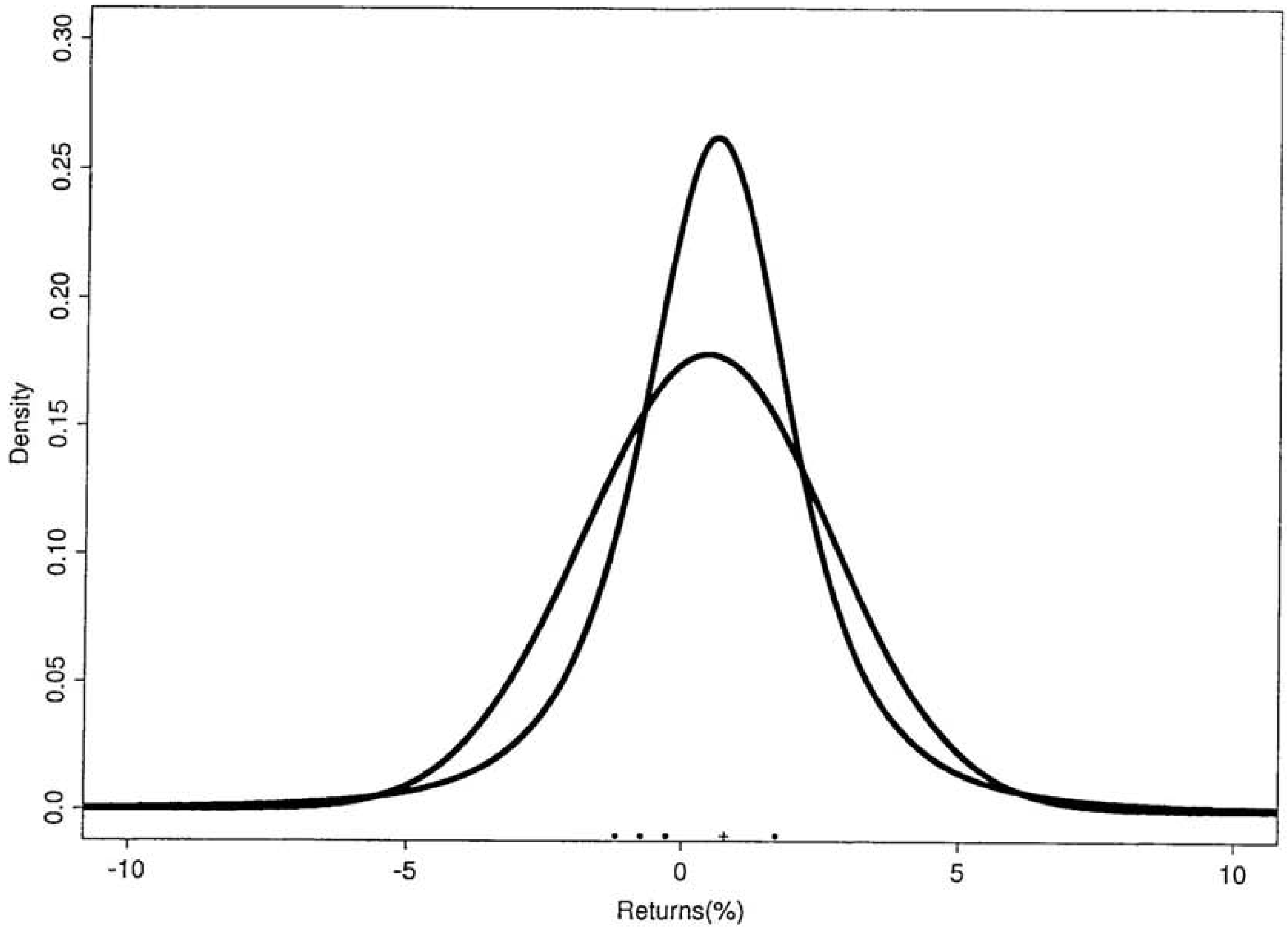
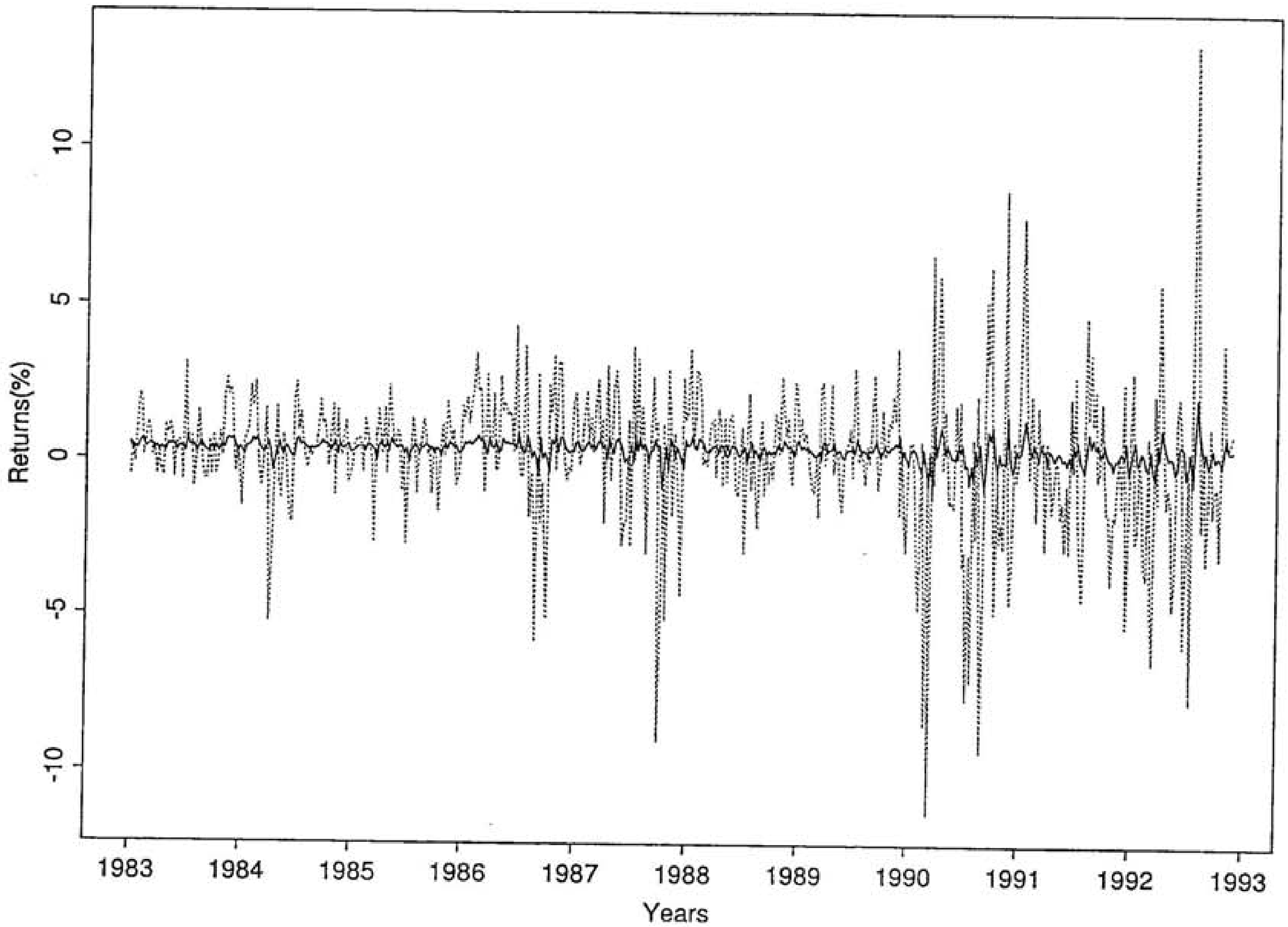


Figure 3.7: The time series of expected returns and actual returns



Chapter 4

Non-Gaussian Distribution for Stock Returns and Related Stochastic Differential Equation

4.1 Introduction

Measuring the market risk of financial assets such as stocks and bonds is very important problem for financial institutions and regulators. In this area, the normal distribution is commonly used because of its simplicity. But, for the daily returns of stocks, it is well-known that they are actually not distributed as normal distribution. Many researches reported that heavy-tailed and skewed distributions are observed (Kariya, et al. (1995), Nagahara, (1995a)).

Fama (1965) and Mandelbrot (1966) suggested the use of such a heavy-tailed distribution as stable Paretian distribution as a model for daily returns. Blattberg and Gonedes (1974) compared the stable and t-distribution for stock prices of United States. Kariya, et al. (1995) extensively studied on this feature of the Japanese market (see also Kariya (1993)) and showed that the distribution of market returns is not only leptokurtic but also skewed. In our time series analysis below, to take into account of this feature, we use the Pearson type IV family of distributions which is a noncentral version of the type VII distributions and forms a broad class of distributions including t-distribution and Cauchy distribution (Pearson, K. (1914), Pearson, E.S. and Hartrey (1954), Pearson, E.S. (1963), Johnson and Kotz (1970), Nagahara (1994, 1995a)).

We also consider a stochastic differential equation whose stationary distribution is the type VII or IV of Pearson system (Wong, 1963). This reveals the relationship between the stationary distribution of stock returns and related stochastic differential equation.

Wong pointed out the transition probability density function of stochastic differential equation for the Type VII Pearson distribution. It can be written more explicitly for t-distribution which have even degrees of freedom. In this chapter, we derive the transition probability density function for a more general family of distributions, and present a method for estimating the parameters of its stochastic differential equation. Furthermore, in order to estimate the parameters of stochastic differential equation which corresponds to the Pearson IV type (asymmetric distribution and continuous shape parameter), we consider the local linearization method (Ozaki (1985a, 1985b, 1989, 1992a, 1992b, 1993), Biscay, Jimenez, Riera, and Valdes (1994), Shoji and Ozaki (1994), and Nagahara (1995b)).

In section 4.2, heavy-tailed and non-central distributions are introduced and the maximum likelihood method for parameter estimation is shown. In section 4.3, the stochastic differntial equations related to the stationary distribution in section 4.2 are introduced. The parameters are estimated by maximum likelihood method using the transition probability density function and the local linearization method. In section 4.4, the conclusion of the previous section is shown. In section 4.5, a comparison of parameter estimation methods for stochastic differential equation is shown.

4.2 Non-Gaussian Distribution for Stock Returns

4.2.1 The VII and IV Type of Pearson System

In this chapter, x_t denotes the daily returns of stock price index defined by

$$x_t = \log \frac{P_t}{P_{t-1}}, \tag{4.1}$$

where P_t is the closing price of t-th day. We shall use the type IV Pearson family of probability distributions as the distribution of the daily stock returns. The type IV Pearson distribution can discribe heavy-tailed and skewed distribution. It is easy to compare with symmetric distributions such as t-distribution and the type VII Pearson distribution. Furtheremore, since the type IV distribution belongs to Pearson System, it is possible to introduce a stochastic differential equation which generates the distribution as the stationary distribution (Wong (1963), Ozaki (1985a)).

The density function $p(x)$ of the type IV distribution is defined as the solution to the differential equation

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta)}{(\frac{x-\mu}{\tau})^2 + 1}, \tag{4.2}$$

where μ, τ and δ are the location parameter, the scale parameter and the non-centrality parameter, respectively. The solution is explicitly given by

$$p(x|\mu, \tau, \delta, b) = \frac{C \exp\{2b\delta \arctan(\frac{x-\mu}{\tau})\}}{\{(x-\mu)^2 + \tau^2\}^b}, \tag{4.3}$$

where $b > \frac{1}{2}$, $\tau > 0$, $-\infty < \delta < \infty$. Note that, if $\delta = 0$, then the solution defines the type VII distributions of Pearson system.

In the past research, this distribution has not been introduced in this way, but used as an approximation to non-central t distribution by using moment ratio (Merrington and Pearson (1958)). As far as the author knows, this type IV family has not been used for actual data analysis because of its difficulty in computation. In particular, it is necessary for empirical use to develop a practical method for calculating normalizing constant, C. In order to compute the normalizing constant, analytic expressions of the normalizing constant are obtained by using recursive formula for $b = n$ and $b = n + 1/2$, $n = 1, 2, 3, \ldots$ (Chapter 2). In this section, we model the distribution of x_t by using (4.3) as the probability density function. Namely, x_t are assumed to be independently and identically distributed according to the density function (4.3).

Given T observations $x_1, \ldots, x_T$, the log-likelihood of this model is given by

$$\begin{aligned} l(\mu, \tau, \delta, b) &= \sum_{t=1}^{T} \log p(x_t) \qquad (4.4) \\ &= \sum_{t=1}^{T} \log C + 2b\delta \sum_{t=1}^{T} \arctan\left(\frac{x_t - \mu}{\tau}\right) - b \sum_{t=1}^{T} \log\{(x_t - \mu)^2 + \tau^2\}. \end{aligned}$$

We use DE (double exponential formula) to calculate the normalizing constant, C, by numerical integration for any b (Mori (1987)). The accuracy of numerical integration was checked by comparing with the analytic solution for some specific values of b (Chapter 2). To estimate the parameters and to identify the models, the procedure based on quasi-Newton method is used (Kitagawa (1993)). Model is selected by AIC (Akaike Information Criterion) defined by

$$\text{AIC} = -2(\text{maximum log-likelihood}) + 2(\text{number of free parameters}). \tag{4.5}$$

4.2.2 Results

The daily rate of returns were obtained from the daily index of Nikkei 225 index, TOPIX index and Standard and Poor's 500 index. The data of Japanese market consists of returns from January 5, 1970 to December 30, 1994 with 6912 observations. The data of U.S. market consists of returns from January 5, 1975 to December 30, 1994 with 5053 observations.

The results are shown in Table 4.1 - Table 4.6. The $*$ indicates the minimum values of AIC. LL stands for log-likelihood. According to Table 4.1, in Japanese market, the estimated parameter b over the entire period is around 1.75. It suggests that the distribution of the returns has heavier tail than the normal distribution. The AIC values of asymmetric distributions (Pearson IV type) is smaller than that of symmetric distributions (Pearson VII type). Table 4.2 - Table 4.6 show the results of fitting the models to subintervals of five year span. Whether asymmetric distribution is selected or not, depends on its period. From 1970 to 1974, both of the AIC values of asymmetric distribution are much smaller than symmetric ones. In other periods, there are smaller differences. As for TOPIX, according to the AIC criterion, the asymmetric distribution is considered to be better than the symmetric one for the first two periods, 1970-1974 and 1975-1979. Then the symmetric distribution is selected for the rest three periods 1980-1984, 1985-1989, 1990-1994. As for NK225, the asymmetric distribution is selected for the first four intervals, 1970-1974, 1975-1979, 1980-1984, 1985-1989. However, the symmetric distribution is selected in the last interval, 1990-1994.

According to Table 4.1, in U.S. market, the estimated parameter b is around 2.70. It also suggests that the distribution of the returns has heavier tail than the normal distribution, but is not so heavier than the one for the Japanese market. The AIC values of symmetric distribution (Pearson VII type) become smaller than that of asymmetric distribution (Pearson IV type). Only at the interval, 1980-1984, the AIC value of asymmetric one becomes smaller than the symmetric one, but the difference is not so much. From these results, for long term such as 5 years, it is reasonable that the stationary distributions of U.S. market are symmetric, especially from 1975 to 1994.

Figure 4.1 shows the density function of real returns of Nikkei 225 (a broken line) and the density function estimated by the maximum likelihood method (a solid line). Figure 4.2 shows the density function of real returns of SP500 (a broken line) and the density function estimated by the maximum likelihood method (a solid line). In both cases, the estimated distributions are well fitted with actual distributions. As the characteristic, each distribution is leptokurtic. And there is a little skewness for Nikkei 225.

Table 4.1: Comparison of symmetric and asymmetric estimated distributions (Japan 1970-94, U.S. 1975-1994)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	0.00064	0.00885	0.00	1.78268	23235.95	-46465.90
	(asymmetric)	0.00152	0.00881	-0.06384	1.78297	23246.50	-46485.00*
TOPIX	(symmetric)	0.00053	0.00745	0.00	1.75860	24305.54	-48605.08
	(asymmetric)	0.00105	0.00743	-0.04468	1.75832	24310.99	-48613.98*
SP500	(symmetric)	0.00041	0.01390	0.00	2.69815	16970.88	-33935.76*
	(asymmetric)	0.00033	0.01391	0.00410	2.69998	16970.91	-33933.82

Table 4.2: Comparison of symmetric and asymmetric estimated distributions (1970-1974)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	0.00096	0.01102	0.00	1.97605	4787.22	-9568.44
	(asymmetric)	0.00291	0.01096	-0.11871	1.98996	4793.90	-9579.79*
TOPIX	(symmetric)	0.00092	0.00805	0.00	1.85563	5135.51	-10265.02
	(asymmetric)	0.00257	0.00797	-0.13534	1.86905	5144.96	-10281.92*

Table 4.3: Comparison of symmetric and asymmetric estimated distributions (1975-1979)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	0.00048	0.01033	0.00	3.10319	5389.06	-10772.13
	(asymmetric)	0.00144	0.01017	-0.07050	3.06240	5390.70	-10773.40*
TOPIX	(symmetric)	0.00041	0.00954	0.00	3.64864	5688.67	-11371.35
	(asymmetric)	0.00118	0.00946	-0.06361	3.62304	5689.78	-11371.56*
SP500	(symmetric)	0.00030	0.02405	0.00	6.56703	4382.97	-8759.95*
	(asymmetric)	-0.00110	0.02454	0.04990	6.79835	4383.29	-8758.58

Table 4.4: Comparison of symmetric and asymmetric estimated distributions (1980-1984)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	0.00054	0.00918	0.00	2.59165	5332.04	-10658.08
	(asymmetric)	0.00149	0.00913	-0.07508	2.59013	5334.09	-10660.17*
TOPIX	(symmetric)	0.00052	0.00791	0.00	2.49152	5490.92	-10975.85*
	(asymmetric)	0.00070	0.00791	-0.01615	2.49293	5491.03	-10974.06
SP500	(symmetric)	0.00019	0.02278	0.00	4.40873	4120.55	-8235.09
	(asymmetric)	-0.00328	0.02292	0.12381	4.51945	4123.37	-8238.74*

Table 4.5: Comparison of symmetric and asymmetric estimated distributions (1985-1989)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	0.00120	0.00929	0.00	2.02186	4713.14	-9420.27
	(asymmetric)	0.00203	0.00933	-0.05990	2.03760	4714.65	-9421.31*
TOPIX	(symmetric)	0.00101	0.00906	0.00	1.93043	4672.03	-9338.05*
	(asymmetric)	0.00130	0.00907	-0.02115	1.93354	4672.23	-9336.46
SP500	(symmetric)	0.00101	0.01038	0.00	1.86298	4102.91	-8199.82*
	(asymmetric)	0.00145	0.01043	-0.02782	1.87138	4103.27	-8198.55

Table 4.6: Comparison of symmetric and asymmetric estimated distributions (1990-1994)

Index		μ	τ	δ	b	LL	AIC
NK225	(symmetric)	-0.00065	0.02137	0.00	2.34822	3441.07	-6876.14*
	(asymmetric)	-0.00071	0.02138	0.00179	2.34926	3441.07	-6874.14
TOPIX	(symmetric)	-0.00070	0.01440	0.00	1.96979	3691.40	-7376.80*
	(asymmetric)	-0.00120	0.01440	0.02347	1.97130	3691.66	-7375.32
SP500	(symmetric)	0.00024	0.01178	0.00	2.64030	4429.81	-8853.62*
	(asymmetric)	0.00055	0.01180	-0.01874	2.64502	4429.93	-8851.86

Figure 4.1: The density function of Nikkei225

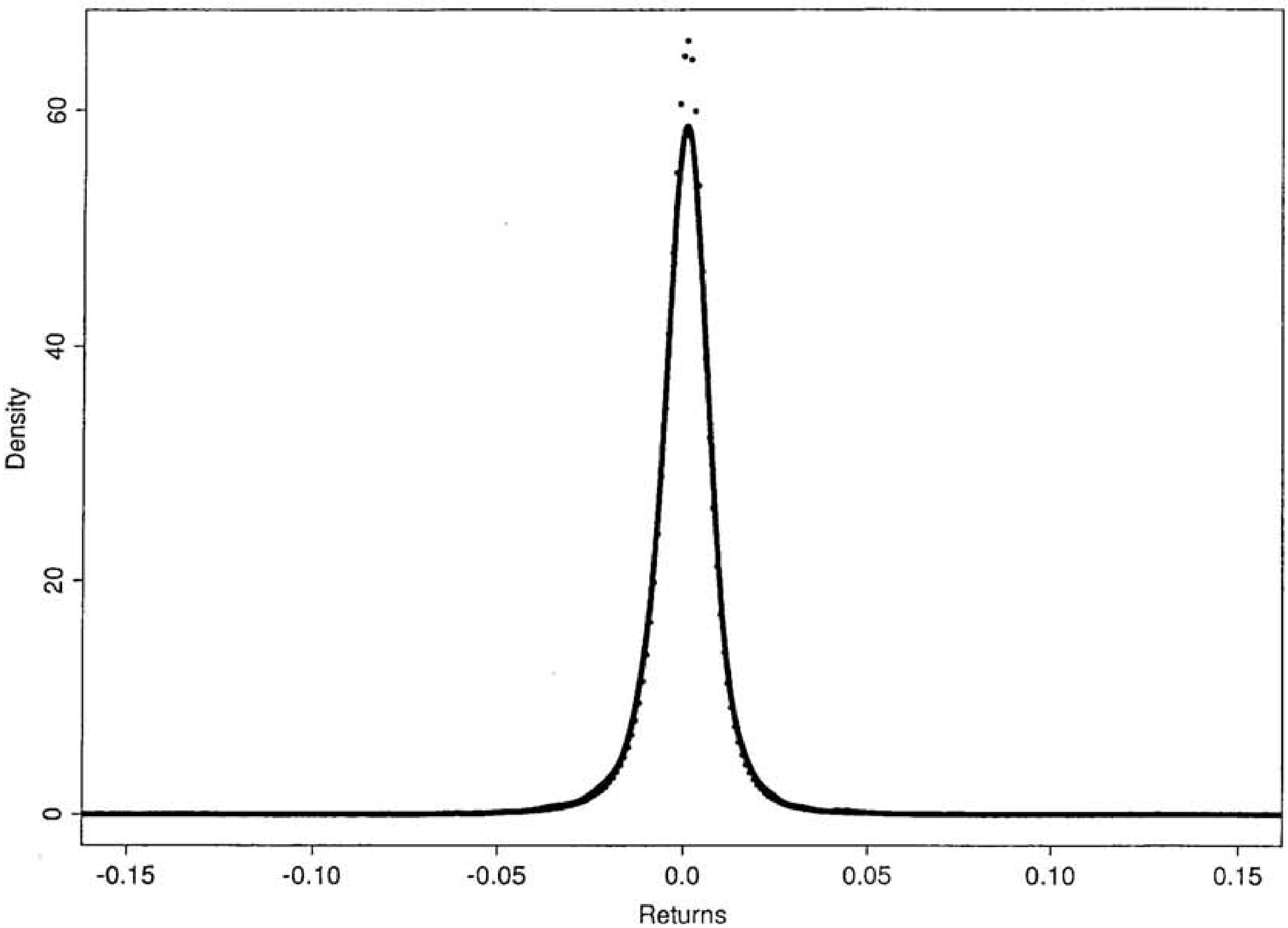
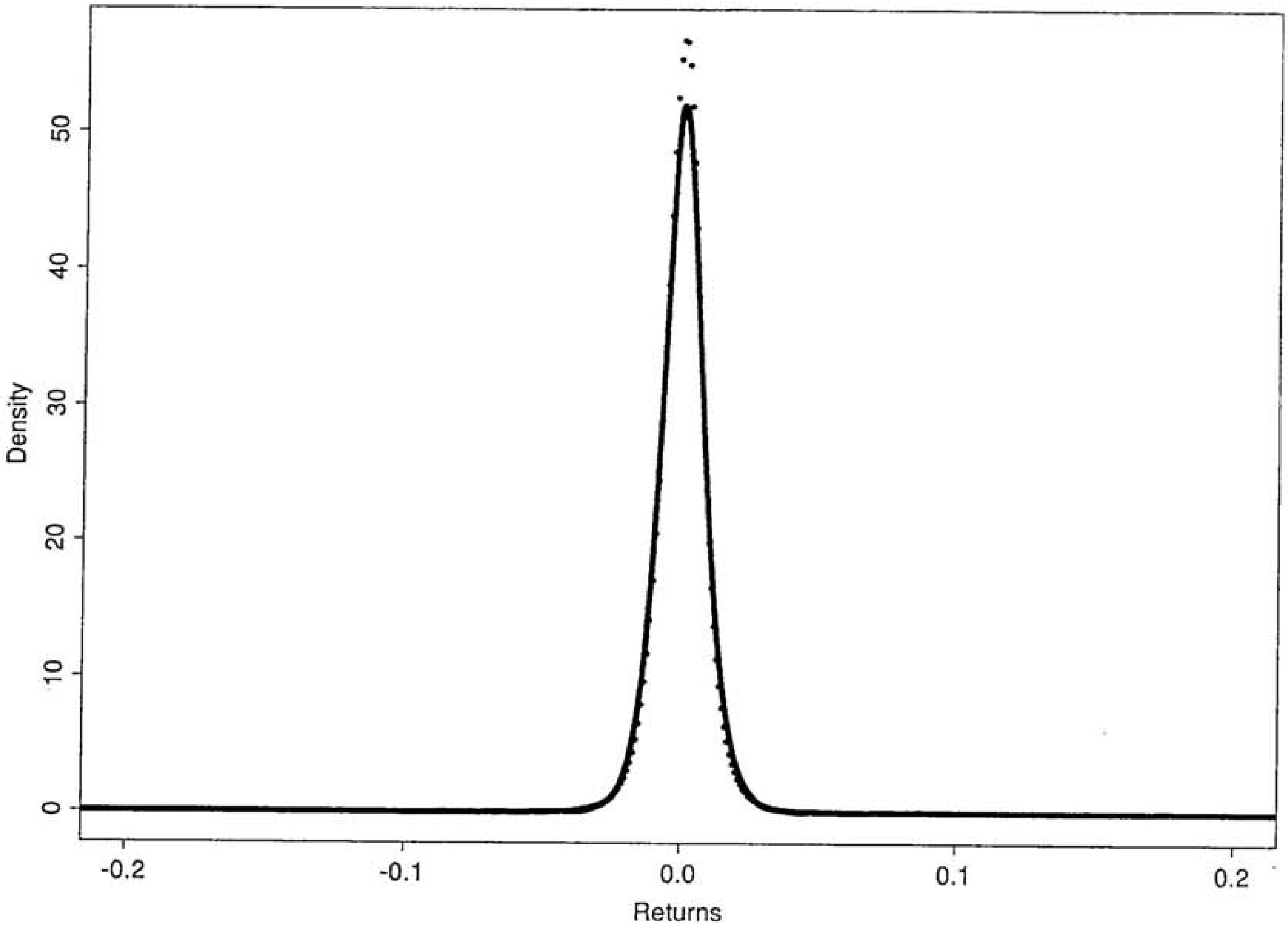


Figure 4.2: The density function of SP500



4.3 Related Stochastic Differential Equation Model

4.3.1 Stochastic Differential Equation which has Non-central Stationary Distribution

In the previous section, we consider the type VII and IV of Pearson system to represent skewed and heavy-tailed distribution. According to Wong (1963), there is a relationship between the Pearson system and Markov diffusion process. He showed that for any distribution $p(x)$ which belongs to the Pearson system, defined by the following equation,

$$-\frac{dp(x)}{dx} = \frac{c_0 + c_1 x}{d_0 + d_1 x + d_2 x^2} p(x), \tag{4.6}$$

it is possible to derive a diffusion process whose marginal distribution is $p(x)$.

According to Ozaki (1985a), it can be extended to any distribution $p(x)$ defined by proper analytic functions $c(x)$ and $d(x)$ in a distribution system given by,

$$\frac{dp(x)}{dx} = \frac{c(x)}{d(x)} p(x). \tag{4.7}$$

He also showed that the corresponding Markov diffusion process is given, with $c(x)$ and $d(x)$, by the following Fokker-Planck equation for the transition density function $q(x|x_0, t)$,

$$\frac{\partial q(x|x_0,t)}{\partial t} = -\frac{\partial}{\partial x}[\{c(x) + d'(x)\}q(x|x_0,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[2d(x)q(x|x_0,t)]. \tag{4.8}$$

According to Mortensen (1979), the diffusion process $x(t)$ defined by (4.8) has a stochastic differential equation representation,

$$\dot{x} = a(x) + \sqrt{b(x)}n(t), \tag{4.9}$$

where

$$a(x) = c(x) + d'(x), \tag{4.10}$$

$$b(x) = 2d(x), \tag{4.11}$$

in the Ito form of stochastic calculus.

In this chapter, we adopted the type VII and IV of Pearson system to express the symmetric and asymmetric heavy-tailed distibution of returns, respectively. Correspondingly, we consider the diffusion process defined by

$$-\frac{p'}{p} = \frac{\frac{2b}{\tau}(\frac{x-\mu}{\tau} - \delta) \times \frac{\sigma^2}{2}}{\{(\frac{x-\mu}{\tau})^2 + 1\} \times \frac{\sigma^2}{2}} = \frac{2b(x - \mu - \tau\delta) \times \frac{\sigma^2}{2}}{\{(x-\mu)^2 + \tau^2\} \times \frac{\sigma^2}{2}}. \tag{4.12}$$

In this case, the diffusion process is given by

$$dx = \sigma^2\{(1-b)(x-\mu) + b\tau\delta\}dt + \sigma\sqrt{(x-\mu)^2+\tau^2}dw. \tag{4.13}$$

This diffusion process is reasonable because the time series data have the first order positive autocorrelation (AR term).

Furthermore, with the variable transformation for the equation of (4.9),

$$y = \int_{-\infty}^{x} \frac{1}{\sqrt{b(\xi)}} d\xi, \tag{4.14}$$

we can replace by a dynamical system driven by Gaussian white noise $n(t)$ as follows:

$$x = h(y), \tag{4.15}$$

$$\dot{y} = f(y) + n(t), \tag{4.16}$$

where $h(y)$ is a smooth function given by the inverse function of (4.14) and

$$f(y) = \frac{a(x)}{\sqrt{b(x)}} - \frac{1}{4}\frac{b'(x)}{\sqrt{b(x)}}. \tag{4.17}$$

We can consider the various diffusion processes which have stationary distribution like Pearson type (Ozaki 1985a, 1992b). We apply the same method to our case, and obtain the following transformation.

Specifically for the following diffusion process

$$dx = \sigma^2\{(1-b)(x-\mu) + b\tau\delta\}dt + \sigma\tau\sqrt{(\frac{x-\mu}{\tau})^2+1^2}dw, \tag{4.18}$$

the transformation (4.14) is given by

$$\begin{aligned} y &= \int_{-\infty}^{x} \frac{d\xi}{\sigma\sqrt{(\frac{\xi-\mu}{\tau})^2+1^2}} && (4.19)\\ &= \frac{\tau}{\sigma}\log\{\frac{(x-\mu)}{\tau} + \sqrt{(\frac{x-\mu}{\tau})^2+1^2}\} && (4.20)\\ &= \frac{\tau}{\sigma}\text{arcsinh}\left(\frac{x-\mu}{\tau}\right). && (4.21)\end{aligned}$$

Therefore, x is given by:

$$x = \mu + \tau\sinh\left(\frac{\sigma y}{\tau}\right), \tag{4.22}$$

and we define the $f(y)$ as follows,

$$f(y) = \tau\sigma\left(\frac{1}{2} - b\right)\tanh\left(\frac{\sigma y}{\tau}\right) + \frac{\sigma b\tau\delta}{\cosh(\frac{\sigma y}{\tau})}, \tag{4.23}$$

we obtain the transformation of (4.18) defined by

$$dy = f(y)dt + \tau dw. \tag{4.24}$$

We apply the local linearization method to this formula in subsection 4.3.4.

4.3.2 The Transition Probability Density Function and Estimation of the Parameters

In this subsection, we obtain the transition probability density function of the symmetric case in order to define the likelihood for estimating the parameters of the stochastic differential equation. In the case that $\delta = 0$, Wong (1963) studied a symmetric version of the diffusion process given by

$$dx = \{(1-2\alpha)x\}dt + \sqrt{2(x^2+1)}dw. \tag{4.25}$$

Wong pointed out the transition probability density function of this stochastic differential equation is given by

$$\begin{aligned} p(x|x_0,t) &= (1+x^2)^{(\alpha+1/2)}\left\{\frac{1}{\pi}\sum_{n=0}^{N}\frac{(\alpha-n)}{n!\Gamma(2\alpha+1-n)}e^{-n(2\alpha-n)t}\theta_n(x_0)\theta_n(x)\right. \\ &\left.+\frac{1}{2\pi}\int_0^\infty e^{-(\alpha^2+\mu^2)t}[\psi(\mu,x_0)\psi(-\mu,x)+\psi(-\mu,x_0)\psi(\mu,x)]d\mu\right\}, \end{aligned} \tag{4.26}$$

where

$$\theta_n(x) = 2^{\alpha-n}\Gamma(\alpha-n+\frac{1}{2})(-1)^n(1+x^2)^{\alpha+1/2}\frac{d^n}{dx^n}[(1+x^2)^{n-\alpha-1/2}], \tag{4.27}$$

are polynomials of degree n, and $\psi(\mu,x)$ is given by

$$\psi(\mu,x) = (x+\sqrt{1+x^2})^{i\mu}(1+x^2)^{1/2}{}_2F_1\left(-\alpha,\alpha+1;1+i\mu;\frac{1}{2}+\frac{1}{2}\frac{x}{\sqrt{1+x^2}}\right), \tag{4.28}$$

with ${}_2F_1$ being the Gauss hypergeometric series (Abramowits and Stegun(1965)).

For α equals to a positive integer, K, it is possible to get more explicit form. These are the cases when the stationary distributions belong to t-distributions which have even degrees of freedom and we calculate the transition probability density function of our stochastic differential equation. The transition probability density functions for $\alpha = K, (b = 1/2 + K, K = 1, 2, 3, ...)$ are given by,

$$\begin{aligned} p(x|x_0, t) &= \frac{1}{\tau\{(\frac{x-\mu}{\tau})^2 + 1\}^{\frac{1}{2}+K}} \left[\left\{ (\frac{x-\mu}{\tau})^2 + 1 \right\}^{\frac{K}{2}} \left\{ (\frac{x_0-\mu}{\tau})^2 + 1 \right\}^{\frac{K}{2}} \right. \\ &\quad \left. \times \frac{1}{2\sqrt{\pi \times \frac{\sigma^2}{2} t}} e^{-K^2 \times \frac{\sigma^2}{2} t} e^{-u^2} + \frac{1}{\sqrt{\pi}} \sum_{i=1}^{K} \{ a_{i,K}(x, x_0) \int_{u-i\sqrt{\frac{\sigma^2}{2}t}}^{u+i\sqrt{\frac{\sigma^2}{2}t}} e^{-z^2} dz \} \right], \end{aligned}$$

where

$$u = \frac{\operatorname{arcsinh}(\frac{x-\mu}{\tau}) - \operatorname{arcsinh}(\frac{x_0-\mu}{\tau})}{2\sqrt{\frac{\sigma^2}{2} t}}. \tag{4.29}$$

The analytic solution of the parameters, $a_{i,K}(x, x_0)$, are defined as follows:

(1)$K = 1$

$$a_{1,1}(x, x_0) = \frac{1}{2} \tag{4.30}$$

(2)$K = 2$

$$a_{1,2}(x, x_0) = \frac{3}{2} \left(\frac{x-\mu}{\tau} \right) \left(\frac{x_0-\mu}{\tau} \right) e^{-3 \times \frac{\sigma^2}{2} t} \tag{4.31}$$

$$a_{2,2}(x, x_0) = \frac{3}{4} = \frac{3 \times 1}{2^2} \tag{4.32}$$

(3)$K = 3$

$$a_{1,3}(x, x_0) = \frac{3}{16} \left\{ 4 \left(\frac{x-\mu}{\tau} \right)^2 - 1 \right\} \left\{ 4 \left(\frac{x_0-\mu}{\tau} \right)^2 - 1 \right\} e^{-8 \times \frac{\sigma^2}{2} t} \tag{4.33}$$

$$a_{2,3}(x, x_0) = \frac{15}{4} \left(\frac{x-\mu}{\tau} \right) \left(\frac{x_0-\mu}{\tau} \right) e^{-5 \times \frac{\sigma^2}{2} t} \tag{4.34}$$

$$a_{3,3}(x, x_0) = \frac{15}{16} = \frac{5 \times 3}{2^4} \tag{4.35}$$

(4) $K=4$

$$\begin{aligned}
a_{1,4}(x,x_0) &= \frac{5}{16}\left\{4\left(\frac{x-\mu}{\tau}\right)^3 - 3\left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad \times \left\{4\left(\frac{x_0-\mu}{\tau}\right)^3 - 3\left(\frac{x-\mu}{\tau}\right)\right\} e^{-15\times\frac{\sigma^2}{2}t} && (4.36) \\
a_{2,4}(x,x_0) &= \frac{5}{16}\left\{6\left(\frac{x-\mu}{\tau}\right)^2 - 1\right\}\left\{6\left(\frac{x_0-\mu}{\tau}\right)^2 - 1\right\} e^{-12\times\frac{\sigma^2}{2}t} && (4.37) \\
a_{3,4}(x,x_0) &= \frac{105}{16}\left(\frac{x-\mu}{\tau}\right)\left(\frac{x_0-\mu}{\tau}\right) e^{-7\times\frac{\sigma^2}{2}t} && (4.38) \\
a_{4,4}(x,x_0) &= \frac{35}{32} = \frac{7\times 5}{2^5} && (4.39)
\end{aligned}$$

(5) $K=5$

$$\begin{aligned}
a_{1,5}(x,x_0) &= \frac{15}{128}\left\{8\left(\frac{x-\mu}{\tau}\right)^4 - 12\left(\frac{x-\mu}{\tau}\right)^2 + 1\right\} \\
&\quad \times \left\{8\left(\frac{x_0-\mu}{\tau}\right)^4 - 12\left(\frac{x-\mu}{\tau}\right)^2 + 1\right\} e^{-24\times\frac{\sigma^2}{2}t} && (4.40) \\
a_{2,5}(x,x_0) &= \frac{105}{16}\left\{2\left(\frac{x-\mu}{\tau}\right)^3 - \left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad \times \left\{2\left(\frac{x_0-\mu}{\tau}\right)^3 - \left(\frac{x-\mu}{\tau}\right)\right\} e^{-21\times\frac{\sigma^2}{2}t} && (4.41) \\
a_{3,5}(x,x_0) &= \frac{105}{256}\left\{8\left(\frac{x-\mu}{\tau}\right)^2 - 1\right\}\left\{8\left(\frac{x_0-\mu}{\tau}\right)^2 - 1\right\} e^{-16\times\frac{\sigma^2}{2}t} && (4.42) \\
a_{4,5}(x,x_0) &= \frac{315}{32}\left(\frac{x-\mu}{\tau}\right)\left(\frac{x_0-\mu}{\tau}\right) e^{-9\times\frac{\sigma^2}{2}t} && (4.43) \\
a_{5,5}(x,x_0) &= \frac{315}{256} = \frac{9\times 7\times 5}{2^8} && (4.44)
\end{aligned}$$

(6) $K=6$

$$\begin{aligned}
a_{1,6}(x,x_0) &= \frac{21}{128}\left\{8\left(\frac{x-\mu}{\tau}\right)^5 - 20\left(\frac{x-\mu}{\tau}\right)^3 + 5\left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad \times \left\{8\left(\frac{x_0-\mu}{\tau}\right)^5 - 20\left(\frac{x-\mu}{\tau}\right)^3 + 5\left(\frac{x-\mu}{\tau}\right)\right\} e^{-35\times\frac{\sigma^2}{2}t} && (4.45)
\end{aligned}$$

$$
\begin{aligned}
a_{2,6}(x,x_0) &= \frac{105}{512}\left\{16\left(\frac{x-\mu}{\tau}\right)^4-16\left(\frac{x-\mu}{\tau}\right)^2+1\right\} \\
&\quad\times\left\{16\left(\frac{x_0-\mu}{\tau}\right)^4-16\left(\frac{x-\mu}{\tau}\right)^2+1\right\}e^{-32\times\frac{\sigma^2}{2}t} && (4.46) \\
a_{3,6}(x,x_0) &= \frac{315}{256}\left\{8\left(\frac{x-\mu}{\tau}\right)^3-3\left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad\times\left\{8\left(\frac{x_0-\mu}{\tau}\right)^3-3\left(\frac{x-\mu}{\tau}\right)\right\}e^{-27\times\frac{\sigma^2}{2}t} && (4.47) \\
a_{4,6}(x,x_0) &= \frac{63}{128}\left\{10\left(\frac{x-\mu}{\tau}\right)^2-1\right\}\left\{10\left(\frac{x_0-\mu}{\tau}\right)^2-1\right\}e^{-20\times\frac{\sigma^2}{2}t} && (4.48) \\
a_{5,6}(x,x_0) &= \frac{3465}{256}\left(\frac{x-\mu}{\tau}\right)\left(\frac{x_0-\mu}{\tau}\right)e^{-11\times\frac{\sigma^2}{2}t} && (4.49) \\
a_{6,6}(x,x_0) &= \frac{693}{512}=\frac{11\times 9\times 7}{2^9} && (4.50)
\end{aligned}
$$

$(7) K = 7$

$$
\begin{aligned}
a_{1,7}(x,x_0) &= \frac{7}{2048}\left\{64\left(\frac{x-\mu}{\tau}\right)^6-240\left(\frac{x-\mu}{\tau}\right)^4+120\left(\frac{x-\mu}{\tau}\right)^2-5\right\} \\
&\quad\times\left\{64\left(\frac{x_0-\mu}{\tau}\right)^6-240\left(\frac{x_0-\mu}{\tau}\right)^4+120\left(\frac{x_0-\mu}{\tau}\right)^2-5\right\}e^{-48\frac{\sigma^2}{2}t} \\
& && (4.51) \\
a_{2,7}(x,x_0) &= \frac{21}{512}\left\{48\left(\frac{x-\mu}{\tau}\right)^5-80\left(\frac{x-\mu}{\tau}\right)^3+15\left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad\times\left\{48\left(\frac{x_0-\mu}{\tau}\right)^5-80\left(\frac{x-\mu}{\tau}\right)^3+15\left(\frac{x-\mu}{\tau}\right)\right\}e^{-45\times\frac{\sigma^2}{2}t} && (4.52) \\
a_{3,7}(x,x_0) &= \frac{63}{2048}\left\{80\left(\frac{x-\mu}{\tau}\right)^4-60\left(\frac{x-\mu}{\tau}\right)^2+3\right\} \\
&\quad\times\left\{80\left(\frac{x_0-\mu}{\tau}\right)^4-60\left(\frac{x-\mu}{\tau}\right)^2+3\right\}e^{-40\times\frac{\sigma^2}{2}t} && (4.53) \\
a_{4,7}(x,x_0) &= \frac{231}{128}\left\{10\left(\frac{x-\mu}{\tau}\right)^3-3\left(\frac{x-\mu}{\tau}\right)\right\} \\
&\quad\times\left\{10\left(\frac{x_0-\mu}{\tau}\right)^3-3\left(\frac{x-\mu}{\tau}\right)\right\}e^{-33\times\frac{\sigma^2}{2}t} && (4.54) \\
a_{5,7}(x,x_0) &= \frac{1155}{2048}\left\{12\left(\frac{x-\mu}{\tau}\right)^2-1\right\}\left\{12\left(\frac{x_0-\mu}{\tau}\right)^2-1\right\}e^{-24\times\frac{\sigma^2}{2}t} && (4.55) \\
a_{6,7}(x,x_0) &= \frac{9009}{512}\left(\frac{x-\mu}{\tau}\right)\left(\frac{x_0-\mu}{\tau}\right)e^{-13\times\frac{\sigma^2}{2}t} && (4.56)
\end{aligned}
$$

$$a_{7,7}(x, x_0) = \frac{3003}{2048} = \frac{13 \times 11 \times 7 \times 3}{2^{11}} \tag{4.57}$$

4.3.3 Estimation of Parameters by Maximum Likelihood Method

We can obtain the exact log-likelihood by using the transition probability density function. Therefore, by maximizing this log-likelihood we can obtain the maximum likelihood estimates of the parameters of stochastic differential equation. The results are summarized as follows. In the tables the $*$ denotes the maximum of the log-likelihood.

The results for entire interval, i.e., 1970-1994 for Japan and 1975-1994 for United States, are shown in Table 4.7. Figure 4.3 and 4.4 show the time series data of Nikkei 225 and SP500, respectively. For Japan data, the minimum AIC values are achieved at $b = 1.5$. They are close to the results of the stationary distribution, $b = 1.75 - 1.78$ given in section 4.2.

For U.S. data, the minimum AIC value is achieved at $b = 2.5$. It is also close to the result of the stationary distribution, $b = 2.70$. The difference between Japan and U.S. is that the distribution of Japanese stock returns has heavier tail than that of U.S.. The AIC values of stochastic differential equation are smaller than those of the stationary distibution models.

According to Table 4.9 - Table 4.16 which show the parameters estimated by maximum likelihood method, in Japanese market, the minimum AIC values are achieved at $b = 3.5$ (Table 4.9), $b = 2.5$ (Table 4.10), $b = 2.5$ (NK225, Table 4.11), $b = 1.5$ (TOPIX, Table 4.11), $b = 2.5$ (Table 4.12). The number of b are close to the results of the stationary distributions. In U.S. market, the minimum AIC values are achieved at $b = 6.5$ (Table 4.9), $b = 4.5$ (Table 4.10), $b = 1.5$ (Table 4.11), $b = 2.5$ (Table 4.12). The number of b are also close to the results of the stationary distributions.

The summary of results in this subsection is given as follows.

First, the estimated parameters, μ, τ, and b, of stochastic differential equation are close to the parameters of the stationary distribution. Second, the distribution of Japanese stock returns is heavier tailed than that of the United States. Third, for long term such as 5 years, by the AIC values, it can be concluded that the stationary distributions are symmetric, especially from 1975 to 1994. Fourth, this stochastic differential equation and its maximum likelihood methods are very useful for describing the stock price behavior and for the risk management. Fifth, the AIC values of stochastic differential equation are consistently smaller than the stationary distribution except for SP500 of 1985-89, Nikkei

225 of 1990-94, SP500 of 1990-94. This clearly suggests the necessity of modeling the autocorrelation by adopting the stochastic differential equation.

Figure 4.3: The time series data of Nikkei225 returns

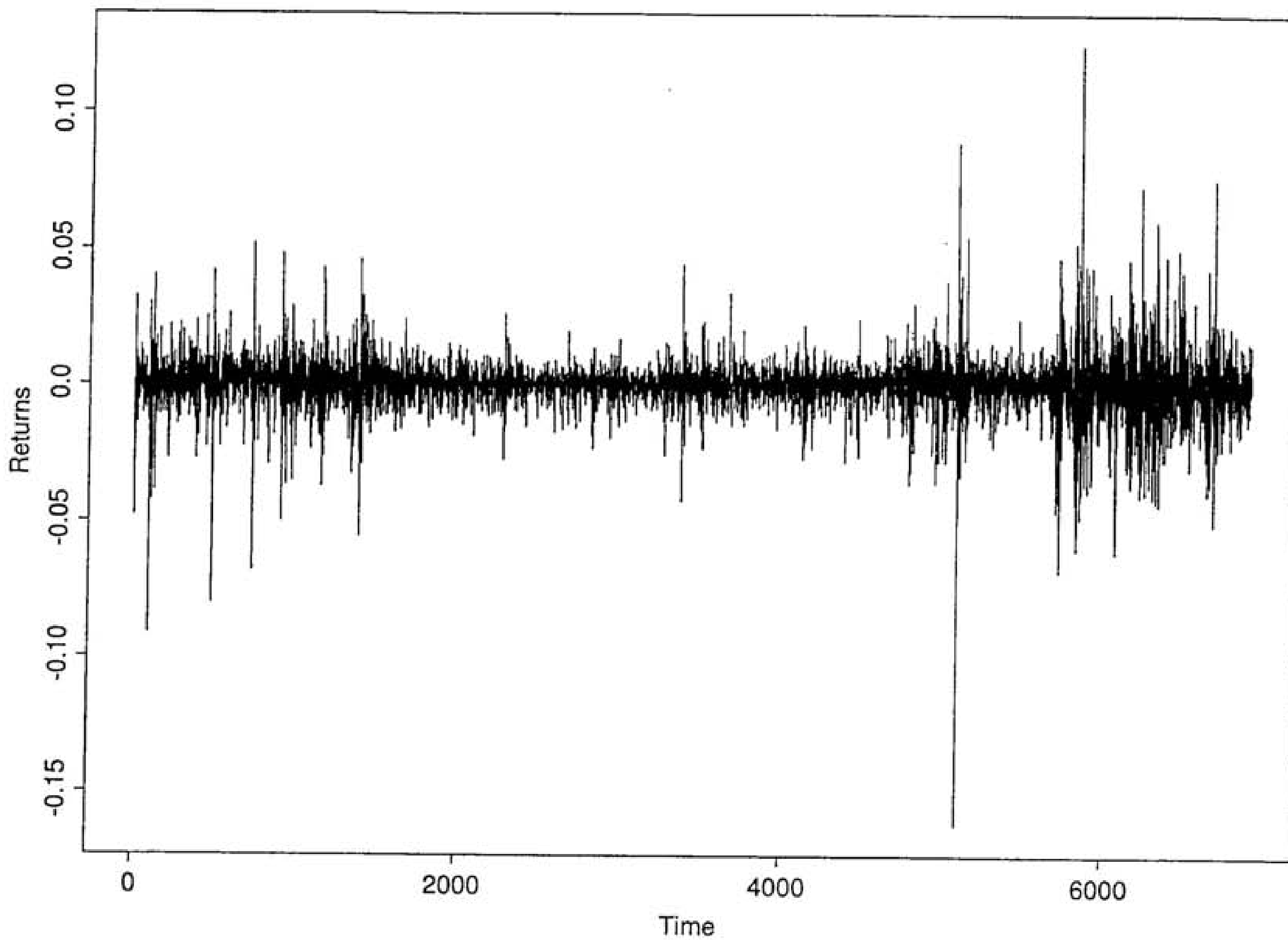


Figure 4.4: The time series data of SP500 returns

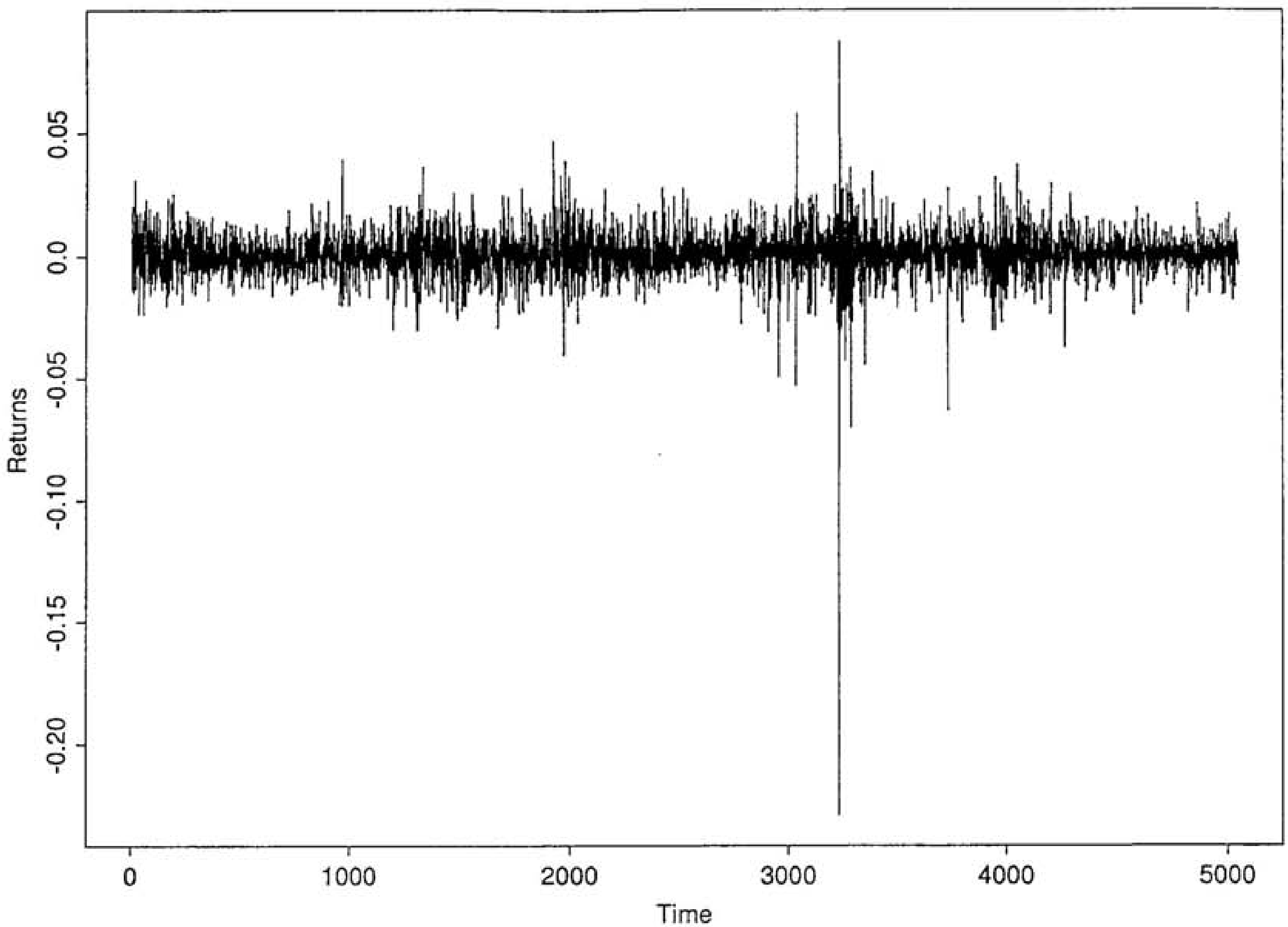


Table 4.7: The parameters estimated by maximum likelihood method (Japan 1970-94, U.S. 1975-1994)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	0.00071	0.00728	1.5	1.81926	23259.72*	-46513.44
	K=2	0.00062	0.01234	2.5	1.24933	23216.29	-46426.58
	K=3	0.00057	0.01659	3.5	0.99371	23086.13	-46166.25
	K=4	0.00053	0.02030	4.5	0.84705	22973.96	-45941.91
	K=5	0.00050	0.02364	5.5	0.75012	22883.21	-45760.41
	K=6	0.00048	0.02669	6.5	0.68052	22809.88	-45613.76
	K=7	0.00046	0.02952	7.5	0.62682	22748.05	-45490.10
TOPIX	K=1	0.00059	0.00624	1.5	1.56155	24435.43*	-48864.86
	K=2	0.00053	0.01048	2.5	1.09030	24378.49	-48750.98
	K=3	0.00050	0.01406	3.5	0.87067	24241.41	-48476.82
	K=4	0.00048	0.01720	4.5	0.74208	24122.57	-48239.13
	K=5	0.00046	0.02004	5.5	0.65608	24025.38	-48044.76
	K=6	0.00045	0.02265	6.5	0.59393	23946.60	-47887.19
	K=7	0.00043	0.02539	7.5	0.50387	23865.39	-47724.78
SP500	K=1	0.00040	0.00794	1.5	1.81676	16876.15	-33746.29
	K=2	0.00041	0.01302	2.5	1.27592	16982.50*	-33959.01
	K=3	0.00041	0.01710	3.5	1.01967	16972.93	-33939.85
	K=4	0.00042	0.02058	4.5	0.87124	16949.36	-33892.73
	K=5	0.00042	0.02365	5.5	0.77245	16925.50	-33845.00
	K=6	0.00042	0.02644	6.5	0.69865	16904.13	-33802.27
	K=7	0.00043	0.02901	7.5	0.64984	16883.94	-33761.88

Table 4.8: The parameters estimated by maximum likelihood method (1970-1974)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	0.00113	0.00825	1.5	1.68441	4788.74	-9571.48
	K=2	0.00097	0.01370	2.5	1.17504	4793.23*	-9580.46
	K=3	0.00088	0.01823	3.5	0.93737	4773.13	-9540.26
	K=4	0.00083	0.02217	4.5	0.79846	4753.82	-9501.64
	K=5	0.00078	0.02571	5.5	0.70557	4737.46	-9468.92
	K=6	0.00075	0.02895	6.5	0.63837	4723.82	-9441.64
	K=7	0.00072	0.03196	7.5	0.58638	4712.26	-9418.53
TOPIX	K=1	0.00112	0.00642	1.5	1.48838	5163.69*	-10321.37
	K=2	0.00095	0.01065	2.5	1.05014	5158.58	-10311.17
	K=3	0.00087	0.01422	3.5	0.84111	5134.23	-10262.47
	K=4	0.00081	0.01735	4.5	0.71708	5112.41	-10218.82
	K=5	0.00076	0.02017	5.5	0.63366	5094.37	-10182.75
	K=6	0.00073	0.02275	6.5	0.57393	5079.51	-10153.01
	K=7	0.00070	0.02515	7.5	0.52601	5067.06	-10128.11

Table 4.9: The parameters estimated by maximum likelihood method (1975-1979)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	0.00062	0.00532	1.5	1.47270	5379.50	-10753.00
	K=2	0.00056	0.00863	2.5	1.03991	5409.44	-10812.89
	K=3	0.00052	0.01130	3.5	0.83288	5410.42*	-10814.83
	K=4	0.00049	0.01358	4.5	0.71088	5407.48	-10808.96
	K=5	0.00048	0.01559	5.5	0.62908	5404.15	-10802.29
	K=6	0.00046	0.01740	6.5	0.56843	5401.23	-10796.46
	K=7	0.00046	0.01907	7.5	0.52405	5398.44	-10790.87
TOPIX	K=1	0.00055	0.00445	1.5	1.33340	5686.27	-11366.54
	K=2	0.00049	0.00709	2.5	0.95121	5720.01	-11434.03
	K=3	0.00045	0.00924	3.5	0.76458	5723.87*	-11441.73
	K=4	0.00043	0.01107	4.5	0.65368	5723.01	-11440.02
	K=5	0.00042	0.01268	5.5	0.57915	5721.25	-11436.51
	K=6	0.00041	0.01414	6.5	0.52431	5719.63	-11433.26
	K=7	0.00040	0.01548	7.5	0.48295	5717.81	-11429.62
SP500	K=1	0.00036	0.00788	1.5	1.37360	4340.60	-8675.19
	K=2	0.00030	0.01228	2.5	0.99520	4385.34	-8764.69
	K=3	0.00028	0.01583	3.5	0.80426	4395.58	-8785.16
	K=4	0.00028	0.01884	4.5	0.69000	4398.71	-8791.41
	K=5	0.00028	0.02150	5.5	0.61278	4399.65	-8793.30
	K=6	0.00028	0.02389	6.5	0.55402	4400.05*	-8794.10
	K=7	0.00028	0.02609	7.5	0.51313	4399.54	-8793.09

Table 4.10: The parameters estimated by maximum likelihood method (1980-1984)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	0.00065	0.00547	1.5	1.61547	5313.39	-10620.78
	K=2	0.00057	0.00890	2.5	1.14336	5339.70*	-10673.41
	K=3	0.00054	0.01169	3.5	0.91662	5334.93	-10663.87
	K=4	0.00051	0.01409	4.5	0.78355	5327.01	-10648.02
	K=5	0.00049	0.01622	5.5	0.69479	5319.53	-10633.07
	K=6	0.00048	0.01815	6.5	0.63207	5313.05	-10620.10
	K=7	0.00047	0.01993	7.5	0.58105	5307.40	-10608.80
TOPIX	K=1	0.00054	0.00492	1.5	1.41833	5496.29	-10986.58
	K=2	0.00053	0.00793	2.5	1.00799	5517.82*	-11029.63
	K=3	0.00052	0.01041	3.5	0.80936	5511.43	-11016.87
	K=4	0.00052	0.01256	4.5	0.69201	5502.69	-10999.37
	K=5	0.00051	0.01447	5.5	0.61310	5494.73	-10983.46
	K=6	0.00051	0.01619	6.5	0.55298	5488.15	-10970.29
	K=7	0.00051	0.01781	7.5	0.51192	5482.11	-10958.22
SP500	K=1	-0.00005	0.00927	1.5	1.68928	4072.40	-8138.80
	K=2	0.00006	0.01486	2.5	1.20058	4114.12	-8222.23
	K=3	0.00012	0.01933	3.5	0.95910	4121.18	-8236.36
	K=4	0.00016	0.02311	4.5	0.81796	4122.13*	-8238.26
	K=5	0.00019	0.02644	5.5	0.72392	4121.54	-8237.08
	K=6	0.00021	0.02945	6.5	0.65511	4120.56	-8235.13
	K=7	0.00022	0.03221	7.5	0.60358	4119.48	-8232.97

Table 4.11: The parameters estimated by maximum likelihood method (1985-1989)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	0.00126	0.00682	1.5	1.69636	4710.57	-9415.14
	K=2	0.00122	0.01125	2.5	1.19616	4716.33*	-9426.65
	K=3	0.00119	0.01493	3.5	0.96360	4697.16	-9388.32
	K=4	0.00116	0.01814	4.5	0.82915	4678.24	-9350.47
	K=5	0.00114	0.02102	5.5	0.74008	4661.86	-9317.72
	K=6	0.00112	0.02367	6.5	0.67932	4647.84	-9289.67
	K=7	0.00111	0.02611	7.5	0.62726	4635.80	-9265.59
TOPIX	K=1	0.00102	0.00695	1.5	1.52129	4688.47*	-9370.94
	K=2	0.00102	0.01147	2.5	1.08440	4686.06	-9366.13
	K=3	0.00102	0.01526	3.5	0.87677	4663.03	-9320.05
	K=4	0.00101	0.01858	4.5	0.75470	4641.52	-9277.04
	K=5	0.00100	0.02157	5.5	0.67310	4623.24	-9240.48
	K=6	0.00099	0.02431	6.5	0.61322	4607.99	-9209.82
	K=7	0.00099	0.02686	7.5	0.56904	4594.52	-9183.04
SP500	K=1	0.00102	0.00815	1.5	2.23198	4093.98*	-8181.95
	K=2	0.00104	0.01383	2.5	1.49805	4093.07	-8180.13
	K=3	0.00103	0.01847	3.5	1.18130	4072.62	-8139.24
	K=4	0.00102	0.02249	4.5	1.00214	4053.45	-8100.91
	K=5	0.00101	0.02609	5.5	0.88396	4036.97	-8067.94
	K=6	0.00100	0.02937	6.5	0.80370	4022.81	-8039.63
	K=7	0.00098	0.03241	7.5	0.74302	4010.57	-8015.13

Table 4.12: The parameters estimated by maximum likelihood method (1990-1994)

models		μ	τ	b	σ	LL	AIC
NK225	K=1	-0.00061	0.01350	1.5	2.03554	3421.98	-6837.96
	K=2	-0.00065	0.02261	2.5	1.40528	3438.85*	-6871.70
	K=3	-0.00067	0.02997	3.5	1.12119	3432.57	-6859.14
	K=4	-0.00068	0.03624	4.5	0.95973	3425.14	-6844.28
	K=5	-0.00068	0.04179	5.5	0.85294	3418.69	-6831.39
	K=6	-0.00068	0.04680	6.5	0.77453	3413.33	-6820.65
	K=7	-0.00068	0.05140	7.5	0.71612	3408.83	-6811.66
TOPIX	K=1	-0.00076	0.01077	1.5	1.60129	3700.29	-7394.57
	K=2	-0.00070	0.01803	2.5	1.10309	3703.28*	-7400.56
	K=3	-0.00068	0.02404	3.5	0.87593	3688.98	-7371.97
	K=4	-0.00066	0.02924	4.5	0.74445	3675.87	-7345.75
	K=5	-0.00065	0.03390	5.5	0.65710	3665.16	-7324.33
	K=6	-0.00064	0.03812	6.5	0.59348	3656.54	-7307.08
	K=7	-0.00064	0.04203	7.5	0.54563	3649.34	-7292.68
SP500	K=1	0.00026	0.00672	1.5	2.08554	4402.90	-8799.80
	K=2	0.00026	0.01125	2.5	1.40256	4426.27*	-8846.55
	K=3	0.00024	0.01485	3.5	1.10224	4424.01	-8842.02
	K=4	0.00024	0.01790	4.5	0.93495	4419.32	-8832.64
	K=5	0.00023	0.02060	5.5	0.82528	4414.90	-8823.81
	K=6	0.00023	0.02303	6.5	0.74633	4411.13	-8816.26
	K=7	0.00023	0.02526	7.5	0.68676	4407.94	-8809.88

4.3.4 Estimation of Parameters by Local Linearization Methods

In order to estimate the parameters of stochastic differential equation of Pearson type IV (with asymmetric distribution and continuous shape parameter), we consider the local linearization method by using the first order Taylor's expansion (Ozaki (1985a, 1989, 1992a, 1992b, 1993), Biscay, Jimenez, Riera, and Valdes (1994), Nagahara (1995b)). In this chapter, we call it simply extended local linearization method. We consider the following diffusion process,

$$dx = \sigma^2\{(1-b)(x-\mu) + b\tau\delta\}dt + \sigma\tau\sqrt{(\frac{x-\mu}{\tau})^2 + 1^2}dw. \tag{4.58}$$

According to subsection 4.3.1, we apply the following transformation

$$y = \int_{-\infty}^{x} \frac{d\xi}{\sigma\sqrt{(\frac{\xi-\mu}{\tau})^2 + 1^2}} = \frac{\tau}{\sigma}\text{arcsinh}\left(\frac{x-\mu}{\tau}\right). \tag{4.59}$$

Then, x is given by

$$x = \mu + \tau\sinh\left(\frac{\sigma y}{\tau}\right), \tag{4.60}$$

and we obtain the equation for y,

$$dy = \left\{\tau\sigma\left(\frac{1}{2} - b\right)\tanh\left(\frac{\sigma y}{\tau}\right) + \frac{\sigma b\tau\delta}{\cosh(\frac{\sigma y}{\tau})}\right\}dt + \tau dw. \tag{4.61}$$

According to the local linearization method, we estimate the parameters by the simple version of the log-likelihood as follows. This simply extended local linearization and log-likelihood is as follows.

We defined the simply extended local linearization of the following equation,

$$dx_t = f(x_t)dt + \sigma dw_t. \tag{4.62}$$

For $t > s$, we approximate $f(x_t)$ by the first order Taylor expansion,

$$\begin{aligned} f(x_t) &\approx f(x_s) + \frac{\partial f(x_s)}{\partial x_s} \times (x_t - x_s) \\ &\approx L_s x_t + M_s, \end{aligned} \tag{4.63}$$

where

$$
\begin{aligned}
L_s &= \frac{\partial f(x_s)}{\partial x_s}, && (4.64)\\
M_s &= f(x_s) - \frac{\partial f(x_s)}{\partial x_s} x_s \\
&= f(x_s) - L_s x_s. && (4.65)
\end{aligned}
$$

Therefore, (4.62) can be approximated by

$$
\begin{aligned}
dx_t &= f(x_t)dt + \sigma dw_t, \\
&\approx (L_s x_t + M_s)dt + \sigma dw_t. && (4.66)
\end{aligned}
$$

We assume that L_s and M_s are constants for a small interval.

In general, if we have the ordinary differential equation as follows,

$$
\frac{dx}{du} = P(u)x + Q(u), \tag{4.67}
$$

the solution in the case that $x(s) = x_s$ is

$$
x_t = e^{\int_s^t P(u)du} \left(\int_s^t Q(u) e^{\int_s^u -P(v)dv} du + x_s \right). \tag{4.68}
$$

In this case, we assume that $P(u)$ is L_s and $Q(u)$ is $M_s + \sigma \frac{dw}{du}$.

Therefore, we obtain the following solution.

$$
\begin{aligned}
x_t &= e^{\int_s^t Ldu} \left\{ \int_s^t (M_s + \sigma \frac{dw}{du}) e^{\int_s^u -Ldv} du + x_s \right\} \\
&= e^{L(t-s)} \left\{ \int_s^t (M_s + \sigma \frac{dw}{du}) e^{\int_s^u -Ldv} du + x_s \right\} \\
&= e^{L(t-s)} \left\{ e^{Ls} \frac{(e^{-Lt} - e^{-Ls})}{-L} M_s + e^{Ls} \int_s^t \sigma e^{-Lu} dw + x_s \right\} \\
&= e^{L(t-s)} \left\{ e^{Ls} \frac{(e^{-Lt} - e^{-Ls})}{-L} (f(x_s) - L_s x_s) + x_s + e^{Ls} \int_s^t \sigma e^{-Lu} dw \right\} \\
&= e^{L(t-s)} \left\{ e^{Ls} \frac{(e^{-Lt} - e^{-Ls})}{-L} f(x_s) + (e^{L(s-t)} - 1)x_s + x_s + e^{Ls} \int_s^t \sigma e^{-Lu} dw \right\} \\
&= x_s + \frac{f(x_s)}{L_s} (e^{L_s(t-s)} - 1) + \sigma \int_s^t e^{L_s(t-u)} dw. && (4.69)
\end{aligned}
$$

By the Ito's rule, it is obvious that this is the solution of (4.66).

In another way, we can also consider the following procedure.

By the Girsanov theorem (Girsanov (1960)), we obtain

$$dx_t = (L_s x_t)dt + \sigma d\tilde{w}_t, \tag{4.70}$$

where $\tilde{w}_t$ is the Wiener process after a change of measure defined by

$$\begin{aligned} \tilde{w}_t &= w_t - \int_s^t \beta(u)du, & (4.71) \\ \beta(u) &= -\frac{1}{\sigma}M_s. & (4.72) \end{aligned}$$

Transformation of $y_t = e^{-L_s t}x_t$ yields

$$dx_t = (L_s x_t)dt + e^{L_s t}dy_t. \tag{4.73}$$

Therefore, we have

$$dy_t = \sigma e^{-L_s t}d\tilde{w}_t. \tag{4.74}$$

Solving this formula, we obtain

$$\begin{aligned} y_t &= y_s + \sigma \int_s^t e^{-L_s u}d\tilde{w}_u \\ &= y_s + \int_s^t M_s e^{-L_s u}du + \sigma \int_s^t e^{-L_s u}dw_u \\ &= y_s + \frac{f(x_s) - L_s x_s}{-L_s}(e^{-L_s t} - e^{-L_s s}) + \sigma \int_s^t e^{-L_s u}dw_u. & (4.75) \end{aligned}$$

By using transformation of $y_t = e^{-L_s t}x_t$, the equation for x_t is obtained by

$$x_t = x_s + \frac{f(x_s)}{L_s}(e^{L_s(t-s)} - 1) + \sigma \int_s^t e^{L_s(t-u)}dw_u, \tag{4.76}$$

where

$$Var_s(x_t) = \frac{\{\exp(2L_s(t-s)) - 1\}}{2L_s}\sigma^2. \tag{4.77}$$

This conditional variance differs from that of ordinary local linearization (Ozaki (1992b), Biscay, Jimenez, Riera, and Valdes (1994), Shoji and Ozaki (1994), Nagahara (1995b)). The $Var_s(x_t)$ of ordinary local linearization is given by

$$Var_s(x_t) = \frac{\{\exp(2K_s(t-s)) - 1\}}{2K_s}\sigma^2, \tag{4.78}$$

where

$$K_s = \frac{1}{t-s}\log\left\{1 + \frac{f(x_s)}{x_s L_s}(e^{L_s(t-s)} - 1)\right\}. \tag{4.79}$$

If $x_1, \ldots, x_T$, are equisampled observation with time interval Δt, the log-likelihood is obtained by

$$\begin{aligned}
\log(p(x_1, \ldots, x_T)) &= \log(p(x_1)) + \sum_{t=1}^{T-1} \log p(x_{t+1}|x_t), \\
&= \log(p(x_1)) + \sum_{t=1}^{T-1} \frac{1}{2V_t}\{x_{t+1} - x_t - \frac{f_t}{L_t}(\exp L_t \Delta t - 1)\}^2 \\
&\quad -\frac{T-1}{2}\log(2\pi V_t)
\end{aligned} \tag{4.80}$$

Furthermore, if the transformation of $y = \psi(x)$ exists, the likelihood is given by

$$p(x_1, ..., x_T) = p(y_1, ..., y_T)\left|\frac{\partial(y_1, ..., y_T)}{\partial(x_1, ..., x_T)}\right|. \tag{4.81}$$

Therefore, the log-likelihood is given by

$$\log(p(x_1, ..., x_T)) = \log(p(y_1, ..., y_T)) + \sum_{t=1}^{T-1} \log(\frac{d\psi}{dx}) \tag{4.82}$$

In this chapter, since the estimated parameters, especially shape parameter, of this simply extended local linearization are close to the parameters estimated by the exact maximum likelihood method using the transition probability density function, we adopt this simply extended local linearization whose difference with ordinary local linearization is only the conditional variance of x_t, V_t. The detail of these comparison is given in the section 4.5.

According to this local linearization method, we have

$$L = \frac{\partial f}{\partial y} = \left\{\left(\frac{1}{2} - b\right) - b\delta \sinh\left(\frac{\sigma y}{\tau}\right)\right\} \times \frac{\sigma^2}{\cosh^2(\frac{\sigma y}{\tau})}, \tag{4.83}$$

$$V_t = \frac{\{\exp(2L_t\Delta t) - 1\}}{2L_t}\tau^2, \tag{4.84}$$

and the log-likelihood is given by

$$\begin{aligned}
1(\theta) &= -\sum_{t=1}^{T-1} \frac{1}{2V_t}\{y_{t+1} - y_t - \frac{f_t}{L_t}(\exp L_t\Delta t - 1)\}^2 \\
&\quad -\frac{T-1}{2}\log(2\pi V_t) + \sum_{t=1}^{T-1} \log\frac{1}{\sigma\sqrt{(\frac{x_t-\mu}{\tau})^2 + 1^2}}.
\end{aligned} \tag{4.85}$$

In this chapter, the time interval, Δt is set to 1. And we estimate the parameters of stochastic differential equation by maximizing this log-likelihood.

4.3.5 Stochastic Differential Equation Estimated by Local Linearization

We estimate the parameters and compute AIC for each model. The results for each term are shown in Table 4.13–Table 4.18. The $*$ shows the minimum AIC. The results of the parameters of stochastic differential equation are close to the parameters of the stationary symmetric and asymmetric distributions. Therefore, it may be concluded that the local linearization methods are useful to estimate the parameters of stochastic differential equation of Pearson IV type (asymmetric distribution and continuous shape parameter) and Pearson VII type (symmetric distribution and continuous shape parameter).

The details of results are as follows. The results for entire interval, i.e., 1970-1994 for Japan and 1975-1994 for the United States, are shown in Table 4.13. For Japan data, the minimum AIC values are achieved at $b = 1.66$ which are close to the results of the stationary distribution, $b = 1.75 - 1.78$ and the transition probability density method $b = 1.5$. The asymmetric distribution has smaller AIC. For the U.S. data, the minimum AIC value is achieved at $b = 2.56$ which is close to the result of the stationary distribution, $b = 2.70$ and the transition probability density method $b = 2.5$. The symmetric distribution has smaller AIC.

The results of only Japanese data for the term, 1970-1974, are shown in Table 4.14. The minimum AIC values of Nikkei 225 and TOPIX are achieved at $b = 1.80$ and $b = 1.73$, respectively. They are close to the results of the stationary distribution, NK225 ($b = 1.98$), TOPIX ($b = 1.56$).

The results of Japan and U.S. data for the term, 1975-1979, are shown in Table 4.15. For Japan data, the minimum AIC values of Nikkei 225 and TOPIX are achieved at $b = 2.29$ and $b = 2.78$, respectively. They are close to the results of the stationary distribution, NK225 ($b = 3.06$), TOPIX ($b = 3.62$). For the U.S. data, the minimum AIC value is achieved at $b = 5.42$ which is close to the result of the stationary distribution, $b = 6.57$.

The results of Japan and U.S. data for the term, 1980-1984, are shown in Table 4.16. In Japan, the minimum AIC values of Nikkei 225 and TOPIX are achieved at $b = 2.24$ and $b = 2.05$, respectively. They are close to the results of the stationary distribution, NK225 ($b = 2.59$), TOPIX ($b = 2.49$). For the U.S. data, the minimum AIC value is achieved at $b = 4.86$ which is close to the result of the stationary distribution, $b = 4.41$.

The results of Japan and U.S. data for the term, 1985-1989, are shown in Table 4.17. The minimum AIC values of Nikkei 225 and TOPIX are achieved at $b = 1.91$ and $b = 1.72$, respectively. They are close to the results of the stationary distribution, NK225 ($b = 2.02$), TOPIX ($b = 1.93$). For the U.S. data, the minimum AIC value is achieved at $b = 1.85$ which is close to the result of the stationary distribution, $b = 1.86$.

The results of Japan and U.S. data for the term, 1990-1994, are shown in Table 4.18. For Japan data, The minimum AIC values of Nikkei 225 and TOPIX are achieved at $b = 2.05$ and $b = 1.72$, respectively. They are close to the results of the stationary distribution, NK225 ($b = 2.34$), TOPIX ($b = 1.97$). For the U.S. data, the minimum AIC value is achieved at $b = 2.34$ which is close to the result of the stationary distribution, $b = 2.64$.

The summary of conclusion in this subsection is as follows. The estimated parameters of stochastic differential equation by local linearization method, especially b, are close to the estimated parameters obtained by the maximum likelihood method of the transition probability density function. Therefore, local linearization method is useful to estimate the parameters of stochastic differential equation which corresponds to the Pearson IV (asymmetric and continuous shape parameter) and the Pearson VII (symmetric and continuous shape parameter).

4.4 Conclusion

The summary of conclusion in this chapter are as follows. First, the stationary distributions of stock returns are heavier tailed than normal distributions. Especially, the Pearson type VII and IV are well fitted to the data distributions. Second, the stochastic differential equation related to the stationary distribution of the Pearson type VII and IV are valid for the generating process of stock returns. Third, the maximum likelihood method by using the transition probability density function and the local linearization method are very useful to estimate the parameters of this stochastic differential equation. Forth, when the shape parameter, b, is low, there is a possibility that the big price change occur. The fact that this shape parameter differs with the terms indicates that the stability of the stock markets is changing.

Table 4.13: The parameters estimated by local linearization method (Japan 1970-1994, U.S. 1975-1994)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	0.00070	0.00796	0.00	1.65285	1.25206	23517.14	-47026.28
	(Type IV)	0.00159	0.00808	-0.07493	1.66125	1.24950	23530.10	-47050.20*
TOPIX	(Type VII)	0.00058	0.00682	0.00	1.60539	1.18916	24648.47	-49288.94
	(Type IV)	0.00109	0.00681	-0.05163	1.60736	1.18916	24654.42	-49298.84*
SP500	(Type VII)	0.00040	0.01389	0.00	2.56937	0.92077	16972.14	-33936.28*
	(Type IV)	0.00040	0.01389	0.00009	2.56875	0.92080	16972.14	-33934.28

Table 4.14: The parameters estimated by local linearization method (1970-1974)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	0.00107	0.01001	0.00	1.77457	1.14736	4829.82	-9651.64
	(Type IV)	0.00283	0.01005	-0.12262	1.80520	1.14030	4836.16	-9662.33*
TOPIX	(Type VII)	0.00105	0.00766	0.00	1.69666	1.10196	5190.99	-10373.98
	(Type IV)	0.00252	0.00769	-0.13637	1.73057	1.09247	5198.57	-10387.14*

Table 4.15: The parameters estimated by local linearization method (1975-1979)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	0.00048	0.00820	0.00	2.34039	0.90087	5431.51	-10855.03
	(Type IV)	0.00136	0.00803	-0.08355	2.29912	0.90962	5433.73	-10857.46*
TOPIX	(Type VII)	0.00044	0.00801	0.00	2.83069	0.75897	5733.60	-11459.20
	(Type IV)	0.00131	0.00786	-0.08813	2.78494	0.76540	5735.55	-11461.10*
SP500	(Type VII)	0.00026	0.02123	0.00	5.30824	0.56114	4406.54	-8805.08*
	(Type IV)	-0.00189	0.02141	0.08823	5.42218	0.55584	4407.39	-8804.79

Table 4.16: The parameters estimated by local linearization method (1980-1984)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	0.00061	0.00816	0.00	2.20870	0.97090	5353.97	-10699.94
	(Type IV)	0.00175	0.00822	-0.10472	2.24466	0.96267	5357.60	-10705.20*
TOPIX	(Type VII)	0.00058	0.00679	0.00	2.04545	0.95693	5535.13	-11062.26*
	(Type IV)	0.00095	0.00682	-0.04081	2.05843	0.95345	5535.72	-11061.44
SP500	(Type VII)	0.00022	0.02475	0.00	4.77957	0.64507	4116.64	-8225.27
	(Type IV)	-0.00413	0.02470	0.15339	4.86736	0.64019	4119.69	-8229.37*

Table 4.17: The parameters estimated by local linearization method (1985-1989)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	0.00121	0.00877	0.00	1.89859	1.10823	4757.90	-9507.79
	(Type IV)	0.00228	0.00882	-0.08698	1.91826	1.10137	4760.56	-9511.11*
TOPIX	(Type VII)	0.00093	0.00810	0.00	1.72100	1.12247	4733.85	-9459.70*
	(Type IV)	0.00110	0.00810	-0.01437	1.72093	1.12244	4733.93	-9457.86
SP500	(Type VII)	0.00114	0.01086	0.00	1.80873	1.20719	4085.78	-8163.56
	(Type IV)	0.00218	0.01110	-0.06526	1.85269	1.18973	4087.33	-8164.67*

Table 4.18: The parameters estimated by local linearization method (1990-1994)

models		μ	τ	δ	b	σ	LL	AIC
NK225	(Type VII)	-0.00054	0.01886	0.00	2.05312	1.10258	3461.17	-6914.34*
	(Type IV)	-0.00045	0.01882	-0.00351	2.04989	1.10374	3461.18	-6912.36
TOPIX	(Type VII)	-0.00059	0.01270	0.00	1.72272	1.14109	3733.86	-7459.72*
	(Type IV)	-0.00086	0.01270	0.01455	1.72296	1.14042	3733.94	-7457.88
SP500	(Type VII)	0.00022	0.01096	0.00	2.34282	1.00644	4424.27	-8840.54*
	(Type IV)	0.00052	0.01093	-0.02001	2.33639	1.00827	4424.40	-8838.80

4.5 A Comparison of Parameter Estimation Methods for a Stocahastic Differential Equation

First, we compare the parameters of stationary distribution by maximum likelihood method to the parameters of stochastic differential equation by local linearization method. Second, we use the transition probability density function of stochastic differential equation and compare the parameters of stochastic differential equation by maximum likelihood method using the transition probability density function with the parameters of stochastic differential equation by local linearization method.

We estimate the parameters of the stochastic differential equation by local linearization method researched by Ozaki (1985a, 1989, 1992a, 1992b, 1993), Biscay, Jimenez, Riera and Valdes (1994), Shoji and Ozaki (1994), Nagahara (1995b). In this section, two different local linearization method ; one is based on the first order Taylor's expansion (we call it simply extended local linearization in this note) (Biscay, Jimenez, Riera and Valdes (1994), Nagahara (1995b)), another is based on Ito's expansion (new local linearization, Shoji and Ozaki (1994)), are shown. We compare the exact maximum likelihood method by using the transition probability density function and two local linearization. In the result, the Taylor's expansion local linearization method (simply extended local linearization) has closer estimated parameters to that of the exact maximum likelihood method by using this transition probability density function than the Ito's expansion local linearization method (new local linearization).

In subsection 4.5.1 and 4.5.2, the parameters of stationary distribution estimated by maximum likelihood method and the parameters of related stochastic differntial equation by local linearization are shown. In subsection 4.5.3 and 4.5.4, the transition probability of related stochastic differntial equation are shown. And the parameters estimated by maximum likelihood method using this transition probability density function and local linearization are compared.

4.5.1 Parameter Estimation by Simply Extended Local Linearization and New Local linearization

In order to estimate the parameters of stochastic differential equation of Pearson IV type (asymmetric distribution and continuous shape parameter), we consider simply extended local linearization method and new local linearization(Ozaki (1985a, 1989, 1992a, 1992b, 1993), Biscay, Jimenez, Riera and Valdes (1994), and Nagahara (1995b)) and new local linearization (Shoji and Ozaki (1994)). We consider the following diffusion process,

$$dx = \sigma^2\{(1-b)(x-\mu) + b\tau\delta\}dt + \sigma\tau\sqrt{(\frac{x-\mu}{\tau})^2 + 1^2}dw. \tag{4.86}$$

According to subsection 4.3.2, we apply the following transformation:

$$y = \int_{-\infty}^{x} \frac{d\xi}{\sigma\sqrt{(\frac{\xi-\mu}{\tau})^2 + 1^2}} \tag{4.87}$$

$$= \frac{\tau}{\sigma}\text{arcsinh}(\frac{x-\mu}{\tau}). \tag{4.88}$$

Then, x is given by

$$x = \mu + \tau\sinh(\frac{\sigma y}{\tau}), \tag{4.89}$$

and we obtain the equation for y,

$$dy = \left\{\tau\sigma(\frac{1}{2} - b)\tanh(\frac{\sigma y}{\tau}) + \frac{\sigma b\tau\delta}{\cosh(\frac{\sigma y}{\tau})}\right\}dt + \tau dw. \tag{4.90}$$

In this section, we compare two local linearization methods, one is based on the first order Taylor's expansion of the drift function, (simply extnded local linerization), another is based on the Ito's expansion of the drift function (new local linerization, Shoji and Ozaki (1994)).

First, we introduce the log-likelihood of the simply extended local linearization method. According to the last section, we obtain the following log-likelihood.

$$L = \frac{\partial f}{\partial y} = \{(\frac{1}{2} - b) - b\delta\sinh(\frac{\sigma y}{\tau})\} \times \frac{\sigma^2}{\cosh^2(\frac{\sigma y}{\tau})}. \tag{4.91}$$

$$V_t = \frac{\{\exp(2L_t\Delta t) - 1\}}{2L_t}\tau^2. \tag{4.92}$$

$$(\log - \text{likelihood}) = -\sum_{t=1}^{T-1}\frac{1}{2V_t}\{y_{t+1} - y_t - \frac{f_t}{L_t}(\exp L_t\Delta t - 1)\}^2 \tag{4.93}$$

$$-\frac{T-1}{2}\log(2\pi V_t) + \sum_{t=1}^{T-1}\log\frac{1}{\sigma\sqrt{(\frac{x_t-\mu}{\tau})^2 + 1^2}} \tag{4.94}$$

In this section, Δt equals to 1. And we estimate the parameters of stochastic differential equation by maximizing this log-likelihood.

Second, we introduce the log-likelihood of the new local linearization method. According to Shoji and Ozaki (1994), new local linearization is defined by the following,

$$dx_t = f(x_t, t)dt + \sigma dw_t. \tag{4.95}$$

In the case of $t > s$, they approximate $f(x_t, t)$ by the Ito's lemma.

$$\begin{aligned} f(x_t, t) &= f(x_s, s) + \left(\frac{\sigma^2}{2}\frac{\partial^2 f(x_s, s)}{\partial x_s^2} + \frac{\partial f(x_s, s)}{\partial s}\right)(t - s) + \frac{\partial f(x_s, s)}{\partial x_s} \times (x_t - x_s) \\ &= L_s x_t + M_s t + N_s, \end{aligned} \tag{4.96}$$

$$L_s = \frac{\partial f(x_s, s)}{\partial x_s}. \tag{4.97}$$

$$M_s = \frac{\sigma^2}{2}\frac{\partial^2 f(x_s, s)}{\partial x_s^2} + \frac{\partial f(x_s, s)}{\partial s} \tag{4.98}$$

$$N_s = f(x_s, s) - L_s x_s - M_s s. \tag{4.99}$$

Therefore,

$$\begin{aligned} dx_t &= f(x_t, t)dt + \sigma dw_t, \\ &= (L_s x_t + M_s t + N_s)dt + \sigma dw_t. \end{aligned} \tag{4.100}$$

They assume that L_s, M_s and N_s are constant for a small interval.

By the Girsanov theorem,

$$dx_t = (L_s x_t)dt + \sigma d\tilde{w}_t, \tag{4.101}$$

where $\tilde{w}_t$ is the Wiener process after a change of mesure defined by

$$\tilde{w}_t = w_t - \int_s^t \beta(u)du, \tag{4.102}$$

$$\beta(u) = -\frac{1}{\sigma}(M_s u + N_s). \tag{4.103}$$

Transformation of $y_t = e^{-L_s t}x_t$,

$$dx_t = (L_s x_t)dt + e^{L_s t}dy_t. \tag{4.104}$$

Therefore,

$$dy_t = \sigma e^{-L_s t}d\tilde{w}_t. \tag{4.105}$$

Solving this formula,

$$\begin{aligned} y_t &= y_s + \sigma\int_s^t e^{-L_s u} d\tilde{w}_t, \\ &= y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dw_u. \end{aligned} \tag{4.106}$$

By transformation of $y_t = e^{-L_s t} x_t$

$$x_t = x_s + \frac{f(x_s, s)}{L_s}(e^{L_s(t-s)} - 1) + \frac{M_s}{L_s^2}\{(\exp L_s \Delta t - 1) - L_s \Delta t\} + \sigma \int_s^t e^{L_s(t-u)} dw_u, \tag{4.107}$$

where

$$Var_s(x_t) = \frac{\{\exp(2L_s(t-s)) - 1\}}{2L_s}\sigma^2. \tag{4.108}$$

If $x_1, ..., x_T$ are given as the same time interval Δt, the log-likelihood is the following.

$$\begin{aligned} \log(p(x_1, ..., x_T)) &= \log(p(x_1)) + \sum_{t=1}^{T-1} \log p(x_{t+1}|x_t), \\ &= \log(p(x_1)) + \sum_{t=1}^{T-1} \frac{1}{2V_t}\Bigg[x_{t+1} - x_t - \frac{f_t}{L_t}(\exp L_t \Delta t - 1) \\ &\quad - \frac{M_t}{L_t^2}\{(\exp L_t \Delta t - 1) - L_t \Delta t\Bigg]^2 \\ &\quad - \frac{T-1}{2}\log(2\pi V_t). \end{aligned} \tag{4.109}$$

Furthermore, if the transformation of $y = \psi(x)$ is existing, the likelihood is given by

$$p(x_1, ..., x_T) = p(y_1, ..., y_T)\left|\frac{\partial(y_1, ..., y_T)}{\partial(x_1, ..., x_T)}\right|. \tag{4.110}$$

Therefore, the log-likelihood is

$$\log(p(x_1, ..., x_T)) = \log(p(y_1, ..., y_T)) + \sum_{t=1}^{T-1} \log(\frac{d\psi}{dx}). \tag{4.111}$$

According to our case, we obtain the following log-likelihood,

$$L = \frac{\partial f}{\partial y} = \left\{(\frac{1}{2} - b) - b\delta \sinh(\frac{\sigma y}{\tau})\right\} \times \frac{\sigma^2}{\cosh^2(\frac{\sigma y}{\tau})}, \tag{4.112}$$

$$\frac{\partial L}{\partial y} = \frac{\partial^2 f}{\partial y^2} = \frac{\sigma^3}{\tau \cosh(\frac{\sigma y}{\tau})}\left\{-b\delta + (2b-1)\frac{\tanh(\frac{\sigma y}{\tau})}{\cosh(\frac{\sigma y}{\tau})} + 2b\delta \tanh^2(\frac{\sigma y}{\tau})\right\}. \tag{4.113}$$

$$M_t = \frac{\tau^2}{2}\frac{\partial L}{\partial y} = \frac{\tau^2}{2}\frac{\partial^2 f}{\partial y^2}, \tag{4.114}$$

$$V_t = \frac{\{\exp(2L_t\Delta t) - 1\}}{2L_t}\tau^2, \tag{4.115}$$

$$\begin{aligned}(\log - \text{likelihood}) \quad = \quad & -\sum_{t=1}^{T-1}\frac{1}{2V_t}\left[y_{t+1} - y_t - \frac{f_t}{L_t}(\exp L_t\Delta t - 1) - \frac{M_t}{L_t^2}\{(\exp L_t\Delta t - 1) - L_t\Delta t\}\right]^2 \\ & -\frac{T-1}{2}\log(2\pi V_t) + \sum_{t=1}^{T-1}\log\frac{1}{\sigma\sqrt{(\frac{x_t-\mu}{\tau})^2 + 1^2}}. \end{aligned} \tag{4.116}$$

And we estimate the parameters of stochastic differential equation by maximizing this log-likelihood.

4.5.2 Empirical Results

The daily rate of returns were obtained from the daily index of Nikkei 225 index, TOPIX index and Standard and Poor's 500 index. The data of Japanese market consist of returns from 5, January, 1970 to 29 December , 1994 for a total of 6912 observations. And, the data of U.S. market consists of returns from 5, January, 1975 to 29 December, 1994 for a total of 4801 observations.

The Table 4.19 - Table 4.35 are the comparison of empirical results. S.D., SELL, and NLL stand for stationary distribution, simply extended local linearization, and new local linearization, respectively. The new local linearization method tends to over-estimate the level of b more than maximum likelihood method for the stationary distribution and the simply extended local lenearization method. And the parameters of the simply extended local lenearization method is nearer to that of the stationary distribution than the new local linearization method. Furthermore, we compare two local linearization to the maximum likelihood method by using the transition probability density function in the next section.

Table 4.19: Comparison of S.D, SELL and NLL (NK225) (1970-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00064	0.00885	0.00	1.78268		23235.95	-46465.90
	(IV type)	0.00152	0.00881	-0.06384	1.78297		23246.50	-46485.00
SELL	(VII type)	0.00070	0.00796	0.00	1.65285	1.25206	23517.14	-47026.28
	(IV type)	0.00159	0.00808	-0.07493	1.66125	1.24950	23530.10	-47050.20
NLL	(VII type)	0.00088	0.01134	0.00	2.36502	1.01548	23278.26	-46548.52
	(IV type)	0.00254	0.01131	-0.10116	2.39635	1.01496	23301.99	-46593.98

Table 4.20: Comparison of S.D, SELL and NLL (NK225) (1970-1974)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00096	0.01102	0.00	1.97605		4787.22	-9568.44
	(IV type)	0.00291	0.01096	-0.11871	1.98996		4793.90	-9579.79
SELL	(VII type)	0.00107	0.01001	0.00	1.77457	1.14736	4829.82	-9651.64
	(IV type)	0.00283	0.01005	-0.12262	1.80520	1.14030	4836.16	-9662.33
NLL	(VII type)	0.00131	0.01343	0.00	2.49348	0.96837	4792.77	-9577.54
	(IV type)	0.00433	0.01383	-0.15722	2.65900	0.94979	4803.47	-9596.94

Table 4.21: Comparison of S.D, SELL and NLL (NK225) (1975-1979)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00048	0.01033	0.00	3.10319		5389.06	-10772.13
	(IV type)	0.00144	0.01017	-0.07050	3.06240		5390.70	-10773.40
SELL	(VII type)	0.00048	0.00820	0.00	2.34039	0.90087	5431.51	-10855.03
	(IV type)	0.00136	0.00803	-0.08355	2.29912	0.90962	5433.73	-10857.46
NLL	(VII type)	0.00056	0.01150	0.00	3.71735	0.77557	5418.91	-10829.82
	(IV type)	0.00238	0.01121	-0.12253	3.66112	0.78827	5423.73	-10837.46

Table 4.22: Comparison of S.D, SELL and NLL (NK225) (1980-1984)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00054	0.00918	0.00	2.59165		5332.04	-10658.08
	(IV type)	0.00149	0.00913	-0.07508	2.59013		5334.09	-10660.17
SELL	(VII type)	0.00061	0.00816	0.00	2.20870	0.97090	5353.97	-10699.94
	(IV type)	0.00175	0.00822	-0.10472	2.24466	0.96267	5357.60	-10705.20
NLL	(VII type)	0.00065	0.01303	0.00	4.06027	0.75808	5332.96	-10657.92
	(IV type)	0.00344	0.01273	-0.16703	4.08068	0.76865	5341.15	-10672.30

Table 4.23: Comparison of S.D, SELL and NLL (NK225) (1985-1989)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00120	0.00929	0.00	2.02186		4713.14	-9420.27
	(IV type)	0.00203	0.00933	-0.05990	2.03760		4714.65	-9421.31
SELL	(VII type)	0.00121	0.00877	0.00	1.89859	1.10823	4757.90	-9507.79
	(IV type)	0.00228	0.00882	-0.08698	1.91826	1.10137	4760.56	-9511.11
NLL	(VII type)	0.00141	0.01278	0.00	2.96960	0.90490	4728.97	-9449.94
	(IV type)	0.00430	0.01277	-0.16005	3.06772	0.90644	4737.44	-9464.88

Table 4.24: Comparison of S.D, SELL and NLL (NK225) (1990-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	-0.00065	0.02137	0.00	2.34822		3441.07	-6876.14
	(IV type)	-0.00071	0.02138	0.00179	2.34926		3441.07	-6874.14
SELL	(VII type)	-0.00054	0.01886	0.00	2.05312	1.10258	3461.17	-6914.34
	(IV type)	-0.00045	0.01882	-0.00351	2.04989	1.10374	3461.18	-6912.36
NLL	(VII type)	-0.00057	0.03406	0.00	4.18618	0.76177	3416.07	-6824.14
	(IV type)	-0.00218	0.03391	0.03588	4.16568	0.76343	3416.37	-6822.74

Table 4.25: Comparison of S.D, SELL and NLL (TOPIX) (1970-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00053	0.00745	0.00	1.75860		24305.54	-48605.08
	(IV type)	0.00105	0.00743	-0.04468	1.75832		24310.99	-48613.98
SELL	(VII type)	0.00058	0.00682	0.00	1.60539	1.18916	24648.47	-49288.94
	(IV type)	0.00109	0.00681	-0.05163	1.60736	1.18916	24654.42	-49298.84
NLL	(VII type)	0.00075	0.00882	0.00	2.15907	1.04596	24546.44	-49084.89
	(IV type)	0.00159	0.00884	-0.06798	2.17899	1.04491	24558.10	-49106.20

Table 4.26: Comparison of S.D, SELL and NLL (TOPIX) (1970-1974)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00092	0.00805	0.00	1.85563		5135.51	-10265.02
	(IV type)	0.00257	0.00797	-0.13534	1.86905		5144.96	-10281.92
SELL	(VII type)	0.00105	0.00766	0.00	1.69666	1.10196	5190.99	-10373.98
	(IV type)	0.00252	0.00769	-0.13637	1.73057	1.09247	5198.57	-10387.14
NLL	(VII type)	0.00118	0.00968	0.00	2.28119	1.00094	5181.07	-10354.14
	(IV type)	0.00324	0.00985	-0.15187	2.39563	0.98821	5192.32	-10374.64

Table 4.27: Comparison of S.D, SELL and NLL (TOPIX) (1975-1979)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00041	0.00954	0.00	3.64864		5688.67	-11371.35
	(IV type)	0.00118	0.00946	-0.06361	3.62304		5689.78	-11371.56
SELL	(VII type)	0.00044	0.00801	0.00	2.83069	0.75897	5733.60	-11459.20
	(IV type)	0.00131	0.00786	-0.08813	2.78494	0.76540	5735.55	-11461.10
NLL	(VII type)	0.00048	0.01048	0.00	4.27462	0.68967	5729.08	-11450.16
	(IV type)	0.00205	0.01019	-0.12026	4.18228	0.70318	5732.85	-11455.70

Table 4.28: Comparison of S.D, SELL and NLL (TOPIX) (1980-1984)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00052	0.00791	0.00	2.49152		5490.92	-10975.85
	(IV type)	0.00070	0.00791	-0.01615	2.49293		5491.03	-10974.06
SELL	(VII type)	0.00058	0.00679	0.00	2.04545	0.95693	5535.13	-11062.26
	(IV type)	0.00095	0.00682	-0.04081	2.05843	0.95345	5535.72	-11061.44
NLL	(VII type)	0.00071	0.00916	0.00	3.03576	0.84410	5524.89	-11041.77
	(IV type)	0.00168	0.00914	-0.08019	3.05920	0.84735	5527.38	-11044.77

Table 4.29: Comparison of S.D, SELL and NLL (TOPIX) (1985-1989)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00101	0.00906	0.00	1.93043		4672.03	-9338.05
	(IV type)	0.00130	0.00907	-0.02115	1.93354		4672.23	-9336.46
SELL	(VII type)	0.00093	0.00810	0.00	1.72100	1.12247	4733.85	-9459.70
	(IV type)	0.00110	0.00810	-0.01437	1.72093	1.12244	4733.93	-9457.86
NLL	(VII type)	0.00109	0.01046	0.00	2.36577	1.00762	4723.11	-9438.22
	(IV type)	0.00163	0.01051	-0.03604	2.38138	1.00619	4723.69	-9437.38

Table 4.30: Comparison of S.D, SELL and NLL (TOPIX) (1990-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	-0.00070	0.01440	0.00	1.96979		3691.40	-7376.80
	(IV type)	-0.00120	0.01440	0.02347	1.97130		3691.66	-7375.32
SELL	(VII type)	-0.00059	0.01270	0.00	1.72272	1.14109	3733.86	-7459.72
	(IV type)	-0.00086	0.01270	0.01455	1.72296	1.14042	3733.94	-7457.88
NLL	(VII type)	-0.00031	0.01779	0.00	2.49840	0.94422	3703.63	-7399.26
	(IV type)	-0.00046	0.01778	0.00609	2.49699	0.94428	3703.65	-7397.29

Table 4.31: Comparison of S.D, SELL and NLL (SP500) (1975-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00041	0.01390	0.00	2.69815		16970.88	-33935.76
	(IV type)	0.00033	0.01391	0.00410	2.69998		16970.91	-33933.82
SELL	(VII type)	0.00040	0.01389	0.00	2.56937	0.92077	16972.14	-33936.28
	(IV type)	0.00040	0.01389	0.00009	2.56875	0.92080	16972.14	-33934.28
NLL	(VII type)	0.00042	0.02426	0.00	5.58025	0.65787	16868.28	-33728.55
	(IV type)	0.00082	0.02436	-0.01328	5.61711	0.65646	16868.40	-33726.80

Table 4.32: Comparison of S.D, SELL and NLL (SP500) (1975-1979)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00030	0.02405	0.00	6.56703		4382.97	-8759.95
	(IV type)	-0.00110	0.02454	0.04990	6.79835		4383.29	-8758.58
SELL	(VII type)	0.00026	0.02123	0.00	5.30824	0.56114	4406.54	-8805.08
	(IV type)	-0.00189	0.02141	0.08823	5.42218	0.55584	4407.39	-8804.79
NLL	(VII type)	0.00028	0.02633	0.00	7.81047	0.54176	4405.62	-8803.25
	(IV type)	-0.00297	0.02550	0.10760	7.54759	0.55550	4406.98	-8803.96

Table 4.33: Comparison of S.D, SELL and NLL (SP500) (1980-1984)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00019	0.02278	0.00	4.40873		4120.55	-8235.09
	(IV type)	-0.00328	0.02292	0.12381	4.51945		4123.37	-8238.74
SELL	(VII type)	0.00022	0.02475	0.00	4.77957	0.64507	4116.64	-8225.27
	(IV type)	-0.00413	0.02470	0.15339	4.86736	0.64019	4119.69	-8229.37
NLL	(VII type)	0.00035	0.04816	0.00	14.79109	0.42862	4110.19	-8212.37
	(IV type)	-0.00918	0.04472	0.19143	13.50001	0.44076	4112.71	-8215.43

Table 4.34: Comparison of S.D, SELL and NLL (SP500) (1985-1989)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00101	0.01038	0.00	1.86298		4102.91	-8199.82
	(IV type)	0.00145	0.01043	-0.02782	1.87138		4103.27	-8198.55
SELL	(VII type)	0.00114	0.01086	0.00	1.80873	1.20719	4085.78	-8163.56
	(IV type)	0.00218	0.01110	-0.06526	1.85269	1.18973	4087.33	-8164.67
NLL	(VII type)	0.00102	0.01939	0.00	3.47259	0.82814	4024.67	-8041.33
	(IV type)	0.00403	0.01992	-0.11123	3.65364	0.80951	4027.89	-8045.77

Table 4.35: Comparison of S.D, SELL and NLL (SP500) (1990-1994)

models		μ	τ	δ	b	σ	LL	AIC
S.D.	(VII type)	0.00024	0.01178	0.00	2.64030		4429.81	-8853.62
	(IV type)	0.00055	0.01180	-0.01874	2.64502		4429.93	-8851.86
SELL	(VII type)	0.00022	0.01096	0.00	2.34282	1.00644	4424.27	-8840.54
	(IV type)	0.00052	0.01093	-0.02001	2.33639	1.00827	4424.40	-8838.80
NLL	(VII type)	0.00024	0.02979	0.00	9.53724	0.51751	4387.06	-8766.13
	(IV type)	0.00340	0.02995	-0.09137	9.70002	0.51685	4388.06	-8766.12

4.5.3 Comparison of Maximum Likelihood Method by the Transition Probability Density Function and Local Linearization

In this subsection, a comparison of maximum likelihood method by the transition probability density function and local linearization is shown. The Table 4.36 - Table 4.52 are the comparison of empirical results. The $*$ is the best log-likelihood in each methods. T.P. stands for the transition probability density function method. The new local linearization method tends to over-estimate the level of b more than the transition probability density method and the simply extended local lenearization method. In the area of $K = 7$, the result of the new local linearization is close to the results of others. It suggests that, as for the new local linearization, in the area of low b, heavy-tailed, the new local linearization has the problem for the condition of applying the Girsanov theorem. Since the number of the assumed constants for the new local linearization is three, L_s, M_s and N_s, the number of constants is bigger than the simply extended local lenearization method which has two assumed constants, L_s and M_s . Therefore, by this empirical results, it is better to use the simply extended local linearization method in our case.

Table 4.36: Comparison of S.D, SELL and NLL (NK225) (1970-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00071	0.00728	1.5	1.81926	23259.72*	-46513.44
	SELL	0.00072	0.00730	1.5	1.33297	23509.98*	-47013.97
	NLL	0.00104	0.00748	1.5	1.25986	23158.34	-46310.68
K=2	T.P.	0.00062	0.01234	2.5	1.24933	23216.29	-46426.58
	SELL	0.00064	0.01207	2.5	0.97221	23426.95	-46847.90
	NLL	0.00086	0.01191	2.5	0.98983	23276.83*	-46547.66
K=3	T.P.	0.00057	0.01659	3.5	0.99371	23086.13	-46166.25
	SELL	0.00058	0.01615	3.5	0.80232	23278.64	-46551.27
	NLL	0.00073	0.01583	3.5	0.84837	23217.40	-46428.79
K=4	T.P.	0.00053	0.02030	4.5	0.84705	22973.96	-45941.91
	SELL	0.00055	0.01975	4.5	0.69932	23152.50	-46299.01
	NLL	0.00065	0.01936	4.5	0.75609	23132.68	-46259.35
K=5	T.P.	0.00050	0.02364	5.5	0.75012	22883.21	-45760.41
	SELL	0.00052	0.02301	5.5	0.62855	23050.09	-46094.18
	NLL	0.00059	0.02258	5.5	0.68901	23050.84	-46095.69
K=6	T.P.	0.00048	0.02669	6.5	0.68052	22809.88	-45613.76
	SELL	0.00050	0.02600	6.5	0.57615	22966.26	-45926.52
	NLL	0.00055	0.02556	6.5	0.63700	22977.61	-45949.22
K=7	T.P.	0.00046	0.02952	7.5	0.62682	22748.05	-45490.10
	SELL	0.00048	0.02879	7.5	0.53537	22896.52	-45787.05
	NLL	0.00052	0.02835	7.5	0.59494	22913.34	-45820.67

Table 4.37: Comparison of S.D, SELL and NLL (NK225) (1970-1974)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00113	0.00825	1.5	1.68441	4788.74	-9571.48
	SELL	0.00112	0.00846	1.5	1.27533	4825.90*	-9645.80
	NLL	0.00148	0.00854	1.5	1.22938	4764.31	-9522.62
K=2	T.P.	0.00097	0.01370	2.5	1.17504	4793.23*	-9580.46
	SELL	0.00099	0.01373	2.5	0.93492	4818.72	-9631.45
	NLL	0.00130	0.01346	2.5	0.96706	4792.77*	-9579.54
K=3	T.P.	0.00088	0.01823	3.5	0.93737	4773.13	-9540.26
	SELL	0.00093	0.01817	3.5	0.77264	4794.76	-9583.52
	NLL	0.00117	0.01776	3.5	0.82769	4784.34	-9562.67
K=4	T.P.	0.00083	0.02217	4.5	0.79846	4753.82	-9501.64
	SELL	0.00088	0.02207	4.5	0.67347	4773.35	-9540.70
	NLL	0.00106	0.02159	4.5	0.73550	4770.01	-9534.02
K=5	T.P.	0.00078	0.02571	5.5	0.70557	4737.46	-9468.92
	SELL	0.00084	0.02557	5.5	0.60490	4755.57	-9505.14
	NLL	0.00097	0.02508	5.5	0.66739	4755.61	-9505.22
K=6	T.P.	0.00075	0.02895	6.5	0.63837	4723.82	-9441.64
	SELL	0.00081	0.02879	6.5	0.55389	4740.80	-9475.60
	NLL	0.00091	0.02828	6.5	0.61384	4742.48	-9478.95
K=7	T.P.	0.00072	0.03196	7.5	0.58638	4712.26	-9418.53
	SELL	0.00078	0.03177	7.5	0.51403	4728.38	-9450.76
	NLL	0.00085	0.03127	7.5	0.57019	4730.82	-9455.65

Table 4.38: Comparison of S.D, SELL and NLL (NK225) (1975-1979)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00062	0.00532	1.5	1.47270	5379.50	-10753.00
	SELL	0.00055	0.00539	1.5	1.19010	5417.67	-10829.34
	NLL	0.00078	0.00542	1.5	1.17633	5370.23	-10734.46
K=2	T.P.	0.00056	0.00863	2.5	1.03991	5409.44	-10812.89
	SELL	0.00047	0.00868	2.5	0.86685	5431.34*	-10856.68
	NLL	0.00065	0.00842	2.5	0.92867	5412.46	-10818.92
K=3	T.P.	0.00052	0.01130	3.5	0.83288	5410.42*	-10814.83
	SELL	0.00044	0.01135	3.5	0.71594	5426.43	-10846.87
	NLL	0.00058	0.01093	3.5	0.79709	5418.80*	-10831.60
K=4	T.P.	0.00049	0.01358	4.5	0.71088	5407.48	-10808.96
	SELL	0.00043	0.01362	4.5	0.62459	5420.51	-10835.02
	NLL	0.00054	0.01322	4.5	0.71093	5417.86	-10829.73
K=5	T.P.	0.00048	0.01559	5.5	0.62908	5404.15	-10802.29
	SELL	0.00043	0.01562	5.5	0.56170	5415.34	-10824.69
	NLL	0.00051	0.01521	5.5	0.64742	5415.06	-10824.11
K=6	T.P.	0.00046	0.01740	6.5	0.56843	5401.23	-10796.46
	SELL	0.00043	0.01743	6.5	0.51501	5411.03	-10816.07
	NLL	0.00049	0.01701	6.5	0.59663	5411.86	-10817.71
K=7	T.P.	0.00046	0.01907	7.5	0.52405	5398.44	-10790.87
	SELL	0.00042	0.01909	7.5	0.47856	5407.44	-10808.87
	NLL	0.00047	0.01867	7.5	0.55432	5408.75	-10811.49

Table 4.39: Comparison of S.D, SELL and NLL (NK225) (1980-1984)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00065	0.00547	1.5	1.61547	5313.39	-10620.78
	SELL	0.00064	0.00564	1.5	1.23403	5341.89	-10677.78
	NLL	0.00095	0.00580	1.5	1.17924	5278.13	-10550.25
K=2	T.P.	0.00057	0.00890	2.5	1.14336	5339.70*	-10673.41
	SELL	0.00060	0.00909	2.5	0.90350	5353.23*	-10700.46
	NLL	0.00080	0.00902	2.5	0.93111	5323.09	-10640.17
K=3	T.P.	0.00054	0.01169	3.5	0.91662	5334.93	-10663.87
	SELL	0.00055	0.01189	3.5	0.74958	5346.11	-10686.22
	NLL	0.00069	0.01169	3.5	0.80621	5332.17	-10658.35
K=4	T.P.	0.00051	0.01409	4.5	0.78355	5327.01	-10648.02
	SELL	0.00052	0.01427	4.5	0.65675	5338.29	-10670.59
	NLL	0.00062	0.01401	4.5	0.72676	5332.62*	-10659.25
K=5	T.P.	0.00049	0.01622	5.5	0.69479	5319.53	-10633.07
	SELL	0.00050	0.01637	5.5	0.59295	5331.45	-10656.89
	NLL	0.00057	0.01607	5.5	0.66980	5330.30	-10654.59
K=6	T.P.	0.00048	0.01815	6.5	0.63207	5313.05	-10620.10
	SELL	0.00049	0.01827	6.5	0.54558	5325.61	-10645.22
	NLL	0.00054	0.01795	6.5	0.62599	5327.04	-10648.08
K=7	T.P.	0.00047	0.01993	7.5	0.58105	5307.40	-10608.80
	SELL	0.00048	0.02003	7.5	0.50860	5320.61	-10635.23
	NLL	0.00052	0.01968	7.5	0.59071	5323.55	-10641.10

Table 4.40: Comparison of S.D, SELL and NLL (NK225) (1985-1989)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00126	0.00682	1.5	1.69636	4710.57	-9415.14
	SELL	0.00122	0.00698	1.5	1.28104	4751.23	-9496.46
	NLL	0.00167	0.00727	1.5	1.20906	4685.41	-9364.82
K=2	T.P.	0.00122	0.01125	2.5	1.19616	4716.33*	-9426.65
	SELL	0.00119	0.01122	2.5	0.94443	4751.55*	-9497.10
	NLL	0.00148	0.01116	2.5	0.97046	4726.60	-9447.20
K=3	T.P.	0.00119	0.01493	3.5	0.96360	4697.16	-9388.32
	SELL	0.00116	0.01474	3.5	0.78577	4732.10	-9458.20
	NLL	0.00136	0.01449	3.5	0.84725	4727.02*	-9448.04
K=4	T.P.	0.00116	0.01814	4.5	0.82915	4678.24	-9350.47
	SELL	0.00114	0.01782	4.5	0.68924	4712.89	-9419.78
	NLL	0.00128	0.01745	4.5	0.76643	4717.45	-9428.90
K=5	T.P.	0.00114	0.02102	5.5	0.74008	4661.86	-9317.72
	SELL	0.00112	0.02060	5.5	0.62258	4695.96	-9385.93
	NLL	0.00123	0.02016	5.5	0.70692	4705.32	-9404.65
K=6	T.P.	0.00112	0.02367	6.5	0.67932	4647.84	-9289.67
	SELL	0.00111	0.02315	6.5	0.57294	4681.30	-9356.60
	NLL	0.00119	0.02267	6.5	0.66014	4693.01	-9380.02
K=7	T.P.	0.00111	0.02611	7.5	0.62726	4635.80	-9265.59
	SELL	0.00110	0.02553	7.5	0.53409	4668.55	-9331.11
	NLL	0.00116	0.02503	7.5	0.62175	4681.33	-9356.66

Table 4.41: Comparison of S.D, SELL and NLL (NK225) (1990-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	-0.00061	0.01350	1.5	2.03554	3421.98	-6837.96
	SELL	-0.00047	0.01376	1.5	1.34421	3452.69	-6899.38
	NLL	-0.00079	0.01508	1.5	1.19855	3360.86	-6715.72
K=2	T.P.	-0.00065	0.02261	2.5	1.40528	3438.85*	-6871.70
	SELL	-0.00058	0.02254	2.5	0.98217	3459.03*	-6912.07
	NLL	-0.00069	0.02310	2.5	0.94904	3405.36	-6804.71
K=3	T.P.	-0.00067	0.02997	3.5	1.12119	3432.57	-6859.14
	SELL	-0.00060	0.02973	3.5	0.81635	3449.12	-6892.24
	NLL	-0.00061	0.02990	3.5	0.82126	3414.99	-6823.98
K=4	T.P.	-0.00068	0.03624	4.5	0.95973	3425.14	-6844.28
	SELL	-0.00060	0.03591	4.5	0.71683	3439.37	-6872.73
	NLL	-0.00055	0.03587	4.5	0.73909	3415.92*	-6825.83
K=5	T.P.	-0.00068	0.04179	5.5	0.85294	3418.69	-6831.39
	SELL	-0.00058	0.04140	5.5	0.64862	3431.16	-6856.33
	NLL	-0.00051	0.04123	5.5	0.67997	3414.11	-6822.23
K=6	T.P.	-0.00068	0.04680	6.5	0.77453	3413.33	-6820.65
	SELL	-0.00057	0.04639	6.5	0.59806	3424.38	-6840.76
	NLL	-0.00048	0.04613	6.5	0.63451	3411.45	-6816.91
K=7	T.P.	-0.00068	0.05140	7.5	0.71612	3408.83	-6811.66
	SELL	-0.00056	0.05098	7.5	0.55857	3418.73	-6831.47
	NLL	-0.00046	0.05068	7.5	0.59799	3408.61	-6811.22

Table 4.42: Comparison of S.D, SELL and NLL (TOPIX) (1970-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00059	0.00624	1.5	1.56155	24435.43*	-48864.86
	SELL	0.00058	0.00636	1.5	1.24237	24645.15*	-49284.30
	NLL	0.00081	0.00630	1.5	1.23939	24462.27	-48918.54
K=2	T.P.	0.00053	0.01048	2.5	1.09030	24378.49	-48750.98
	SELL	0.00054	0.01044	2.5	0.90581	24549.30	-49092.61
	NLL	0.00072	0.01005	2.5	0.97760	24535.39*	-49064.78
K=3	T.P.	0.00050	0.01406	3.5	0.87067	24241.41	-48476.82
	SELL	0.00052	0.01390	3.5	0.74607	24398.28	-48790.56
	NLL	0.00065	0.01338	3.5	0.83559	24441.87	-48877.75
K=4	T.P.	0.00048	0.01720	4.5	0.74208	24122.57	-48239.13
	SELL	0.00050	0.01697	4.5	0.64871	24269.34	-48532.68
	NLL	0.00059	0.01638	4.5	0.73906	24329.81	-48653.62
K=5	T.P.	0.00046	0.02004	5.5	0.65608	24025.38	-48044.76
	SELL	0.00049	0.01974	5.5	0.58155	24163.57	-48321.14
	NLL	0.00055	0.01913	5.5	0.66579	24225.93	-48445.86
K=6	T.P.	0.00045	0.02265	6.5	0.59393	23946.60	-47887.19
	SELL	0.00048	0.02230	6.5	0.53165	24076.14	-48146.29
	NLL	0.00053	0.02169	6.5	0.60713	24135.02	-48264.03
K=7	T.P.	0.00043	0.02539	7.5	0.50387	23865.39	-47724.78
	SELL	0.00046	0.02469	7.5	0.49277	24002.81	-47999.61
	NLL	0.00051	0.02408	7.5	0.55909	24056.62	-48107.25

Table 4.43: Comparison of S.D, SELL and NLL (TOPIX) (1970-1974)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00112	0.00642	1.5	1.48838	5163.69*	-10321.37
	SELL	0.00108	0.00677	1.5	1.19024	5189.08*	-10372.16
	NLL	0.00136	0.00661	1.5	1.21254	5161.21	-10316.42
K=2	T.P.	0.00095	0.01065	2.5	1.05014	5158.58	-10311.17
	SELL	0.00094	0.01094	2.5	0.87316	5178.26	-10350.52
	NLL	0.00114	0.01047	2.5	0.96041	5180.30*	-10354.60
K=3	T.P.	0.00087	0.01422	3.5	0.84111	5134.23	-10262.47
	SELL	0.00087	0.01441	3.5	0.72190	5154.53	-10303.05
	NLL	0.00100	0.01381	3.5	0.82407	5166.34	-10326.68
K=4	T.P.	0.00081	0.01735	4.5	0.71708	5112.41	-10218.82
	SELL	0.00082	0.01744	4.5	0.62931	5133.25	-10260.50
	NLL	0.00090	0.01678	4.5	0.73018	5147.43	-10288.87
K=5	T.P.	0.00076	0.02017	5.5	0.63366	5094.37	-10182.75
	SELL	0.00078	0.02016	5.5	0.56507	5115.25	-10224.49
	NLL	0.00083	0.01947	5.5	0.65672	5129.00	-10251.99
K=6	T.P.	0.00073	0.02275	6.5	0.57393	5079.51	-10153.01
	SELL	0.00075	0.02267	6.5	0.51706	5099.99	-10193.99
	NLL	0.00078	0.02197	6.5	0.59620	5112.38	-10218.75
K=7	T.P.	0.00070	0.02515	7.5	0.52601	5067.06	-10128.11
	SELL	0.00072	0.02500	7.5	0.47935	5086.95	-10167.91
	NLL	0.00075	0.02432	7.5	0.54621	5097.80	-10189.60

Table 4.44: Comparison of S.D, SELL and NLL (TOPIX) (1975-1979)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00055	0.00445	1.5	1.33340	5686.27	-11366.54
	SELL	0.00053	0.00458	1.5	1.11816	5713.76	-11421.52
	NLL	0.00073	0.00458	1.5	1.12272	5675.42	-11344.83
K=2	T.P.	0.00049	0.00709	2.5	0.95121	5720.01	-11434.03
	SELL	0.00045	0.00725	2.5	0.81594	5733.10*	-11460.20
	NLL	0.00058	0.00703	2.5	0.88659	5718.92	-11431.85
K=3	T.P.	0.00045	0.00924	3.5	0.76458	5723.87*	-11441.73
	SELL	0.00043	0.00940	3.5	0.67342	5732.72	-11459.43
	NLL	0.00051	0.00908	3.5	0.75913	5727.92	-11449.84
K=4	T.P.	0.00043	0.01107	4.5	0.65368	5723.01	-11440.02
	SELL	0.00041	0.01122	4.5	0.58704	5730.03	-11454.05
	NLL	0.00048	0.01087	4.5	0.67272	5729.03*	-11452.05
K=5	T.P.	0.00042	0.01268	5.5	0.57915	5721.25	-11436.51
	SELL	0.00041	0.01283	5.5	0.52752	5727.26	-11448.52
	NLL	0.00045	0.01246	5.5	0.60623	5727.80	-11449.59
K=6	T.P.	0.00041	0.01414	6.5	0.52431	5719.63	-11433.26
	SELL	0.00040	0.01428	6.5	0.48328	5724.80	-11443.60
	NLL	0.00044	0.01390	6.5	0.55220	5725.91	-11445.82
K=7	T.P.	0.00040	0.01548	7.5	0.48295	5717.81	-11429.62
	SELL	0.00040	0.01561	7.5	0.44872	5722.68	-11439.36
	NLL	0.00042	0.01524	7.5	0.50780	5723.96	-11441.92

Table 4.45: Comparison of S.D, SELL and NLL (TOPIX) (1980-1984)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00054	0.00492	1.5	1.41833	5496.29	-10986.58
	SELL	0.00057	0.00503	1.5	1.16544	5526.83	-11047.66
	NLL	0.00083	0.00505	1.5	1.16285	5486.35	-10966.70
K=2	T.P.	0.00053	0.00793	2.5	1.00799	5517.82*	-11029.63
	SELL	0.00057	0.00809	2.5	0.84801	5532.96*	-11059.92
	NLL	0.00075	0.00784	2.5	0.92147	5522.68	-11039.37
K=3	T.P.	0.00052	0.01041	3.5	0.80936	5511.43	-11016.87
	SELL	0.00056	0.01060	3.5	0.69929	5523.23	-11040.45
	NLL	0.00068	0.01023	3.5	0.79163	5523.97*	-11041.95
K=4	T.P.	0.00052	0.01256	4.5	0.69201	5502.69	-10999.37
	SELL	0.00055	0.01274	4.5	0.60951	5513.78	-11021.56
	NLL	0.00063	0.01233	4.5	0.70509	5518.69	-11031.37
K=5	T.P.	0.00051	0.01447	5.5	0.61310	5494.73	-10983.46
	SELL	0.00054	0.01463	5.5	0.54790	5505.86	-11005.72
	NLL	0.00060	0.01420	5.5	0.63988	5512.23	-11018.45
K=6	T.P.	0.00051	0.01619	6.5	0.55298	5488.15	-10970.29
	SELL	0.00053	0.01635	6.5	0.50219	5499.30	-10992.59
	NLL	0.00057	0.01591	6.5	0.58704	5505.98	-11005.95
K=7	T.P.	0.00051	0.01781	7.5	0.51192	5482.11	-10958.22
	SELL	0.00053	0.01792	7.5	0.46658	5493.80	-10981.60
	NLL	0.00056	0.01749	7.5	0.54233	5500.30	-10994.61

Table 4.46: Comparison of S.D, SELL and NLL (TOPIX) (1985-1989)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00102	0.00695	1.5	1.52129	4688.47*	-9370.94
	SELL	0.00092	0.00707	1.5	1.22215	4731.27*	-9456.55
	NLL	0.00105	0.00702	1.5	1.22545	4697.86	-9389.73
K=2	T.P.	0.00102	0.01147	2.5	1.08440	4686.06	-9366.13
	SELL	0.00095	0.01140	2.5	0.89988	4720.35	-9434.70
	NLL	0.00110	0.01096	2.5	0.98450	4722.80*	-9439.60
K=3	T.P.	0.00102	0.01526	3.5	0.87677	4663.03	-9320.05
	SELL	0.00097	0.01508	3.5	0.74587	4694.01	-9382.02
	NLL	0.00110	0.01446	3.5	0.85451	4709.50	-9412.99
K=4	T.P.	0.00101	0.01858	4.5	0.75470	4641.52	-9277.04
	SELL	0.00099	0.01834	4.5	0.65194	4670.49	-9334.98
	NLL	0.00109	0.01764	4.5	0.76723	4690.17	-9374.33
K=5	T.P.	0.00100	0.02157	5.5	0.67310	4623.24	-9240.48
	SELL	0.00099	0.02128	5.5	0.58717	4650.83	-9295.67
	NLL	0.00108	0.02056	5.5	0.70212	4671.32	-9336.63
K=6	T.P.	0.00099	0.02431	6.5	0.61322	4607.99	-9209.82
	SELL	0.00100	0.02399	6.5	0.53912	4634.39	-9262.78
	NLL	0.00106	0.02327	6.5	0.65029	4654.36	-9302.71
K=7	T.P.	0.00099	0.02686	7.5	0.56904	4594.52	-9183.04
	SELL	0.00100	0.02652	7.5	0.50166	4620.48	-9234.95
	NLL	0.00105	0.02581	7.5	0.60693	4639.42	-9272.83

Table 4.47: Comparison of S.D, SELL and NLL (TOPIX) (1990-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	-0.00076	0.01077	1.5	1.60129	3700.29	-7394.57
	SELL	-0.00061	0.01100	1.5	1.24973	3731.74*	-7457.49
	NLL	-0.00042	0.01125	1.5	1.20652	3681.25	-7356.50
K=2	T.P.	-0.00070	0.01803	2.5	1.10309	3703.28*	-7400.56
	SELL	-0.00056	0.01807	2.5	0.90433	3723.66	-7441.31
	NLL	-0.00031	0.01780	2.5	0.94401	3703.63*	-7401.26
K=3	T.P.	-0.00068	0.02404	3.5	0.87593	3688.98	-7371.97
	SELL	-0.00054	0.02402	3.5	0.74229	3704.66	-7403.32
	NLL	-0.00032	0.02351	3.5	0.80458	3697.37	-7388.75
K=4	T.P.	-0.00066	0.02924	4.5	0.74445	3675.87	-7345.75
	SELL	-0.00052	0.02920	4.5	0.64442	3689.07	-7372.14
	NLL	-0.00035	0.02858	4.5	0.71238	3687.00	-7367.99
K=5	T.P.	-0.00065	0.03390	5.5	0.65710	3665.16	-7324.33
	SELL	-0.00051	0.03384	5.5	0.57737	3676.52	-7347.04
	NLL	-0.00037	0.03316	5.5	0.64431	3676.89	-7347.78
K=6	T.P.	-0.00064	0.03812	6.5	0.59348	3656.54	-7307.08
	SELL	-0.00050	0.03807	6.5	0.52781	3666.47	-7326.94
	NLL	-0.00038	0.03736	6.5	0.59084	3667.94	-7329.89
K=7	T.P.	-0.00064	0.04203	7.5	0.54563	3649.34	-7292.68
	SELL	-0.00049	0.04198	7.5	0.48929	3658.29	-7310.57
	NLL	-0.00040	0.04125	7.5	0.54727	3660.23	-7314.46

Table 4.48: Comparison of S.D, SELL and NLL (SP500) (1975-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00040	0.00794	1.5	1.81676	16876.15	-33746.29
	SELL	0.00039	0.00853	1.5	1.26328	16906.41	-33806.82
	NLL	0.00033	0.00911	1.5	1.15273	16600.82	-33195.65
K=2	T.P.	0.00041	0.01302	2.5	1.27592	16982.50*	-33959.01
	SELL	0.00040	0.01359	2.5	0.93477	16972.03*	-33938.06
	NLL	0.00037	0.01382	2.5	0.91653	16788.33	-33570.65
K=3	T.P.	0.00041	0.01710	3.5	1.01967	16972.93	-33939.85
	SELL	0.00041	0.01764	3.5	0.78093	16961.11	-33916.21
	NLL	0.00040	0.01769	3.5	0.79599	16844.61	-33683.22
K=4	T.P.	0.00042	0.02058	4.5	0.87124	16949.36	-33892.73
	SELL	0.00041	0.02109	4.5	0.68722	16941.34	-33876.68
	NLL	0.00041	0.02105	4.5	0.71835	16863.66	-33721.32
K=5	T.P.	0.00042	0.02365	5.5	0.77245	16925.50	-33845.00
	SELL	0.00042	0.02413	5.5	0.62231	16921.81	-33837.62
	NLL	0.00042	0.02403	5.5	0.66214	16868.25*	-33730.50
K=6	T.P.	0.00042	0.02644	6.5	0.69865	16904.13	-33802.27
	SELL	0.00042	0.02689	6.5	0.57381	16904.06	-33802.12
	NLL	0.00042	0.02675	6.5	0.61853	16866.37	-33726.74
K=7	T.P.	0.00043	0.02901	7.5	0.64984	16883.94	-33761.88
	SELL	0.00042	0.02943	7.5	0.53570	16888.18	-33770.36
	NLL	0.00042	0.02926	7.5	0.58308	16861.37	-33716.75

Table 4.49: Comparison of S.D, SELL and NLL (SP500) (1975-1979)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00036	0.00788	1.5	1.37360	4340.60	-8675.19
	SELL	0.00026	0.00821	1.5	1.12694	4363.39	-8720.77
	NLL	-0.00004	0.00853	1.5	1.08148	4313.52	-8621.04
K=2	T.P.	0.00030	0.01228	2.5	0.99520	4385.34	-8764.69
	SELL	0.00022	0.01261	2.5	0.83583	4396.91	-8787.82
	NLL	0.00005	0.01251	2.5	0.87295	4371.49	-8736.99
K=3	T.P.	0.00028	0.01583	3.5	0.80426	4395.58	-8785.16
	SELL	0.00023	0.01612	3.5	0.69619	4404.30	-8802.59
	NLL	0.00013	0.01582	3.5	0.76144	4391.33	-8776.66
K=4	T.P.	0.00028	0.01884	4.5	0.69000	4398.71	-8791.41
	SELL	0.00025	0.01909	4.5	0.61051	4406.25	-8806.50
	NLL	0.00019	0.01868	4.5	0.68803	4399.73	-8793.46
K=5	T.P.	0.00028	0.02150	5.5	0.61278	4399.65	-8793.30
	SELL	0.00026	0.02171	5.5	0.55101	4406.53*	-8807.05
	NLL	0.00022	0.02122	5.5	0.63363	4403.50	-8801.00
K=6	T.P.	0.00028	0.02389	6.5	0.55402	4400.05*	-8794.10
	SELL	0.00027	0.02407	6.5	0.50650	4406.19	-8806.38
	NLL	0.00025	0.02355	6.5	0.58992	4405.10	-8804.20
K=7	T.P.	0.00028	0.02609	7.5	0.51313	4399.54	-8793.09
	SELL	0.00028	0.02624	7.5	0.47153	4405.63	-8805.26
	NLL	0.00026	0.02569	7.5	0.55249	4405.60*	-8805.20

Table 4.50: Comparison of S.D, SELL and NLL (SP500) (1980-1984)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	-0.00005	0.00927	1.5	1.68928	4072.40	-8138.80
	SELL	-0.00005	0.00996	1.5	1.22020	4080.81	-8155.62
	NLL	-0.00017	0.01074	1.5	1.10617	4007.02	-8008.05
K=2	T.P.	0.00006	0.01486	2.5	1.20058	4114.12	-8222.23
	SELL	0.00011	0.01558	2.5	0.90495	4109.97	-8213.95
	NLL	0.00005	0.01592	2.5	0.88415	4063.69	-8121.38
K=3	T.P.	0.00012	0.01933	3.5	0.95910	4121.18	-8236.36
	SELL	0.00017	0.02002	3.5	0.75587	4115.51	-8225.01
	NLL	0.00016	0.02015	3.5	0.76898	4084.95	-8163.91
K=4	T.P.	0.00016	0.02311	4.5	0.81796	4122.13*	-8238.26
	SELL	0.00022	0.02378	4.5	0.66482	4116.60*	-8227.21
	NLL	0.00022	0.02378	4.5	0.69464	4095.32	-8184.64
K=5	T.P.	0.00019	0.02644	5.5	0.72392	4121.54	-8237.08
	SELL	0.00024	0.02709	5.5	0.60176	4116.48	-8226.96
	NLL	0.00026	0.02700	5.5	0.64100	4101.05	-8196.10
K=6	T.P.	0.00021	0.02945	6.5	0.65511	4120.56	-8235.13
	SELL	0.00026	0.03007	6.5	0.55469	4115.98	-8225.96
	NLL	0.00029	0.02993	6.5	0.59968	4104.47	-8202.94
K=7	T.P.	0.00022	0.03221	7.5	0.60358	4119.48	-8232.97
	SELL	0.00028	0.03280	7.5	0.51776	4115.38	-8224.75
	NLL	0.00030	0.03263	7.5	0.56640	4106.60*	-8207.21

Table 4.51: Comparison of S.D, SELL and NLL (SP500) (1985-1989)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00102	0.00815	1.5	2.23198	4093.98*	-8181.95
	SELL	0.00115	0.00896	1.5	1.35443	4082.51*	-8159.03
	NLL	0.00124	0.00969	1.5	1.20208	3987.68	-7969.37
K=2	T.P.	0.00104	0.01383	2.5	1.49805	4093.07	-8180.13
	SELL	0.00110	0.01462	2.5	1.00299	4079.20	-8152.40
	NLL	0.00110	0.01506	2.5	0.95156	4019.92	-8033.83
K=3	T.P.	0.00103	0.01847	3.5	1.18130	4072.62	-8139.24
	SELL	0.00104	0.01923	3.5	0.83850	4063.65	-8121.30
	NLL	0.00101	0.01950	3.5	0.82535	4024.66*	-8043.33
K=4	T.P.	0.00102	0.02249	4.5	1.00214	4053.45	-8100.91
	SELL	0.00099	0.02319	4.5	0.73764	4049.05	-8092.10
	NLL	0.00096	0.02336	4.5	0.74330	4022.13	-8038.26
K=5	T.P.	0.00101	0.02609	5.5	0.88396	4036.97	-8067.94
	SELL	0.00096	0.02672	5.5	0.66723	4036.29	-8066.59
	NLL	0.00092	0.02683	5.5	0.68323	4017.08	-8028.16
K=6	T.P.	0.00100	0.02937	6.5	0.80370	4022.81	-8039.63
	SELL	0.00093	0.02994	6.5	0.61431	4025.14	-8044.29
	NLL	0.00090	0.03001	6.5	0.63629	4011.14	-8016.28
K=7	T.P.	0.00098	0.03241	7.5	0.74302	4010.57	-8015.13
	SELL	0.00091	0.03291	7.5	0.57256	4015.29	-8024.57
	NLL	0.00088	0.03296	7.5	0.59802	4004.95	-8003.90

Table 4.52: Comparison of S.D, SELL and NLL (SP500) (1990-1994)

models		μ	τ	b	σ	LL	AIC
K=1	T.P.	0.00026	0.00672	1.5	2.08554	4402.90	-8799.80
	SELL	0.00024	0.00715	1.5	1.32413	4411.74	-8817.48
	NLL	0.00028	0.00782	1.5	1.17063	4315.66	-8625.32
K=2	T.P.	0.00026	0.01125	2.5	1.40256	4426.27*	-8846.55
	SELL	0.00022	0.01159	2.5	0.96962	4424.09*	-8842.18
	NLL	0.00031	0.01202	2.5	0.91679	4361.16	-8716.31
K=3	T.P.	0.00024	0.01485	3.5	1.10224	4424.01	-8842.02
	SELL	0.00022	0.01519	3.5	0.80516	4419.51	-8833.01
	NLL	0.00030	0.01549	3.5	0.78834	4375.52	-8745.05
K=4	T.P.	0.00024	0.01790	4.5	0.93495	4419.32	-8832.64
	SELL	0.00022	0.01826	4.5	0.70586	4414.13	-8822.27
	NLL	0.00028	0.01848	4.5	0.70702	4381.58	-8757.17
K=5	T.P.	0.00023	0.02060	5.5	0.82528	4414.90	-8823.81
	SELL	0.00022	0.02096	5.5	0.63776	4409.59	-8813.18
	NLL	0.00028	0.02114	5.5	0.64937	4384.49	-8762.97
K=6	T.P.	0.00023	0.02303	6.5	0.74633	4411.13	-8816.26
	SELL	0.00022	0.02340	6.5	0.58738	4405.91	-8805.82
	NLL	0.00027	0.02354	6.5	0.60559	4385.94	-8765.88
K=7	T.P.	0.00023	0.02526	7.5	0.68676	4407.94	-8809.88
	SELL	0.00022	0.02564	7.5	0.54816	4402.92	-8799.84
	NLL	0.00026	0.02574	7.5	0.57086	4386.66*	-8767.32

Chapter 5

Non-Gaussian Stochastic Volatility Model

5.1 Introduction

In the area of finance, the management of the market risk is a very important subject. In this area, ARCH and GARCH models have been used for their easy implementation. On the other hand, the stochastic volatility model was proposed in the finance area by Hull and White (1987) and a method of estimating the parameters of the model was recently developed (Harvey, Ruiz, and Shephard (1992), etc.). The empirical comparative study of GARCH and stochastic volatility model is reported by Heyman et al. (1994). According to their result, for stock returns, the stochastic volatility model is better than GARCH and EGARCH.

In this chapter, in order to analyze the stochastic volatility of the process for the risk management, we consider a non-Gaussian extension of a simple version of a stochastic volatility model. Namely we assume that the observation model has a heavy tailed distribution and the expectation of observation is zero. According to the standard VaR (Value at Risk) model, so-called Risk Metrics by J. P. Morgan (1995a, 1995b) in the 1 day forecast, estimates from the zero-mean and estimates from estimated mean do not differ significantly. Therefore, to measure the VaR, zero-mean estimates are commonly used in this area. Our model can also be introduced in this framework.

Recently, several methods for estimating the parameters of stochastic volatility models have been developed. Firstly, quasi-maximum likelihood method by Harvey, Ruiz, and Shephard (1992) is used because of its easy implementation. They transform the observation equation to linear state space model by multiplying the observation and taking logarithm of it. They approximate the distribution of noise by normal distribution and use the Kalman Filter to maximize a quasi-maximum likelihood. Secondly, Danielsson (1992), Danielsson and Richard (1992) approximate the likelihood using Monte Carlo integration. They proposed an efficient method called accelerated Gaussian importance sampler. Thirdly, Jacquire, Polson, and Rossi (1994) adopt a Bayesian analysis using MCMC (Markov Chain Monte Carlo) method which is based on sampling from the posterior of the model parameters and then from the posterior for the latent variables.

In this chapter, we apply a non-Gaussian nonlinear filter (Kitagawa 1987, 1991) to stochastic volatility model. The method is based on a numerical representation of the arbitrary distributions. By the computer-intensive numerical computations, this method can handle various types of nonlinear or non-Gaussian models including the non-Gaussian stochastic volatility model.

This research focuses on the type VII Pearson family of distributions, including t-distribution, Cauchy distribution and Gaussian distribution. However, the computational method presented in this chapter can be easily applied to other distributions. The research by Nagahara (1995a, 1995b) and others (Mandelbrot (1963), Fama (1965), Kariya, (1993), d Kariya et al. (1995) etc.) for the daily returns of stock index prices conclude that distribution of daily stock returns is heavy-tailed and is better approximated by Pearson type VII or Paretian distribution than the Gaussian distribution. The t-distribution was considered by Harvey, Ruiz, and Shephard (1992). Furthermore, according to Nagahara (1995b), the shape parameter of Pearson type VII distribution tends to be time-varying. These motivated us to develop general state space model of stochastic volatility model and adopt the time-varying shape parameter. We estimate these time-varying parameters by the non-Gaussian smoother developed by Kitagawa (1987, 1991). The outline of this chapter is as follows. Non-Gaussian stochastic volatility models are introduced in section 5.2. Non-Gaussian filter and smoother for the estimation of the unknown parameters of the model such as the time varying volatility and the shape parameter are also given in this subsection. The case study on the analysis of the daily returns of Standard and Poor's 500 Index is shown in section 5.3.

5.2 Stochastic Volatility Models with Gaussian-Noise and Non-Gaussian Noise

5.2.1 Stochastic Volatility Model

The stochastic volatility model with Gaussian noise is defined by

$$y_t \;=\; \sigma_t \epsilon_t, \tag{5.1}$$

where $\epsilon_t \sim NID(0,1)$ and σ_t is the stochastic volatility and its logarithm $h_t = \log \sigma_t^2$ is assumed to be generated by a first order autoregressive model with a constant term

$$h_t \;=\; \alpha_0 + \alpha_1 h_{t-1} + u_t, \tag{5.2}$$

with $u_t \sim NID(0, \mu^2)$ (Harvey, Ruiz, and Shephard (1992)).

In this study, we consider an extension of this stochastic volatility model to non-Gaussian distribution case. As the distribution of the observation noise, we use the Pearson type VII density function defined by

$$p(x|b,\mu,\tau^2) = \frac{\Gamma(b)}{\Gamma(b-1/2)\Gamma(1/2)} \frac{\tau^{2b-1}}{\{\tau^2 + (x-\mu)^2\}^b}, \quad (b > 1/2). \tag{5.3}$$

The details of the distribution is given in Johnson and Kotz (1970). This distribution contains broad class of distributions such as Cauchy distribution ($b = 1$), t-distribution with the degree of freedom k ($b = (k+1)/2$, k: positive integer) and the normal distribution ($b = \infty$).

Based on this Pearson type VII distribution, our non-Gaussian extension of the stochastic volatility model is defined as follows. In this model we assume that the observation y_t is distributed as type VII Pearson family of distributions with mean 0, the shape parameter b and the dispersion parameter τ_t^2,

$$y_t \;\sim\; \frac{\Gamma(b)}{\Gamma(b-1/2)\Gamma(1/2)} \frac{\tau_t^{2b-1}}{(\tau_t^2 + x^2)^b}. \tag{5.4}$$

Here the dispersion of the distribution, τ_t^2, is time-varying and its logarithm $h_t = \log \tau_t^2$ follows an autoregressive model with a constant term,

$$h_t \;=\; \alpha_0 + \alpha_1 h_{t-1} + v_t, \tag{5.5}$$

where v_t is a white noise sequence with $v_t \sim NID(0, \nu^2)$.

This non-Gaussian stochastic volatility model can be further generalized to the case where the shape parameter, b, is also time-varying. In this model, y_t is assumed to be distributed as type VII Pearson family of distributions with the shape parameter b_t and the dispersion parameter τ_t^2,

$$y_t \sim \frac{\Gamma(b_t)}{\Gamma(b_t - 1/2)\Gamma(1/2)} \frac{\tau_t^{2b_t - 1}}{(\tau_t^2 + x^2)^{b_t}}. \tag{5.6}$$

Here both of the shape parameter and the dispersion parameter are time varying, and as above, $h_t = \log \tau_t^2$ follows

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + v_{1,t}, \tag{5.7}$$

with $v_{1,t} \sim NID(0, \xi_1^2)$, and $l_t = \log(b_t - 1/2)$ follows the random walk model

$$l_t = l_{t-1} + v_{2,t}, \tag{5.8}$$

with $v_{2,t} \sim NID(0, \xi_2^2)$. The above transformation, $l_t = \log(b_t - 1/2)$, is used so that the $b_t = e^{l_t} + 1/2$ always satisfies the basic condition for the shape parameter of the type VII Pearson system that $b_t > 1/2$.

5.2.2 General State Space Model

In this subsection, we first derive a general nonlinear non-Gaussian state space representation of the non-Gaussian stochastic volatility model introduced in the previous subsection. This general state space model facilitates to compute the log-likelihood and obtains the maximum likelihood estimates of the non-Gaussian stochastic volatility model. General state space model is defined by

$$\mathbf{x_t} = f(\mathbf{x_{t-1}}, \mathbf{v_t}) \tag{5.9}$$

$$\mathbf{y_t} \sim m(\cdot|\mathbf{x_t}), \tag{5.10}$$

where $\mathbf{y_t}$ is the observation, $\mathbf{x_t}$ is k-dimensional state vector and $\mathbf{v_t}$ is l-dimensional system noise with densities $q(v)$. The $f(\mathbf{x}, \mathbf{v})$ is a possibly nonlinear function of $\mathbf{x}$ and $\mathbf{v}$, and $m(\cdot|\mathbf{x_t})$ indicates the conditional distribution of the observation $\mathbf{y_t}$ given the state $\mathbf{x_t}$. (5.9) and (5.10) are called the state model (or system model) and the observation model, respectively. Our non-Gaussian stochastic volatility model given in the previous section can be expressed in general state space model form by

$$f(\mathbf{x_{t-1}}, \mathbf{v_t}) = \mathbf{f_0} + \mathbf{F}\mathbf{x_{t-1}} + \mathbf{G}\mathbf{v_t}, \tag{5.11}$$

where the state vector $\mathbf{x_t}$, the system noise $\mathbf{v_t}$ and $\mathbf{f_0}$, $\mathbf{F}$, $\mathbf{G}$ are respectively defined by

$$\mathbf{x_t} = \begin{bmatrix} h_t \\ l_t \end{bmatrix}, \quad \mathbf{v_t} = \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}, \quad \mathbf{f_0} = \begin{bmatrix} \alpha_0 \\ 0 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.12}$$

Namely, the state model (5.10) is given as

$$\begin{bmatrix} h_t \\ l_t \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t-1} \\ l_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}. \tag{5.13}$$

On the other hand, the observation model (5.10) is given by

$$m(y|\mathbf{x_t}) = \frac{\Gamma(b_t)}{\Gamma(b_t - 1/2)\Gamma(1/2)} \frac{\tau_t^{2b_t - 1}}{\{\tau_t^2 + y^2\}^{b_t}} = \frac{\Gamma(e^{l_t} + \frac{1}{2})}{\Gamma(e^{l_t})\Gamma(1/2)} \frac{\exp(h_t e^{l_t})}{\{e^{h_t} + y^2\}^{e^{l_t} + \frac{1}{2}}}, \tag{5.14}$$

with

$$\begin{aligned} \tau_t^2 &= \exp\{[1\ 0]\mathbf{x_t}\}, \\ b_t &= \exp\{[0\ 1]\mathbf{x_t}\} + \frac{1}{2}. \end{aligned} \tag{5.15}$$

Based on this state space representation of the non-Gaussian stochastic volatility model, we shall obtain the likelihood of the model and the maximum likelihood estimate of the parameter in the following subsection.

5.2.3 Recursive Estimation of State and Likelihood Computation

In this subsection, we consider the estimation of the state of the general state space model given in (5.9) and (5.10). The estimated state is then used for the computation of the likelihood of the stochastic volatility model. Finally, by the state estimation for the model with the maximum likelihood estimates of the parameters, we can estimate the volatility τ_t at each time point as well as the shape parameter, b_t.

Hereafter, $\mathbf{Y_s}$ is defined to be the set of the observations up to time s, $\mathbf{Y_s} \equiv \{\mathbf{y_1}, ..., \mathbf{y_s}\}$. Then the problem of the state estimation is formulated as to obtain the conditional distribution $p(\mathbf{x_t}|\mathbf{Y_s})$ of the state $\mathbf{x_t}$ given the set of observations $\mathbf{Y_s}$. Corresponding to the three distinct situations, $s < t$, $s = t$ and $s > t$, the conditional density $p(\mathbf{x_t}|\mathbf{Y_s})$ is called the predictor, the filter and the smoother, respectively.

In Kitagawa (1987, 1991), it was shown that, similarly to the Kalman filter and the fixed interval smoother for ordinary state space model (Anderson and Moore (1979)), the recursive formulas for obtaining one step ahead predictor $p(\mathbf{x_t}|\mathbf{Y_{t-1}})$, and the filter $p(\mathbf{x_t}|\mathbf{Y_t})$ can be given as follows.

One step ahead prediction:

$$\begin{aligned}
p(\mathbf{x_t}|\mathbf{Y_{t-1}}) &= \int_{-\infty}^{\infty} p(\mathbf{x_t}, \mathbf{x_{t-1}}|\mathbf{Y_{t-1}})\mathbf{dx_{t-1}} \\
&= \int_{-\infty}^{\infty} p(\mathbf{x_t}|\mathbf{x_{t-1}}, \mathbf{Y_{t-1}})p(\mathbf{x_{t-1}}|\mathbf{Y_{t-1}})\mathbf{dx_{t-1}} \\
&= \int_{-\infty}^{\infty} p(\mathbf{x_t}|\mathbf{x_{t-1}})p(\mathbf{x_{t-1}}|\mathbf{Y_{t-1}})\mathbf{dx_{t-1}}. \qquad (5.16)
\end{aligned}$$

Filtering:

$$\begin{aligned}
p(\mathbf{x_t}|\mathbf{Y_t}) &= p(\mathbf{x_t}|y_t, \mathbf{Y_{t-1}}) \\
&= \frac{p(y_t|\mathbf{x_t}, \mathbf{Y_{t-1}})p(\mathbf{x_t}|\mathbf{Y_{t-1}})}{p(y_t|\mathbf{Y_{t-1}})} \\
&= \frac{p(y_t|\mathbf{x_t})p(\mathbf{x_t}|\mathbf{Y_{t-1}})}{p(y_t|\mathbf{Y_{t-1}})}, \qquad (5.17)
\end{aligned}$$

where $p(y_t|\mathbf{Y_{t-1}})$ is obtained by $\int p(y_t|\mathbf{x_t})p(\mathbf{x_t}|\mathbf{Y_{t-1}})\mathbf{dx_t}$.

Smoothing:

Using the results of non-Gaussian filter, the smoothed density $p(\mathbf{x_t}|\mathbf{Y_T})$ is obtained by

$$\begin{aligned}
p(\mathbf{x_t}|\mathbf{Y_T}) &= \int_{-\infty}^{\infty} p(\mathbf{x_t}, \mathbf{x_{t+1}}|\mathbf{Y_T})\mathbf{dx_{t+1}} \\
&= \int_{-\infty}^{\infty} p(\mathbf{x_{t+1}}|\mathbf{Y_T})p(\mathbf{x_t}|\mathbf{x_{t+1}}, \mathbf{Y_T})\mathbf{dx_{t+1}} \\
&= \int_{-\infty}^{\infty} p(\mathbf{x_{t+1}}|\mathbf{Y_T})p(\mathbf{x_t}|\mathbf{x_{t+1}}, \mathbf{Y_t})\mathbf{dx_{t+1}} \\
&= p(\mathbf{x_t}|\mathbf{Y_t}) \int_{-\infty}^{\infty} \frac{p(\mathbf{x_{t+1}}|\mathbf{Y_T})p(\mathbf{x_{t+1}}|\mathbf{x_t}, \mathbf{Y_t})}{p(\mathbf{x_{t+1}}|\mathbf{Y_t})}\mathbf{dx_{t+1}}. \qquad (5.18)
\end{aligned}$$

For the present general state space models, the conditional density $p(\mathbf{x_t}|\mathbf{Y_s})$ becomes non-Gaussian and cannot be specified by using mean and variance. In the next subsection, we will present a numerical method for handling non-Gaussian state densities (Kitagawa 1991).

In our non-Gaussian stochastic volatility model, the coefficients of the autoregressive model (5.7), α_0 and α_1, and the variances of the system noise, ξ_1^2 and ξ_2^2, are unknown parameters. The vector of the parameter θ is defined by $\theta = (\alpha_0, \alpha_1, \xi_1^2, \xi_2^2)^t$. Then the log-likelihood function is obtained by

$$
\begin{aligned}
l(\theta) &= \log p(y_1, \ldots, y_T) \\
&= \sum_{t=1}^{T} \log p(y_t | y_1, \ldots, y_{t-1}) \\
&= \sum_{t=1}^{T} \log p(y_t | \mathbf{Y_{t-1}}). \qquad (5.19)
\end{aligned}
$$

Since the conditional density, $p(y_t|\mathbf{Y_{t-1}})$, has already been obtained in the filtering step, the log-likelihood of the model is obtained as the byproduct of the recursive non-Gaussian filter. The maximum likelihood estimates of the parameters can be obtained by maximizing this log-likelihood.

The goodness of fit of the model can be evaluated by the AIC (Akaike 1973) defined by

$$
\text{AIC} = -2(\text{maximum log-likelihood}) + 2(\text{number of parameters}). \qquad (5.20)
$$

We shall use this AIC for the comparison of Gaussian and non-Gaussian stochastic volatility models and EGARCH model.

5.2.4 Numerical Implementation of the Nonlinear Smoothing Formulas

Numerical implementation of the non-Gaussian filtering and smoothing formulas for non-Gaussian stochastic volatility model is considered in this subsection. A similar implementation for a nonlinear state space model is given in Kitagawa (1991).

We use a simple step function approximation of an arbitrary function on a two dimensional space, which is specified by the numbers of segments, k_1 and k_2, location of nodes, $w_{1,i}$, $(i = 0, \ldots, k_1)$, $w_{2,j}$, $(j = 0, \ldots, k_2)$ and the value of the density at each node, $p(i,j)$, $(i = 1, \ldots, k_1, j = 1, \ldots, k_2)$. We express the approximated step function by $\{k_1, w_{1,i}, k_2, w_{2,j}, p_{i,j}\}$. Specifically, we use the following notations: $p(\mathbf{x_t}|\mathbf{Y_{t-1}}) \sim p_t(\mathbf{x_t}) \equiv \{k_1, w_{1,i}, k_2, w_{2,j}, p_{i,j}\}$, $p(\mathbf{x_t}|\mathbf{Y_t}) \sim f_t(\mathbf{x_t}) \equiv \{k_1, w_{1,i}, k_2, w_{2,j}, f_{i,j}\}$, $p(\mathbf{x_t}|\mathbf{Y_T}) \sim s_t(\mathbf{x_t}) \equiv \{k_1, w_{1,i}, k_2, w_{2,j}, s_{i,j}\}$.

Similarly, the system noise densities $q_1(x)$ and $q_2(x)$, which are mutually independent, are discretized by using k_{q1} segments, i.e., $q_1(x) \sim \tilde{q}_1(x) \equiv \{k_{q1}, w_{q1,i}, q_{1,i}\}$ and using k_{q2} segments, i.e., $q_2(x) \sim \tilde{q}_2(x) \equiv \{k_{q2}, w_{q2,j}, q_{2,j}\}$, respectively. In the simplest implementation, $w_{1,i} = w_{1,0} + (w_{1,k_1} - w_{1,0})i/k_1$, $(i = 0, \ldots, k_1)$ and $w_{2,j} = w_{2,0} + (w_{2,k_2} - w_{2,0})j/k_2$, $(j = 0, \ldots, k_2)$, $k_{q1} = 2 \times k_1$, $w_{q1,0} = -w_{1,k_1}$, $w_{q1,2\times k_1} = -w_{1,k_1}$, $w_{q1,i} = w_{1,k_1} i/k_1$

and $k_{q2} = 2 \times k_2$, $w_{q2,0} = -w_{2,k_2}$, $w_{q2,2\times k_2} = -w_{2,k_2}$, $w_{q2,j} = w_{2,k_2} j/k_2$. To realize the recursive formulas shown below, it is necessary to develop a numerical method for the linear transformation of variables, the convolution of densities, Bayes formula, and the normalization.

One step ahead prediction:

$p_{i,j}$ $(i = 1, \ldots, k_1, j = 1, \ldots, k_2)$ is obtained by

$$\begin{aligned} p_{i,j} = p_t(w_{1,i}, w_{2,j}) &\simeq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(w_{1,i}, w_{2,j}|h_{t-1}, l_{t-1}) p(h_{t-1}, l_{t-1}|\mathbf{Y_{t-1}}) dh_{t-1} dl_{t-1} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q_1(w_{1,i} - z_1) q_2(w_{2,j} - z_2) \hat{f}_{t-1}(z_1, z_2) dz_1 dz_2 \\ &= \sum_{i=1}^{k_1}\sum_{j=1}^{k_2} \int_{z_{1,i-1}}^{z_{1,i}} \int_{z_{2,j-1}}^{z_{2,j}} \tilde{q}_1(w_{1,i} - z_1) \tilde{q}_2(w_{2,j} - z_2) \hat{f}_{t-1}(z_1, z_2) dz_1 dz_2, \end{aligned} \tag{5.21}$$

with $z_1 \equiv \alpha_0 + \alpha_1 h_{t-1}$ and $z_2 \equiv l_{t-1}$. The density $\hat{f}_{t-1}(z_1, z_2)$ for the transformed state $(z_1, z_2)^t$ can be evaluated numerically by the following algorithm. For simplicity, we assume that the nodes $\{w_{1,i}\}$ are equally spaced and that $\Delta w_1 = (w_{1,k_1} - w_{1,0})/k_1$.

1. For $i = 1$ to k_1 and $j = 1$ to k_2

   put $\hat{f}_{i,j} = 0$.

2. For $i = 1$ to k_1

   (a) $z_{1,0} = \min\{\alpha_0 + \alpha_1 w_{1,i-1}, \alpha_0 + \alpha_1 w_{1,i}\}$

   (b) $z_{1,3} = \max\{\alpha_0 + \alpha_1 w_{1,i-1}, \alpha_0 + \alpha_1 w_{1,i}\}$

   (c) $i_0 = [\frac{z_{1,0} - w_{1,0}}{\Delta w_1}]$, $i_1 = [\frac{z_{1,3} - w_{1,0}}{\Delta w_1}] + 1$

   (d) for $n = i_0 + 1$ to i_1

   $z_{1,1} = \max\{z_{1,0}, w_{1,0} + (j-1)\Delta w_1\}$

   $z_{1,2} = \min\{z_{1,3}, w_{1,0} + j\Delta w_1\}$

   $\hat{f}_{n,j} = \hat{f}_{n,j} + \frac{z_{1,2} - z_{1,1}}{z_{1,3} - z_{1,0}} f_{i,j}$

Using this $\hat{f}_{i,j}$, (5.21) can be evaluated approximately by $p_{i,j} \simeq \sum_{\alpha=-k_1}^{k_1} \tilde{q}_{1,\alpha} \left(\sum_{\beta=-k_2}^{k_2} \tilde{q}_{2,\beta} \hat{f}_{\alpha,\beta}\right)$.

Filtering

$f_{i,j}$ $(i = 1, \ldots, k_1, j = 1, \ldots, k_2)$ is obtained by

$$f_{i,j} = f_t(w_{1,i}, w_{2,j}) = \frac{p_t(w_{1,i}, w_{2,j}) m(y|(w_{1,i}, w_{2,j})^t)}{C} = \frac{p_{i,j} r_{i,j}}{C}. \tag{5.22}$$

Here y is the given observation at time t and $r_{i,j} = m(y|(w_{1,i}, w_{2,j})^t)$ can be evaluated directly from the function $m(v)$. In (5.22), C is the normalized constant given by

$$
\begin{aligned}
C = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_t(\mathbf{x_t})m(y|\mathbf{x_t})d\mathbf{x_t} &\simeq \sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\int_{z_{1,i-1}}^{z_{1,i}}\int_{z_{2,j-1}}^{z_{2,j}} m(y|(w_{1,i}, w_{2,j})^t)\hat{p}_t(z_1, z_2)dz_1dz_2 \\
&\simeq \sum_{i=1}^{k_1}\sum_{j=1}^{k_2} r_{i,j}p_{i,j}\Delta w_1
\end{aligned}
\tag{5.23}
$$

Smoothing

$s_{i,j} \quad (i = 1, \dots, k_1, j = 1, \dots, k_2)$ is obtained by

$$
\begin{aligned}
s_{i,j} = s_t(w_{1,i}, w_{2,j}) &= f_t(w_{1,i}, w_{2,j})\int_{-\infty}^{\infty} \frac{s_{t+1}(z_1, z_2)\hat{q}_1(z_1 - \alpha_0 - \alpha_1 w_{1,i})\hat{q}_2(z_2 - w_{2,j})}{f_{t+1}(z_1, z_2)}dz_1dz_2 \\
&= f_t(w_{1,i}, w_{2,j}) \\
&\quad \times \sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\int_{z_{1,i-1}}^{z_{1,i}}\int_{z_{2,j-1}}^{z_{2,j}} \frac{s_{t+1}(z_1, z_2)\hat{q}_1(z_1 - \alpha_0 - \alpha_1 w_{1,i})\hat{q}_2(z_2 - w_{2,j})}{f_{t+1}(z_1, z_2)}dz_1dz_2
\end{aligned}
\tag{5.24}
$$

The integral in the summation is given approximately by $s_{i,j}\hat{q}_{1,i}q_{2,j}/\hat{f}_{i,j}$, where $s_{i,j} = s_{t+1}(w_{1,i}, w_{2,j})$, $\hat{f}_{i,j} = \hat{f}_{t+1}(w_{1,i}, w_{2,j})$ and $\hat{q}_{1,i}$ can be evaluated either directly or numerically by a similar way as the one for the prediction.

5.3 Results

We analyzed the daily returns of Standard and Poor's 500 Index for 1985-1994 shown in Figure 5.1. We compared six versions of stochastic volatility models and EGARCH(1,1) model (Nelson (1991)) which is defined by

$$
h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2\epsilon_{t-1} + \alpha_3\left(|\epsilon_{t-1}| - \sqrt{2/\pi}\right). \tag{5.25}
$$

In Table 5.1, EGARCH, GSV and NGSV denote EGARCH(1,1) model, Gaussian and non-Gaussian stochastic volatility models, respectively. T denotes restricted trend (or random walk) type model for the volatility, namely the model obtained by assuming $\alpha_0 = 0$, $\alpha_1 = 1$ in (5.2), (5.5) or (5.7). AR denotes unrestricted autoregressive type model with constant term given in (5.2), (5.5) and (5.7). S denotes the most general model with time-varying shape parameter. LL denotes the log-likelihood of the estimated model.

The AIC values of these stochastic volatility models are significantly smaller than that of EGARCH. This result supports the research by Heyman et al.. Within the Gaussian stochastic volatility model, the AIC of GSV-AR is significantly smaller than that of GSV-T. This indicates that it is better to assume that the model does not have a unit root.

The AIC values of all of four non-Gaussian stochastic volatility models are significantly smaller than those of the Gaussian models. Within these four models, the autoregressive type models are better than the restricted trend type model with $\alpha_0 = 0$, $\alpha_1 = 1$. However, the differences of AICs are only about 3 and are not so significant compared with the case of Gaussian stochastic volatility model. The models with time varying shape parameter is better than the ones with time-invariant shape parameter. The overall AIC best model was NGSV-AR-S, namely the non-Gaussian stochastic volatility model with AR type volatility change and time varying shape parameter.

To check the structural changes of the economic system in ten year period, we devide the time interval into two subintervals and fitted the six models to each subset of data. The results of the models for the first half of the data are summarized in Table 5.2. The ones for the latter half are given in Table 5.3.

The most interesting findings with these tables are that the restricted trend type ($\alpha_0 = 0$, $\alpha_1 = 1$) become consistently better than the autoregressive type models in both of the divided data. Probably, this indicates that there is one reverting cycle of the dispersion parameter in ten years not but in the five years. This phenomena relate with business cycle which is ten years long.

In Figure 5.1, the upper graph shows the time series of daily returns in 1985-1994, the lower graph shows the estimated time evolution of the dispersion, σ^2 obtained by GSV-AR model. In Figure 5.2, the upper graph shows the smoothed posterior density of $h_t = \log \tau_t^2$, and the lower graph shows the smoothed posterior density of $l_t = \log(b_t - 1/2)$ obtained by NGSV-AR-S model. The figure shows the mean and $\pm$ one standard deviation intervals of each distribution. When the market crashed down, h_t was rising to the peak and l_t was falling to the bottom. It indicates that the big volatility change was decomposed to the two part depend on scale parameter and shape parameter. In Figure 5.3, the upper graph shows the filtered posterior density of $h_t = \log \tau_t^2$, and the lower graph shows the filtered posterior density of $l_t = \log(b_t - 1/2)$. Figure 5.4 shows the birds-eye view of the filtered distributions (above) and the smoothed distributions (below).

In Figure 5.5, the upper graph shows the time series of daily returns in 1985-1989, and the lower graph shows the estimated time evolution of the dispersion, σ^2 obtained by GSV-AR model. In Figure 5.6, the upper graph shows the smoothed posterior density of $h_t = \log \tau_t^2$, and the lower graph shows the smoothed posterior density of $l_t = \log(b_t - 1/2)$ obtained by NGSV-T-S model. The level of h_t was almost the same during this period and the movement of l_t corresponded to the movement of the market volatility. It indicates that in this period the market volatility was explained by the shape parameter change. In Figure 5.7, the upper graph shows the filtered posterior density of $h_t = \log \tau_t^2$, and the lower graph shows the filtered posterior density of $l_t = \log(b_t - 1/2)$. Figure 5.8 shows the birds-eye view of the filtered distributions (above) and the smoothed distributions (below).

In Figure 5.9, the upper graph shows the time series of daily returns in 1990-1994, and the lower graph shows the estimated time evolution of the dispersion, σ^2 obtained by GSV-AR model. In Figure 5.10, the upper graph shows the smoothed posterior density of $h_t = \log \tau_t^2$, and the lower graph shows the smoothed posterior density of $l_t = \log(b_t - 1/2)$ obtained by NGSV-T model. The level of l_t was the same during this period and the movement of h_t corresponded to the movement of the market volatility. It indicates that in this period the market volatility was explained by the scale parameter change. In Figure 5.11, the upper graph shows the filtered posterior density of $h_t = \log \tau_t^2$ of 1985-1994, and the lower graph shows the filtered posterior density of $l_t = \log(b_t - 1/2)$. Figure 5.12 shows the birds-eye view of the filtered distributions (above) and the smoothed distributions (below).

Table 5.1: Comparison of various types of stochastic volatility model fitted to the entire data set (1985-1994)

Model	α_0	α_1	$\mu^2(\alpha_2)$	$\xi_1^2(\alpha_3)$	ξ_2^2	LL	AIC
EGARCH	-0.262	0.971	-0.070	0.184	—	8453.99	-16900.00
GSV-T	0	1	0.023	—	—	8549.75	-17097.50
GSV-AR	0.185	0.965	0.044	—	—	8566.54	-17127.08
NGSV-T	0	1	—	0.0033	0	8607.37	-17212.74
NGSV-AR	0.025	0.993	—	0.0048	0	8611.03	-17216.06
NGSV-T-S	0	1	—	0.0008	0.0010	8609.97	-17215.94
NGSV-AR-S	0.060	0.984	—	0.0045	0.0004	8613.48	-17218.96

Table 5.2: Comparison of various types of stochastic volatility model fitted to the first half data set (1985-1989)

Model	α_0	α_1	$\mu^2(\alpha_2)$	$\xi_1^2(\alpha_3)$	ξ_2^2	LL	AIC
EGARCH	-0.582	0.934	-0.114	0.232	—	4029.01	-8050.01
GSV-T	0	1	0.043	—	—	4096.81	-8191.62
GSV-AR	0.190	0.965	0.041	—	—	4106.58	-8207.16
NGSV-T	0	1	—	0.0052	0	4132.32	-8262.64
NGSV-AR	0.025	0.993	—	0.0051	0	4134.83	-8263.66
NGSV-T-S	0	1	—	0.0001	0.0024	4136.44	-8268.88
NGSV-AR-S	0.125	0.967	—	0.0018	0.0017	4138.23	-8268.46

Table 5.3: Comparison of various types of stochastic volatility model fitted to the latter half data set (1990-1994)

Model	α_0	α_1	$\mu^2(\alpha_2)$	$\xi_1^2(\alpha_3)$	ξ_2^2	LL	AIC
EGARCH	-0.002	0.999	-0.047	0.001	—	4460.45	-8912.90
GSV-T	0	1	0.006	—	—	4458.15	-8914.30
GSV-AR	0.050	0.990	0.0063	—	—	4461.80	-8917.60
NGSV-T	0	1	—	0.0014	0	4474.97	-8947.94
NGSV-AR	0.025	0.993	—	0.0026	0	4476.10	-8946.20
NGSV-T-S	0	1	—	0.0014	0.00002	4474.97	-8945.94
NGSV-AR-S	0.022	0.994	—	0.0024	0.00003	4476.20	-8944.40

Figure 5.1: Time series of data (1985-1994) and Gaussian S.V.M.'s σ^2

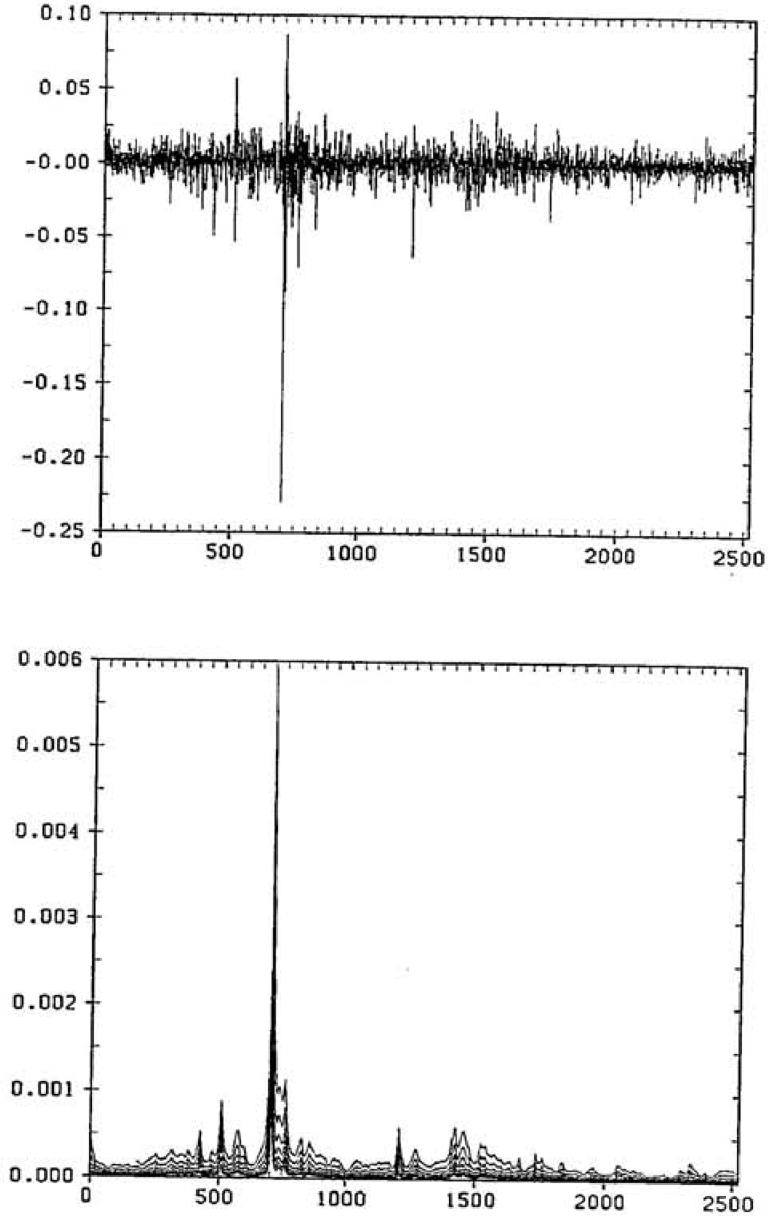


Figure 5.2: The smoothed estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-AR-S (1985-1994)

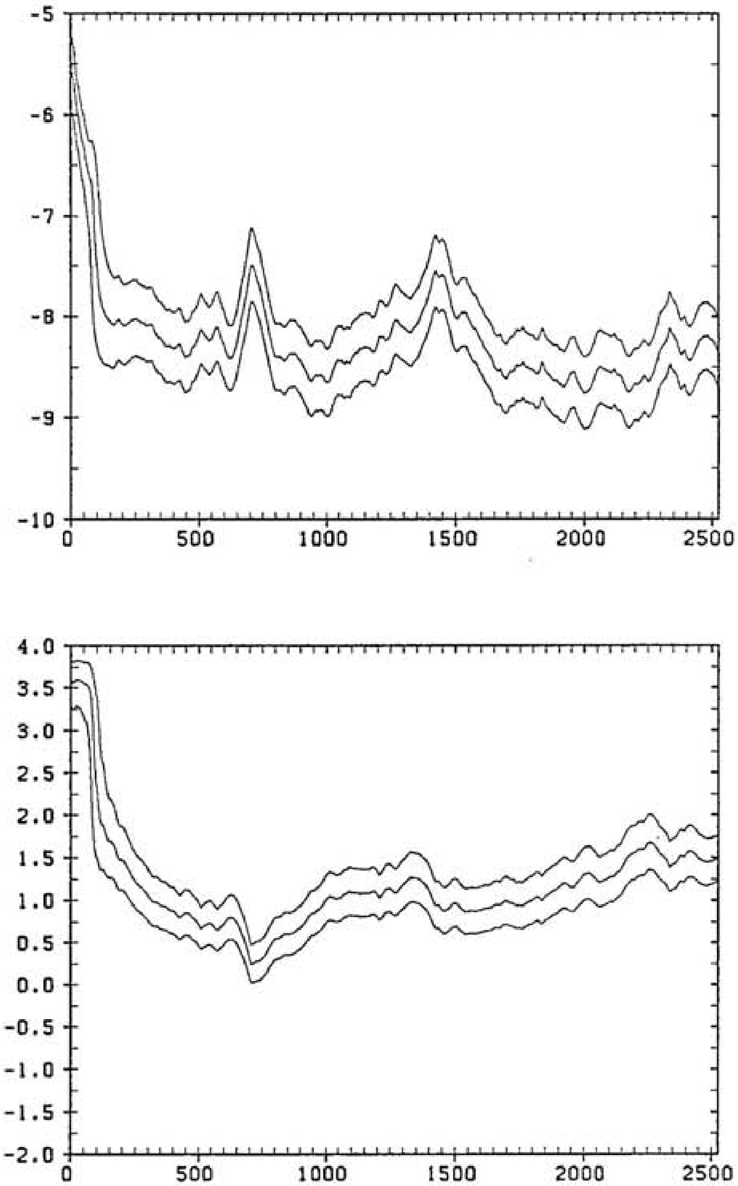


Figure 5.3: The filterd estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-AR-S (1985-1994)

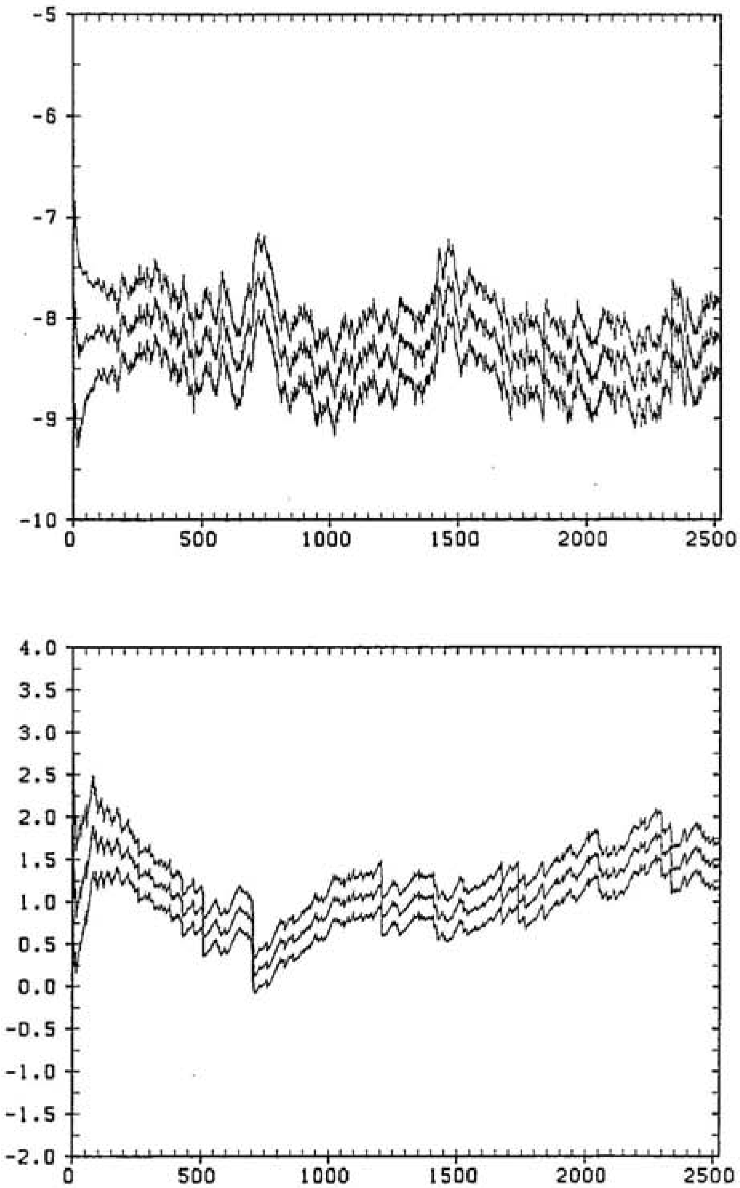


Figure 5.4: The filtered distribution, and the below graph is the smothed distribution (1985-1994)

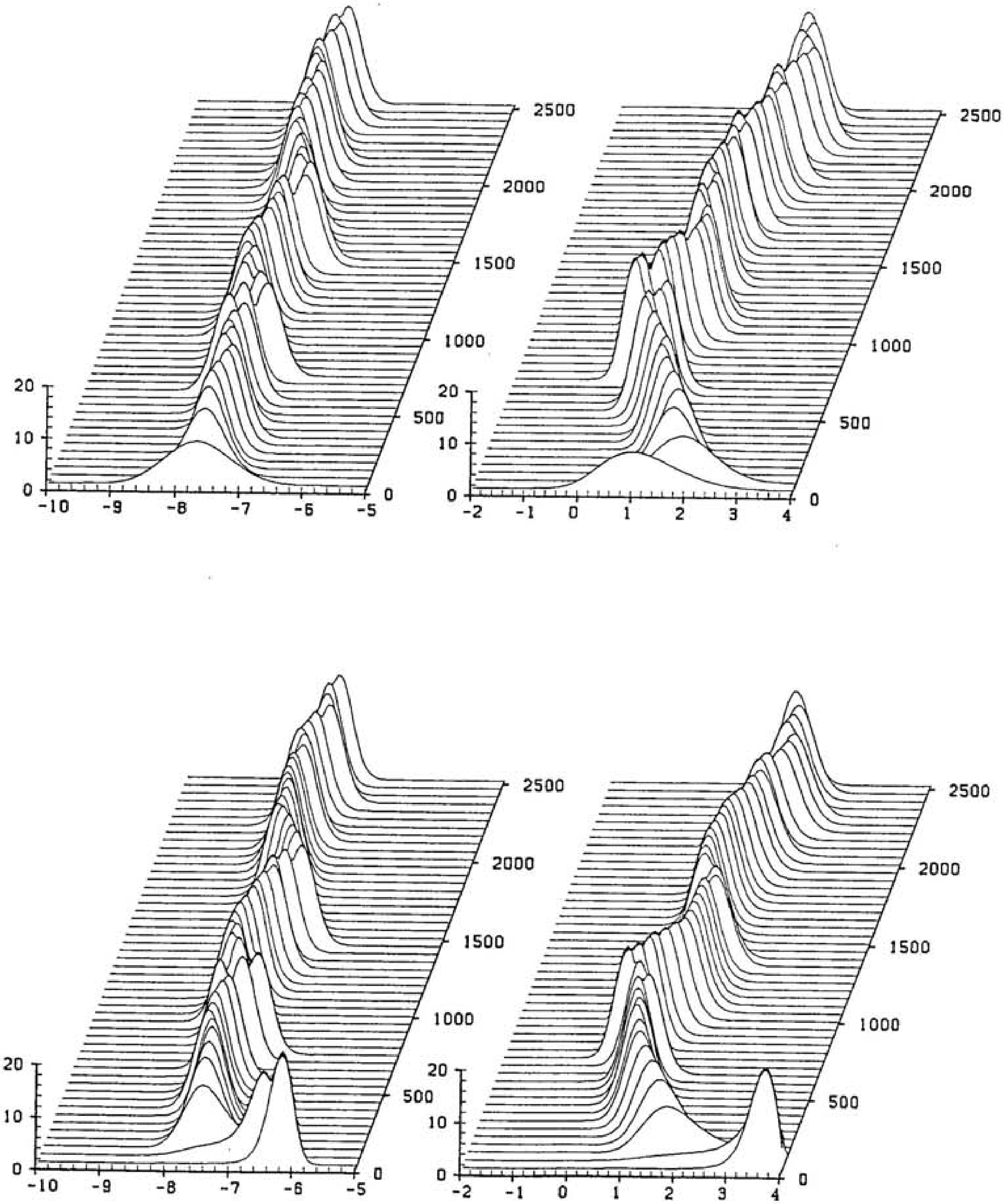


Figure 5.5: Time series of data (1985-1989) and Gaussian S.V.M.'s σ^2

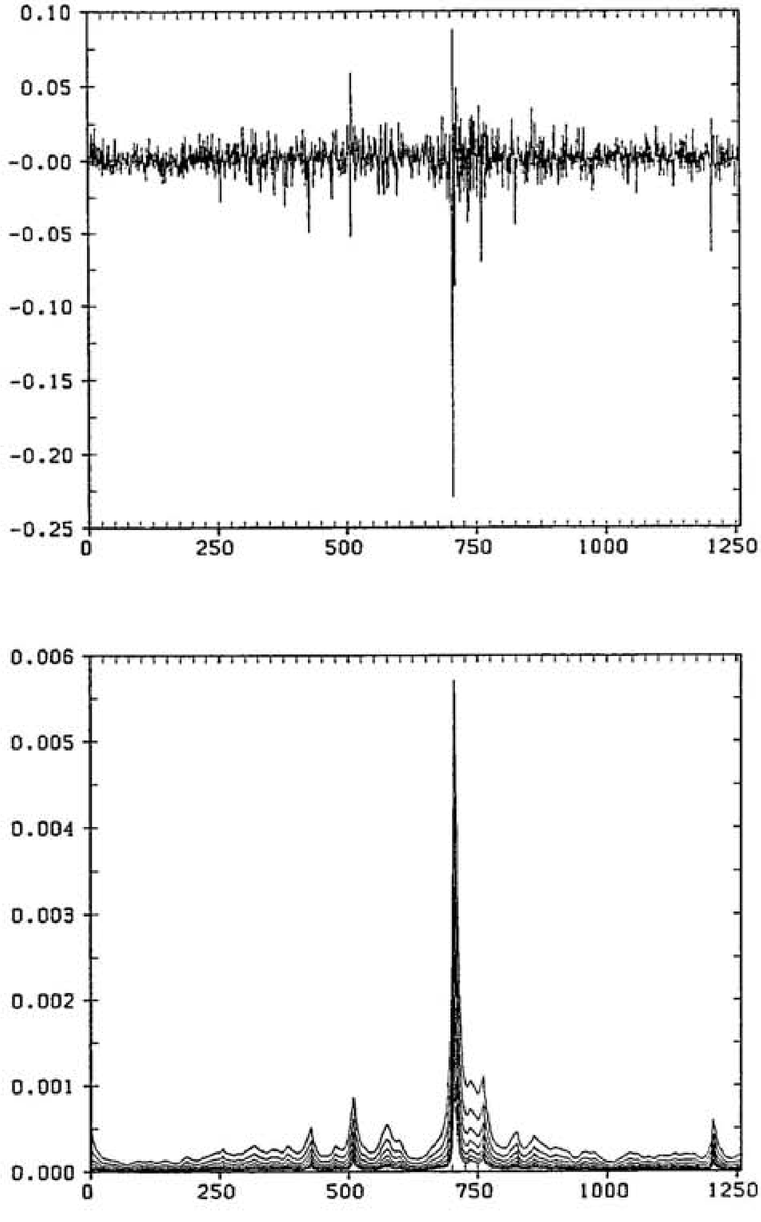


Figure 5.6: The smoothed estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-T-S (1985-1989)

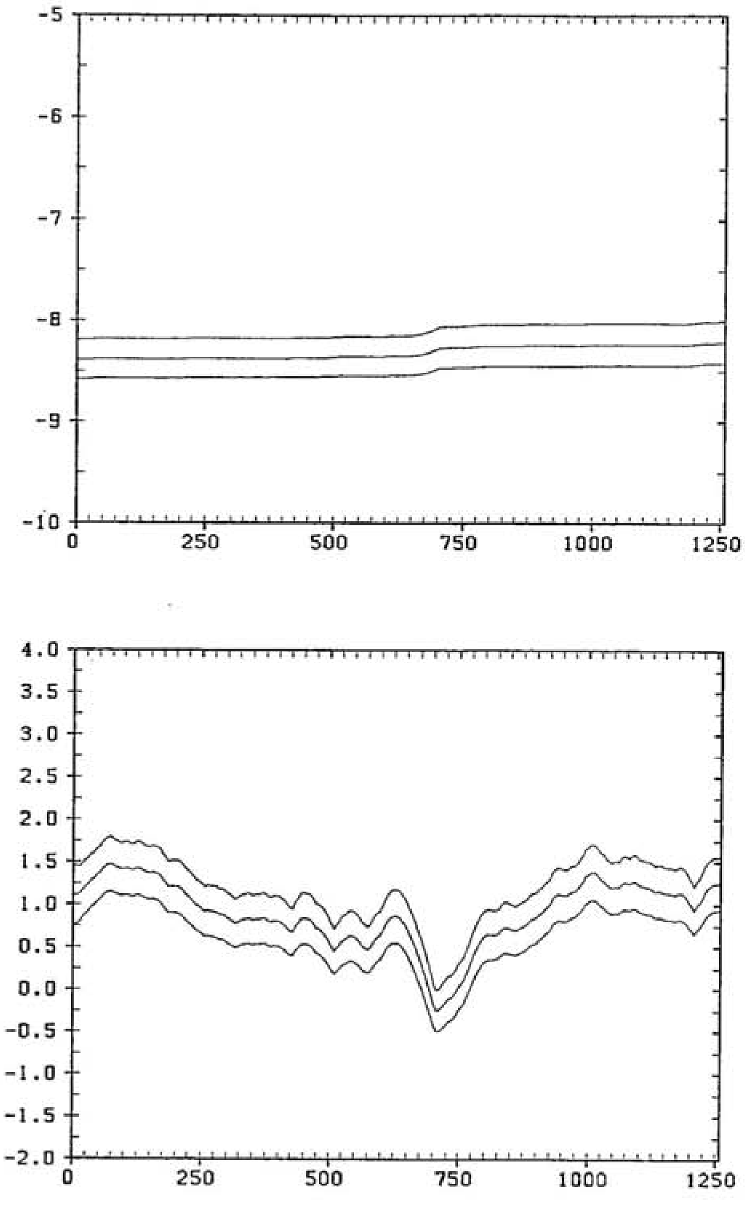


Figure 5.7: The filterd estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-T-S (1985-1989)

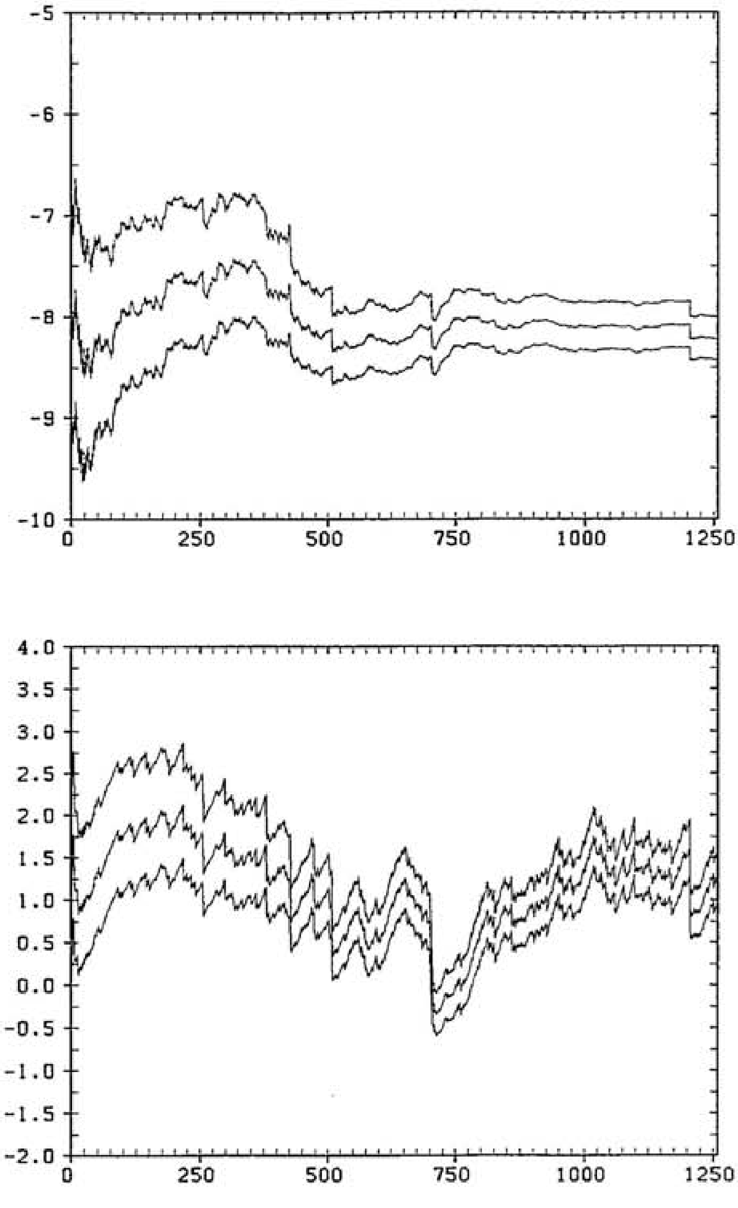


Figure 5.8: The filtered distribution, and the below graph is the smothed distribution (1985-1989)

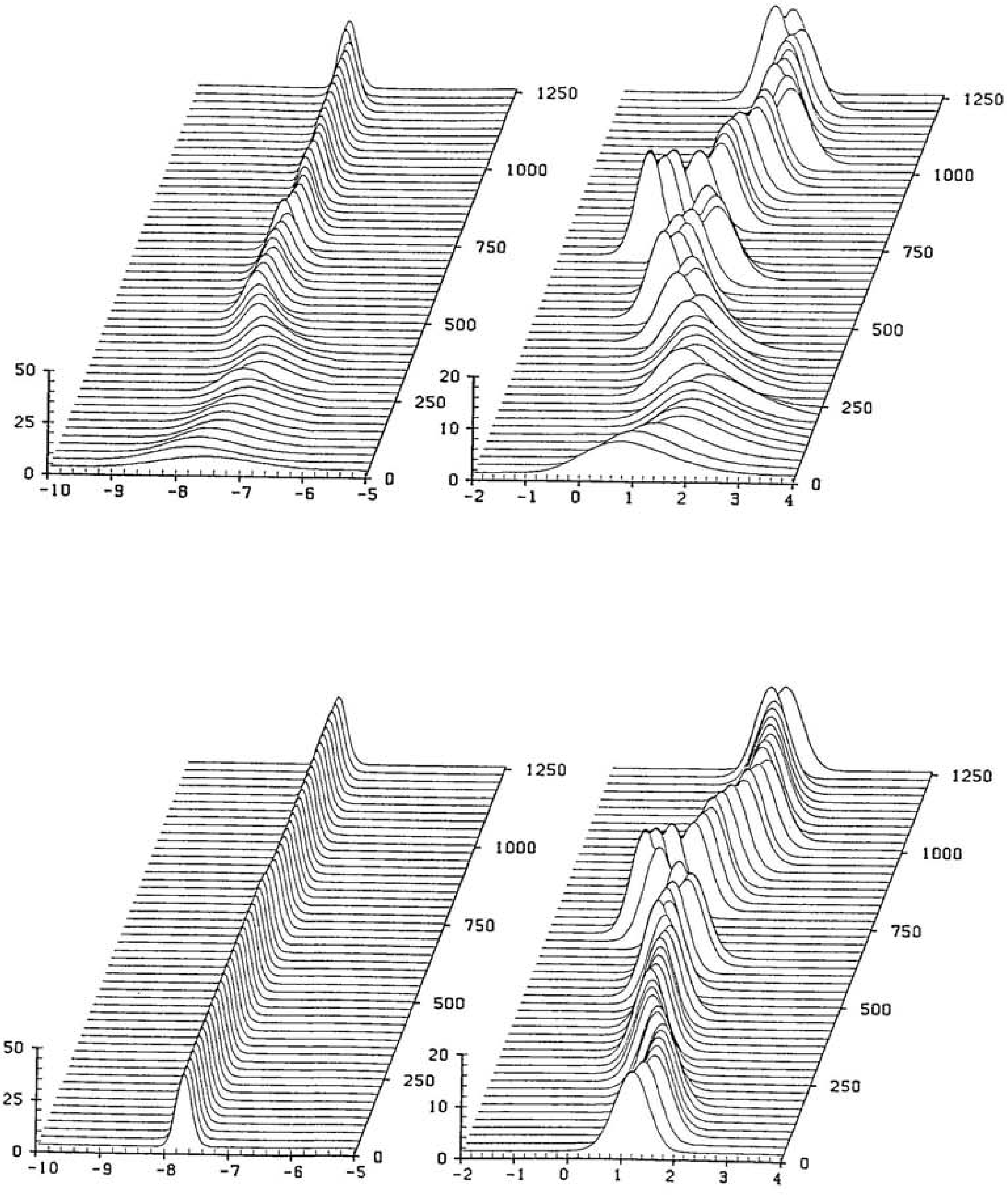


Figure 5.9: Time series of data (1990-1994) and Gaussian S.V.M.'s σ^2

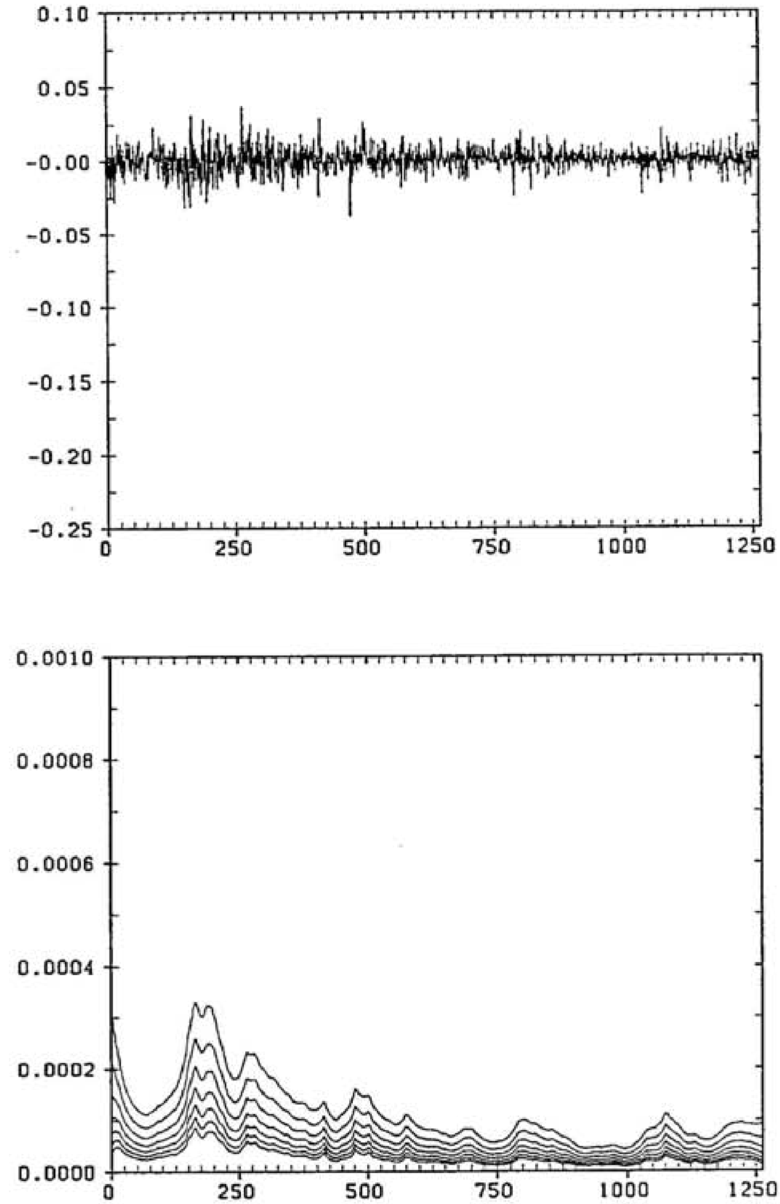


Figure 5.10: The smoothed estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-T (1990-1994)

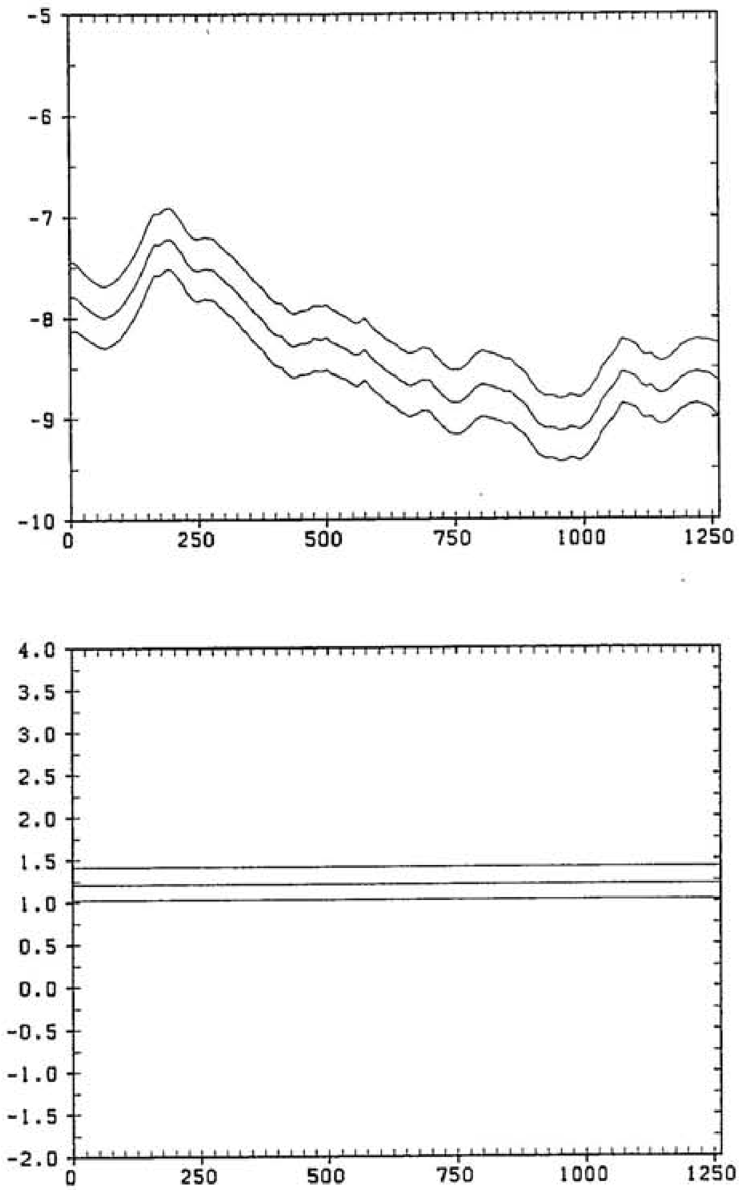
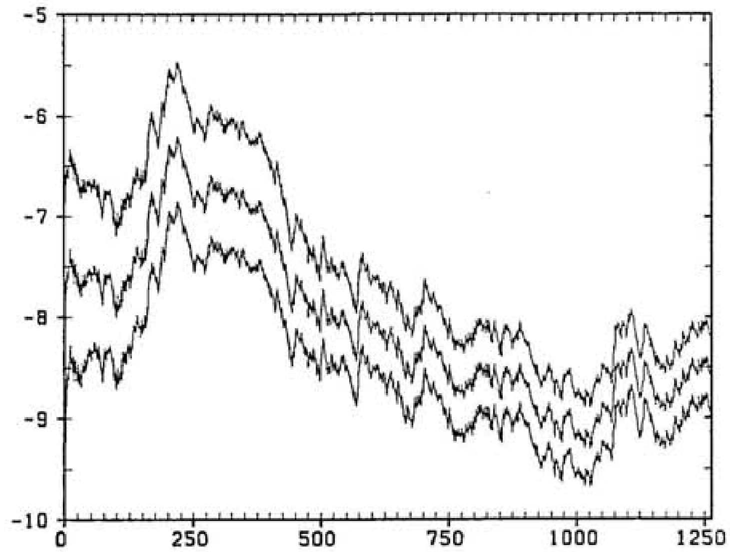


Figure 5.11: The filterd estimate of $h_t = \log \tau_t^2$ (above) and $l_t = \log(b_t - 1/2)$ (below) obtained by the NGSV-T (1990-1994)



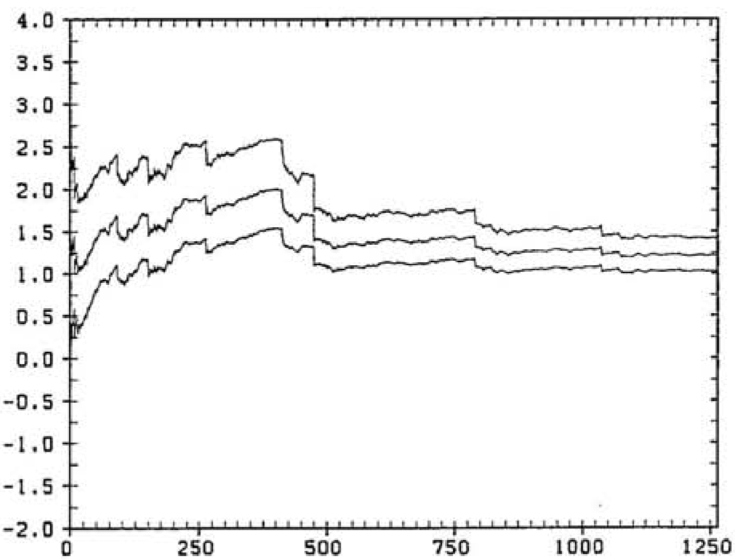
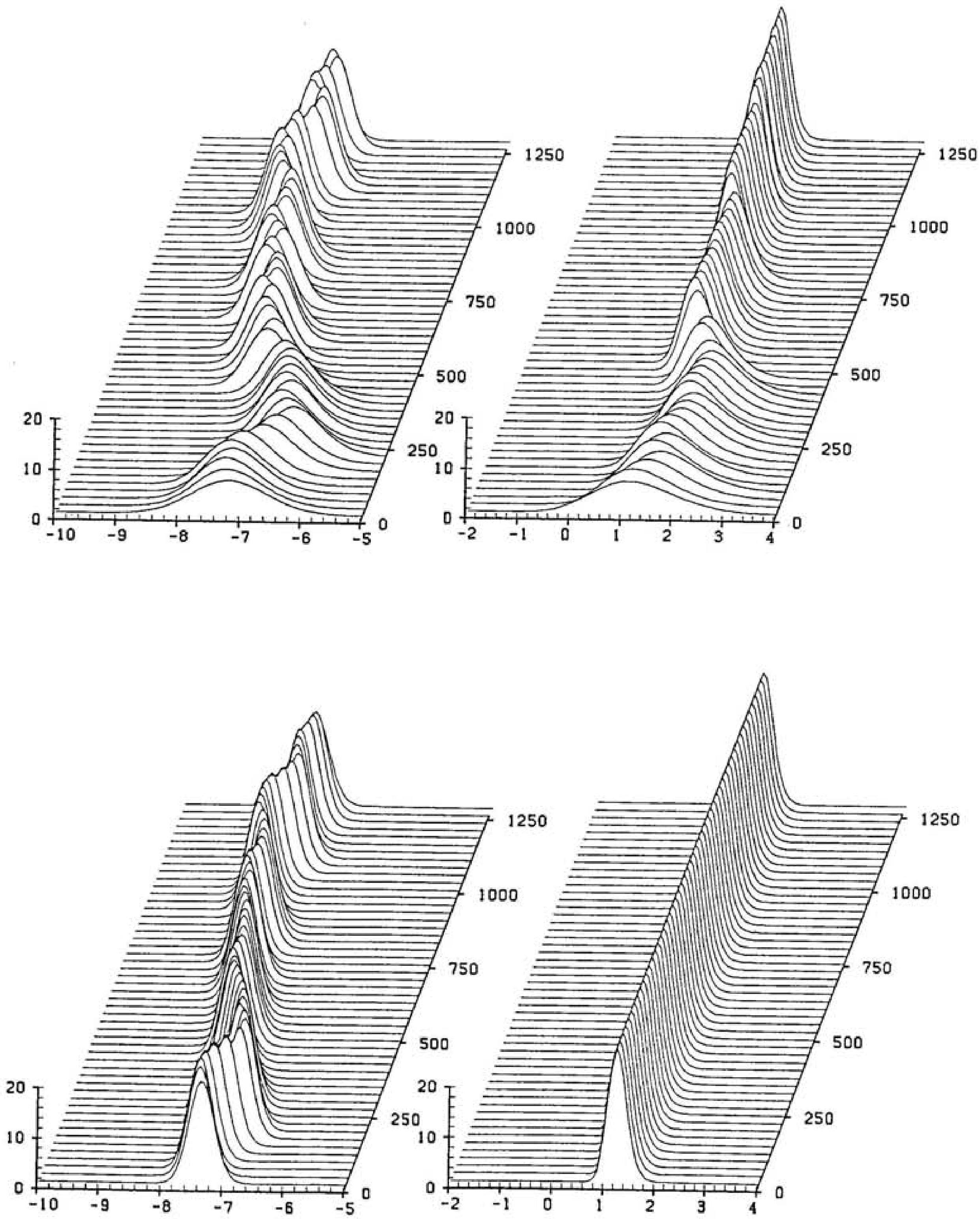


Figure 5.12: The filtered distribution, and the below graph is the smothed distribution (1990-1994)



5.4 Conclusion

Non-Gaussian version of the stochastic volatility model was proposed. The generalization to time-varying shape parameter model is also considered. We showed that the parameters of this model can be estimated by the general state space model and the numerical implementation. According to the AIC criterion, the non-Gaussian stochastic volatility model is better than Gaussian stochastic volatility model. Especially, the difference of the AIC values between two models for the interval, 1985-1989, is much more than that for the interval, 1990-1994. Furthermore, the AIC values of the models with time-varying shape parameter is smaller than that of the model having the constant shape parameter for the interval, 1985-1989 and the entire period. However, for the interval, 1990-1994, the log-likelihood value is almost the same. For the interval, 1985-1989, the movement of the shape parameter indicates the structual change before October, 19, 1987, "Black Monday". Therefore, the model having the time-varying shape parameter is good for predicting the big structural change and has better AIC than the model having the constant shape parameter in some cases.

Chapter 6

Monte Carlo Smoothing Method for Seasonal Adjustment

6.1 Introduction

Various time series models have been developed for seasonal adjustment, e.g., seasonal ARIMA model (Hillmer and Tiao 1982), Bayesian model (Akaike and Ishiguro 1980) and the state space models (Kitagawa 1981, Schlicht 1981, Kitagawa and Gersch 1984, Harvey 1989).

The main advantages of these modeling approach are that the model reveals the assumption used in the analysis explicitly and that there are objective criteria for the estimation of the parameters and the evaluation of the model such as the likelihood and the AIC (Akaike 1973). Therefore, if there arises a problem which needs to be solved for specific data set, we can freely develop a new model to cope with that problem and then compare it with the standard models. The state space model can also be applied to the analysis of count data such as the Poisson or binomially distributed process (Kitagawa 1987, Harvey and Fernandes 1989, Kashiwagi and Yanagimoto 1992). Früwirth-Schnatter (1994) considered seasonal adjustment of count data.

The problem with the non-Gaussian modeling is the computational cost. In any of the recursive filtering algorithms including the non-Gaussian filtering, the necessary amount of comutation is proportional to the data length n, namely $O(n)$. In particular, the Kalman filter algorithm for Gaussian linear structural state space models requires computations proportional to $k \times n$ or $k^2 \times n$, where k is the dimension of the state vector. On the other hand, in the general non-Gaussian filtering (Kitagawa 1987), computationaly costly numerical integration is necessary (i.e., $O(d^k \times n)$ computation is necessary with d being the number of segments for each domain of integration). Thus the direct application of the method to the seasonal adjustment, where the state space model with state dimension over 13 is necessary, is impossible.

To alleviate this problem, West, Harrison and Migon (1985) and Harvey and Fernandes (1989) used conjugate priors. A Gaussian-sum filter and smoother were used to approximate the non-Gaussian filter and smoother in Kitagawa (1989, 1994). Schnatter (1992) developed a numerical integration based Kalman filter, in which the distribution of the state vector is approximated by a Gaussian distribution.

In this chapter, we use a new method for state estimation based on the Monte Carlo filter and smoother developed in Kitagawa (1993). The algorithm is based on the approximation of successive prediction and filtering density functions by many of their realizations. The difference between the present algorithm and other Monte Carlo-Gibbs sampling methods, (Carlin, Polson and Stoffer 1992, Früwirth-Schnatter 1994), is that we use the Monte Carlo method for the entire filtering and smoothing procedure whereas the others are used for numerical integration. The virtue of this algorithm is that it can be applied to very wide class of nonlinear non-Gaussian higher dimensional state space models, if the dimensions of the system noise and the observational noise are low by simply specifying the nonlinear functions and noise densities. It can properly handle the filtering and smoothing problems even when the distribution of the state vector is multimodal or the observations are discrete valued.

In this chapter, the seasonal adjustment of count data and additive and multiplicative, namely mixed type decomposition are considered for seasonal adjustment. To develop a unified procedure for the paramater estimation, the model evaluation and the decomposition, a Monte Carlo filter and smoother (Kitagawa 1993) are used.

6.2 A Non-Gaussian State Space Model for Seasonal Adjustment

6.2.1 Non-Gaussian nonlinear models for seasonal adjustment

The standard state space model for seasonal adjustment is given by (see, for example, Kitagawa 1981, Schlichit 1981 and Kitagawa and Gersch 1984)

$$\begin{aligned} x_n &= Fx_{n-1} + Gv_n \\ y_n &= Hx_n + w_n, \end{aligned} \tag{6.1}$$

where x_n is the $p+1$ dimensional state vector, with p being the period length, F, G and H are $(p+1)\times(p+1)$, $(p+1)\times 2$ and $1\times(p+1)$ matrices defined by

$$
\begin{aligned}
x_n &= \left[\begin{array}{c} T_n \\ T_{n-1} \\ \hline S_n \\ S_{n-1} \\ \vdots \\ S_{n-p+1} \end{array}\right], \quad F = \left[\begin{array}{cc|cccc} 2 & -1 & & & & \\ 1 & 0 & & \mathbf{0} & & \\ \hline & & -1 & \cdots & -1 & -1 \\ & & 1 & & & \\ & \mathbf{0} & & \ddots & & \\ & & & & 1 & 0 \end{array}\right], \quad G = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ \hline 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{array}\right] \\
H &= [\,1\ \ 0\ \ 1\ \ 0\ \ \cdots\ \ 0\,]. \qquad (6.2)
\end{aligned}
$$

In the definition of the state vector, T_n, and S_n are the trend and the seasonal component, respectively. The noise inputs v_n and w_n are mutually independent white noise sequence with mean zero and the covariance matrix $Q = \mathrm{diag}(\tau_1^2, \tau_2^2)$ and the variance σ^2, respectively. Harvey (1989) used a slightly different representation of the model called a structural time series model.

One of the significant merits of the state space modeling is that the model can be easily modified or extended to deal with various situations. Some examples of the extension of the models to include other components than the trend and seasonal ones such as the trading day factor or the stationary component are given in Kitagawa and Gersch (1984). In Kitagawa (1989), the model was generalized to the one with non-Gaussian noise inputs to detect sudden changes of trend or seasonal component and to handle outliers in the observation.

In this chapter, we consider a more general state space model of the form

$$
\begin{aligned}
x_n &= F(x_{n-1}, v_n) \qquad (6.3) \\
y_n &\sim P(\cdot|x_n), \qquad (6.4)
\end{aligned}
$$

where $v_n = (v_{n1}, v_{n2})^t$ is a two dimensional white noise sequence, $v_{n1} \sim q_1(v), v_{n2} \sim q_2(v)$ with q_1 and q_2 being possibly non-Gaussian density functions. P is a distribution function which is specified by the state vector x_n. The function $F(x, v)$ is a possibly nonlinear function of x and v; $R^p \times R^2 \to R^p$. The initial state, x_0, is assumed to be distributed as the density function $p_0(x)$.

If the distribution of the observed process is continuous and has a density function, the observation model (6.4) can be written more explicitly as

$$y_n = h(x_n, w_n), \tag{6.5}$$

where w_n is a 1-dimensional observation noise with density function $r(w)$ and h is a possibly nonlinear function $R^k \times R \to R$.

It is possible to consider a mixed multiplicative and additive model,

$$h(x_n, w_n) = \exp\{T_n\} \times \exp\{S_n\} + w_n. \tag{6.6}$$

In this case, $\exp\{T_n\}$ and $\exp\{S_n\}$ are considered as the trend and the seasonal factor, respectively. Fitting the above multiplicative model with log-normal noise distribution is equivalent to fit an additive model with normal distribution noise after taking logarithm of the original time series. For discrete valued series such as the nonstationary Poisson or binomial process with seasonal variation, we can consider, for example, a model which is expressed by using the conditional distribution

$$\mathrm{Prob}\{y_n = \ell | x_n\} \;=\; \frac{e^{-\lambda_n}\lambda_n^{\ell}}{\ell!} \tag{6.7}$$

where λ_n is the mean value function with trend and seasonality and is expressed by

$$\lambda_n \;=\; \exp\{T_n\} \times \exp\{S_n\}. \tag{6.8}$$

Here $\mathrm{Prob}\{y_n = \ell | x_n\}$ denotes the conditional probability of the observation being ℓ given the state vector x_n defined in (6.2). With this model, it is possible to analyze, a count data with trend and seasonal variation in the intensity function.

6.2.2 State Estimation

The most important problem in state space modeling is the estimation of the state vector x_n from the observations, since many problems in time series analysis such as the likelihood computation for parameter estimation, the prediction, the interpolation and the decomposition of a seasonal time series into the trend and the seasonal components can be handled by estimating the state vector x_n.

The problem of state estimation can be formulated as to evaluate the conditional density $p(x_n|Y_t)$, where Y_t is the set of observations defined by $Y_t = \{y_1, \ldots, y_t\}$. Corresponding to three distinct cases, $n > t$, $n = t$ and $n < t$, the state estimation problem can be classified into three categories and the conditional density $p(x_n|Y_t)$ is called the predictor, the filter and the smoother, respectively.

For the standard linear Gaussian state space model, each density can be expressed by a Gaussian density and its mean vector and the variance-covariance matrix can be obtained by computationally efficient recursive formula such as the Kalman filter and the fixed interval smoothing algorithms (Sage and Melsa 1971, Anderson and Moore 1979).

For nonlinear or non-Gaussian state space models, however, the state distributions become non-Gaussian and various types of approximations to or assumtions on the densities are used to obtain recursive formula for state estimation. Typical examples are the extended Kalman filter (Anderson and Moore 1979), the Gaussian-sum filter (Alspach and Sorenson 1972), dynamic generalized linear model (West, Harrison and Migon 1985), non-Gaussian filter and smoother (Kitagawa 1987, Hodges and Hale 1993, Tanizaki 1993). In this chapter, we shall use a different method based on Monte Carlo filter and smoother (Kitagawa 1993).

6.3 Monte Carlo Filtering and Smoothing

In this section we shall briefly review the algorithms of Monte Carlo filter and smoother (Kitagawa 1993). By the use of these algorithms, it became possible to apply the general nonlinear non-Gaussian state space model to seasonal adjustment problems.

In the Monte Carlo filter and smoother, each state or noise density function is approximated by many of realizations (called "particles") from that distribution. It is equivalent to approximate the distributions by the empirical distribution functions determined by the set of particles. Then it can be shown that a set of particles expressing the one step ahead predictor $p(x_n|Y_{n-1})$ and the filter $p(x_n|Y_n)$ can be obtained recursively.

6.3.1 The Algorithm for Monte Carlo Filtering

In the following algorithms, $p_n^{(j)}$, $f_n^{(j)}$ and $v_n^{(j)}$ denote the j-th independent realizations from $p(x_n|Y_{n-1})$, $p(x_n|Y_n)$ and $q(v_n)$, respectively. We use m realizations (particles) to approximate each density function. For example, given the m particles $v_n^{(1)}, \cdots, v_n^{(m)}$, the cumulative distribution function

$$Q(x) = \int_{-\infty}^{x} q(t)dt \tag{6.9}$$

can be approximated by the emperical distribution function defined by

$$\frac{1}{m}\sum_{j=1}^{m} I(x; v_n^{(j)}), \tag{6.10}$$

where $I(x; v)$ is the indicator function defined by

$$I(x; v) = \begin{cases} 1 & x \geq v \\ 0 & x < v \end{cases}. \tag{6.11}$$

Now, it can be shown that $p_n^{(j)}$ and $f_n^{(j)}$ are obtained recursively by the following Monte Carlo filter algorithm.

[Monte Carlo Filter]

1. *Generate a* $(p+1)$*-dimensional random number* $f_0^{(j)} \sim p_0(x)$ *for* $j = 1, \ldots, m$.

2. *Repeat the following steps for* $n = 1, \ldots, N$.

   (a) *Generate a 2-dimensional random number* $v_n^{(j)} \sim q(v)$ *for* $j = 1, \ldots, m$.

   (b) *Compute* $p_n^{(j)} = F(f_{n-1}^{(j)}, v_n^{(j)})$ *for* $j = 1, \ldots, m$.

   (c) *Compute* $\alpha_n^{(j)} = P(y_n | p_n^{(j)})$ *for* $j = 1, \ldots, m$.

   (d) *Generate* $f_n^{(j)}$ *for* $j = 1, \ldots, m$ *by the resampling with replacement from* $p_n^{(1)}, \ldots, p_n^{(m)}$ *with weights proportional to* $\alpha_n^{(1)}, \cdots, \alpha_n^{(m)}$.

6.3.2 Monte Carlo Smoothing Algorithm

Using the particles, $f_n^{(1)}, \cdots, f_n^{(m)}$, which approximate the filter density $p(x_n|Y_n)$, an approximation to the smoother density can be simply obtained by the following modification of the the Monte Carlo filter algorithm. Hereafter $s_{n|t}^{(j)}$ denotes the j-th independent realization from $p(x_n|Y_t)$ with $t > n$.

(d)' Generate $(s_{n|n}^{(j)}, s_{n-1|n}^{(j)}, \cdots, s_{i|n}^{(j)})^t$ for $j = 1, \ldots, m$ by the resampling of $(p_n^{(j)}, s_{n-1|n-1}^{(j)}, \cdots, s_{i|n-1}^{(j)})^t$ with weights proportional to $\alpha_n^{(1)}, \cdots, \alpha_n^{(m)}$.

Here if we put $i = 1$, then we obtain the approximation to the fixed interval smoother $p(x_t|Y_n)$ for $t = 1, \ldots, n$. However, the repetition of the above algorithm makes the number of distinct particles monotone decreasing. Since the number of distinct particles are at largest m for $\{p_n^{(j)}\}$, it is very likely that the number of distinct particles will decrease rapidly. Therefore it is recommended to stop the above smoothing algorithm after L times, namely to put $i = \max\{1, n - L\}$. According to the authors' experience, it is recommended to take $10 \leq L \leq 50$ (Kitagawa 1993). It is interesting to note that this modified smoothing algorithm yields a fixed lag smoother with lag L.

6.3.3 Likelihood of the Model

The general state space model contains several unknown parameters such as the variances of the noise. The vector consisted of these unknown parameters are denoted by θ. Given the observations $y_1, \ldots, y_N$, the likelihood of the parameter θ of the model is obtained by

$$L(\theta) = p(y_1, \ldots, y_N|\theta) = \prod_{n=1}^{N} p(y_n|y_1, \ldots, y_{n-1}, \theta) = \prod_{n=1}^{N} p(y_n|Y_{n-1}), \tag{6.12}$$

where $p(y_1|Y_0) = p_0(y_1)$. Here, by using the approximation

$$\begin{aligned} p(y_n|Y_{n-1}) &= \int p(y_n, x_n|Y_{n-1})dx_n && (6.13)\\ &= \int P(y_n|x_n)p(x_n|Y_{n-1})dx_n && (6.14)\\ &\cong \frac{1}{m}\sum_{j=1}^{m} P(y_n|p_n^{(j)}) && (6.15)\\ &= \frac{1}{m}\sum_{j=1}^{m} \alpha_n^{(j)}, && (6.16) \end{aligned}$$

the log-likelihood can be approximated by

$$\ell(\theta) = \sum_{n=1}^{N} \log p(y_n|Y_{n-1}) \cong \sum_{n=1}^{N} \log\left(\sum_{j=1}^{m} \alpha_n^{(j)}\right) - N\log m. \tag{6.17}$$

The parameter of the model, θ, can be estimated by maximizing the log-likelihood. However, since the log-likelihood obtained by (6.17) is subject to the sampling error, only a rough approximation to the maximum likelihood estimate is available by this method.

6.4 Seasonal Adjustment of Count Data

In this section, we consider non-Gaussian seasonal adjustment models. A Poisson distribution model with seasonal mean value function is considered. The model was earlier approximated by integration based Kalman filter (Schnatter 1992).

Figure 6.1 (a) shows the monthly number of bankrupted companies with capital over 100 milion yen in Japan provided from Tokyo Shoko Research. The seasonal variation and gradually changing trend is seen. We decomposed this series with the Poisson observation model (6.7). The approximate maximum likelihood of the variances of trend and seasonal system noises were $\hat{\tau}_1^2 = 0.125 \times 10^{-3}$ and $\hat{\tau}_2^2 = 0.312 \times 10^{-3}$. Since the number of bankrupted companies are small, the seasonal parttern in the original series looks variable. Figure 6.1 (b), (c) and (d) respectively show the estimated trend, seasonal factor and the observation noise. From the decomposition it can be seen that the trend increases at the end of the time interval and that the seasonal factor gradually changes with time.

To check the property of this seasonal adjustment of the discrete valued process, we also performed a simulation study by using an artificially generated Poisson process with trend and seasonal variation in the mean value function.

Figure 6.2 (b)–(d) show the intensity of the Poisson process, the assumed trend and the seasonal factors. The intensity was obtained by multiplying the trend and the seasonal factors. A realization of this non-homogeneous Poisson process is also shown in (a). Figure 6.2 (e)–(g) show the estimated trend $\exp(T_n)$, seasonal factor $\exp(S_n)$ and the residual, respectively. Visually, very good reproducton of these components are obtained. Figure 6.2 (h) shows the data obtained when the trend of the intensity is 1/10 of the original data. In the middle part, the observed data become 0 for a consecutive interval. Plots (i), (j) and (k) show estiamted trend, seasonal factor and the residuals obtained from this data. By the present seasonal adjustment procedure, the trend and seasonal factor are reasonably estimated even for such time interval.

6.5 Conclusion

A general state space model for seasonal adjustment which includes the additive and multiplicative models with Gaussian or non-Gaussian noise inputs and discrete valued process model is considered. The Monte Carlo filter and smoother which are based on the approximation of the distributions by many independent realizations, are used for the estimation of the state vector, the decomposition of the seasonal time series into trend, seasonal and irregular components, and the likelihood computation.

Caption

Figure 6.1 (a) Monthly number of bankrupted companies in Japan
(b) Estimated trend
(c) Estimated seasonal factor
(d) Residual of the fitted model

Figure 6.2 Analysis of the artificially generated count data
(a) Assumed intensity function
(b) Assumed intensity of the Poisson process
(c) Assumed trend of the intensity
(d) Assumed seasonality of the intensity
(e) Estimated trend
(f) Estimated seasonal component
(g) Residual of the model
(h) Simulated data when the trend is 1/10 of the original trend shown in (b)
(i) Estimated trend
(j) Estimated seasonal component
(k) Residual of the model

Figure 6.1: Monthly number of bankrupted companies in Japan

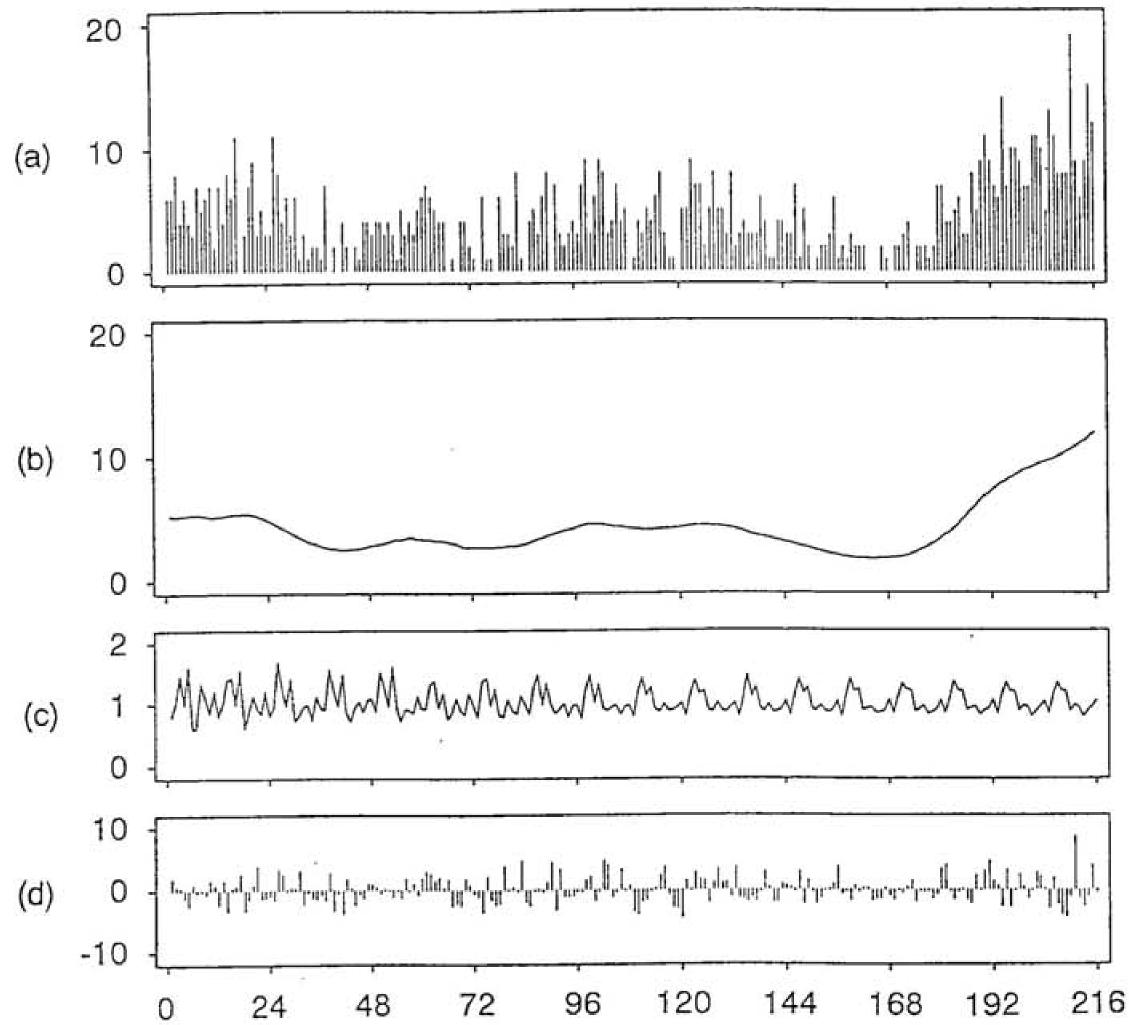
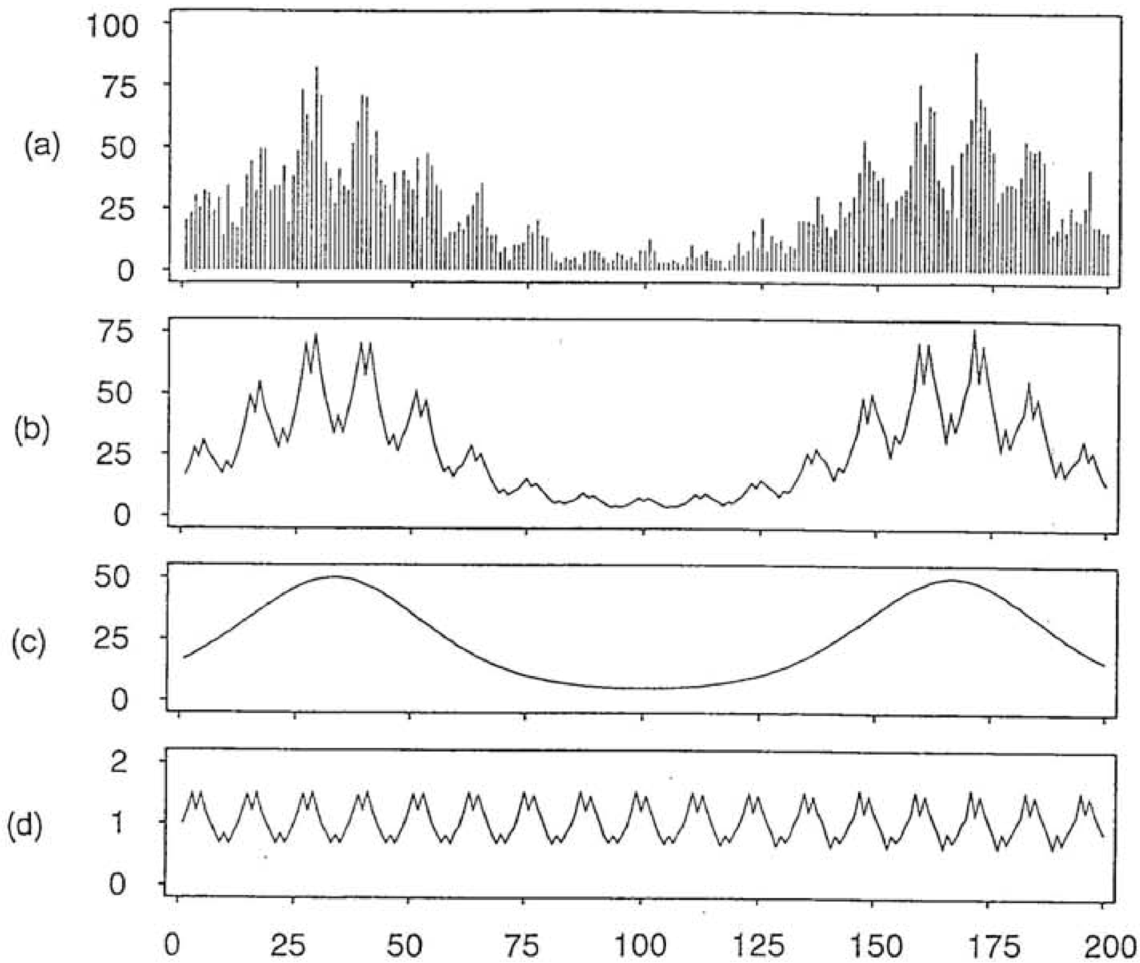
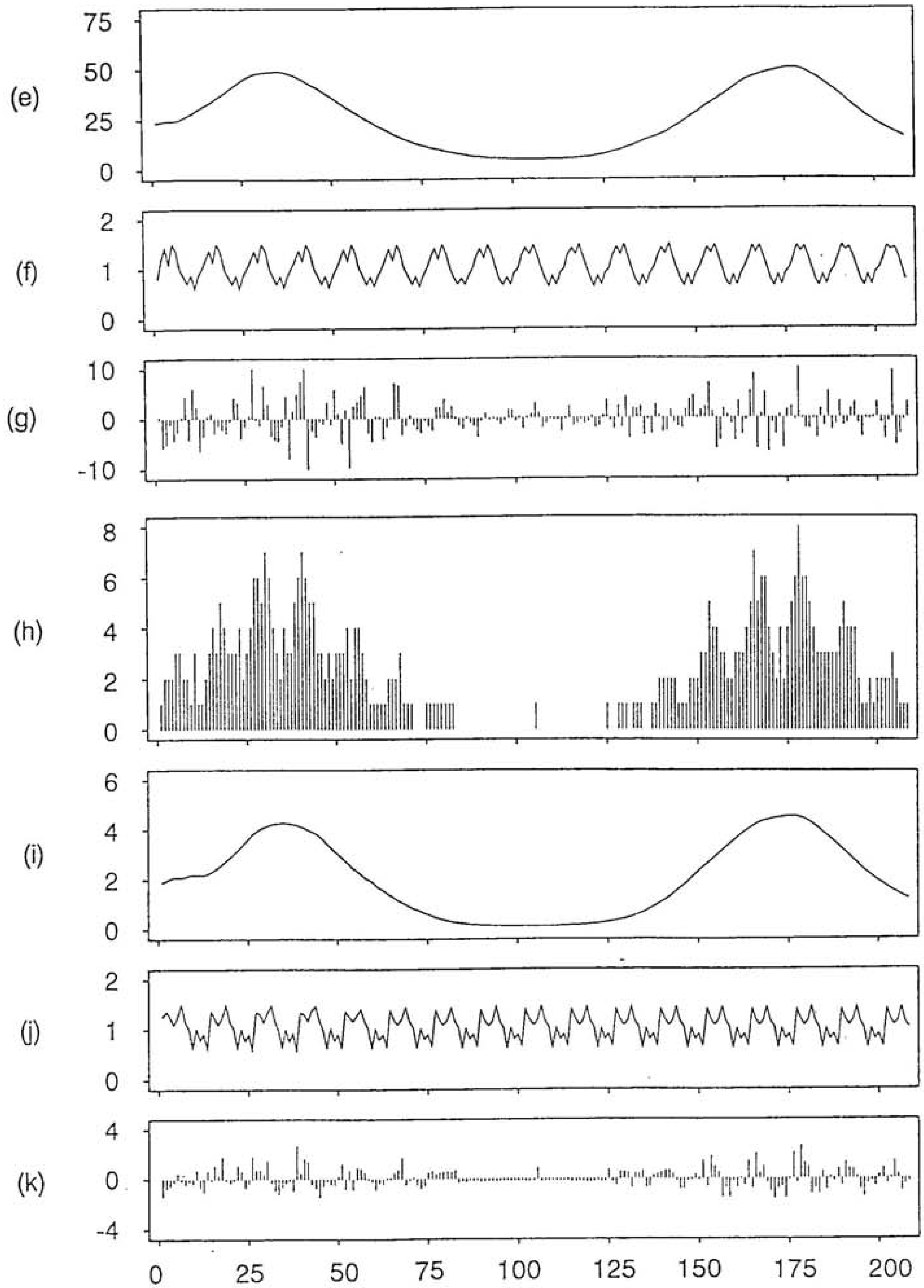


Figure 6.2: Analysis of the artificially generated count data





(e)
75
50
25
0
(f)
2
1
0
(g)
10
0
-10
(h)
8
6
4
2
0
(i)
6
4
2
0
(j)
2
1
0
(k)
4
0
-4
0 25 50 75 100 125 150 175 200

Chapter 7

Conclusion

The conclusion of this study is summarized from two aspects, financial ecomomics and statistics. From an aspect of financial economics, the summary of conclusion is the following. Firstly, the distributions of stock returns in United States and Japan are heavy-tailed and skewed. They are far from normal distributions. The type VII and IV family of Pearson system are fitted well to them. From an aspect of the risk management, Value at Risk which is based on normal distributions is smaller than one which based on actual distributions. Therefore, it is better to use heavy-tailed distributions or the mixture of normal distributions, or to multiply VaR by the adequate constant. Secondly, the asymmetry of the movement of stock markets is explained by the behavior of investors, so-called "industry-effect". This model explained the big price change like a sharp decline or thema-oriented upper trend. Thirdly, from the stationary distribution of daily stock returns, we introduced the stochastic differential equation which has heavy-tailed stationary distribution. This process is valid for the generating process of daily stock returns. Forthly, the truth that the shape parameter had been decreasing before the so-called "Black Monday crash", October in 1987, indicates that the market had been becoming gradually unstable. Therefore, from an aspect of the risk management, this model is superior to the Gaussian stochastic volatility model and the non-Gaussian stochastic volatility model which has time-invariant shape parameter. Fifthly, time series data of the bankruptcy of large companies has seasonal component. This component differs from that of small companies.

From an aspect of statistics, the summary of conclusion is the following. Firstly, we defined the type IV family of Pearson system by introducing the non-central parameter to the type VII family of Pearson system. And the analytic solutions of normalizing constants are introduced. Furthermore, by using double exponential method for the numerical integration of the normalizing constant, we put this distribution to practical use. Secondly, by using heavy-tailed and non-central distribution (the type IV family of Pearson system), we proposed the model superior to ARCH, GARCH model which use the normal distributions. Thirdly, we introduced a stochastic differntial equation whose stationary distribution is the type VII or IV of Pearson system. And in the case that the stationary distribution is the t-distribution which has even degrees of freedom, we calculate the transition probability density function. The parameters of this stochastic differential equation can be estimated by the maximum likelihood method using this transition probability density function. Furthermore, we showed that the local linearization method is useful to estimate the parameters of this stochastic differntial equaiton which correspond to the type IV (asymmetric and continuous shape parameter) and the type VII (symmetric and continuous shape parameter). Forthly, we showed the parameters of stochastic volatility model can be estimated by the non-linear non-Gaussian filter using the numerical approximation of distributions. Furthermore, we extended the observation model to the type VII family distribution of Pearson system. We developed general state space model of the non-Gaussian stochastic volatility model with the time-varying shape parameter. We obtained the non-Gaussian stochastic volatility model which can describe the big structual change superior to the Gaussian stochastic volatility model. Fifthly, we use Monte Carlo filter for seasonal adjustment to estimate the parameters of general state space model in which the observation model is given by the Poisson distribution.

Bibliography

[1] Abramowitz, M and Stegun, I.A. (1965), *Handbook of Mathematical Functions*, Dover Publications, Inc..

[2] Akaike, H. (1973), "Information Theory and an Extention of the Maximum Likelihood Principle", 2nd Inter.Symp on Information Theory (Petrov, B.N. and Csaki, F.eds.), Akademiai Kiado, Budapest, 267-281.

[3] Akaike, H. (1974), "A New Look at the Statistical Model Identification", *IEEE Trans. Autom. Contr.*, AC-19, 716-723.

[4] Anderson, B. D. O. and Moore, J. B. (1979), *Optimal Filtering*, Prentice-Hall, New Jersey.

[5] Bell, W.R. and Hillmer, S.C. (1984), "Issues involved with seasonal adjustment of economic time series", *Journal of Business and Economic Statistics*, Vol. 2, 291-349.

[6] Biscay, J. C., Jimenez, J. C., Riera, J. J. and Valdes, P. A. (1994), "Local Linearization Method for The Numerical Solution of Stochastic Differential Equations", submitted to *Annals of the Institute of Statist.*

[7] Blattberg, R. C. and Gonedes, N. J. (1974), "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices", *Journal of Business*, 47, 244-280.

[8] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

[9] Burridge, P. and Wallis, K.F. (1990), "Seasonal adjustment and Kalman filtering : extension to periodic variances", *Journal of Forecasting*, Vol. 9, No. 2, 109-118.

[10] Carlin, B., Polson, N. G. and Stoffer, D.S. (1992), "A Monte Carlo Approach to Nonnormal and Nonlinear State-Space Modeling", *Journal of the American Statistical Association*, 87, No. 418, 493-500.

[11] Clark, P. K. (1973), "A Subordinated Stochastic Process Model With Finite Variances for Speculative Prices", *Econometrica*, 41, 135-156.

[12] Dagum, E.B. and Quenneville, B. (1993), " Dynamic linear models for time series components", *Journal of Econometrics*, Vol. 55, 333-351.

[13] Danielsson, J. (1994), "Stochastic Volatility in Asset Prices: Estimation With Simulated Maximum Likelihood", *Journal of Econometrics*, 64, 375-400.

[14] Danielsson, J., and Richard, J. F. (1992), "Quadratic Acceleration for Monte Carlo Likelihood Evaluation", unpublished manuscript, University of Iceland, Dept. of Economics.

[15] Engle, Robert F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U. K. Inflation," *Econometrica*, 50, 987-1008.

[16] Fama, E. F. (1965), "The Behavior of Stock Market Prices", *Journal of Business*, 38, 34-105.

[17] Frühwirth-Schnatter, S. (1994a), "Applied state space modelling of non-Gaussian time series using integration-based Kalman filtering", *Statistics and Computing*, Vol.4, 259–269.

[18] Frühwirth-Schnatter, S. (1994b), "Data augmentation and dynamic linear models", *Journal of Time Series Analysis*, Vol. 15, 183–202.

[19] Gersch, W. and Kitagawa, G. (1983), "The Prediction of Time Series With Trends and Seasonalities", *Journal of Business and Economic Statistics*, Vol. 1, 253-264.

[20] Gersch, W. and Kitagawa, G. (1988), "Smoothness priors in time series", *Bayesian Analysis of Time Series and Dynamic Model ed J. C. Spall*, 431-476, Marcel Dekker.

[21] Girsanov, I. V. (1960), "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures", *Theory Probab. Appl.*, 5, 285-301.

[22] Gröbner, W and Hofreiter, N. (1958), *Integraltafel. Teil II, Bestimmte Integrale*, Wien und Innsbruck, Springer-Verlag.

[23] Hansmann, G.H. (1934), "On certain non-normal symmetrical frequency distributions", *Biometrika*, 26, 129-195.

[24] Harrison, P.J. and Stevens, C.F. (1976), "Bayesian Forecasting", (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 34, 1–41.

[25] Harvey, A. C. (1989), *Forecasting, structural Time Series Models and the Kalman Filter*, Cambridge University Press.

[26] Harvey, A. C. and Fernandes, C. (1989), "Time series models for count data or qualitative observations" (with discussion), *Journal of Business and Economic Statistics*, Vol. 7, 407–422.

[27] Harvey, A., Ruiz, E. and Shephard, N. (1994), "Multivariate Stochastic Variance Models", *Review of Economic Studies*, 61, 247-264.

[28] Harvey, A. and Shephard, N. (1993), "The Econometrics of Stochastic Volatility", Working Paper of London School of Economics.

[29] Heynen, R. and Kat, H.M. (1994), "Volatility Prediction: A Comparison of the Stochastic Volatility, GARCH (1,1), and EGARCH (1,1) Models", *Journal of Derivatives*, Winter, 50-65.

[30] Hillmer, S. C., Bell, W. R. and Tiao, G. C. (1983), "Modeling considerations in the seasonal adjustment of economic time series", in *Applied Time Series analysis of Economic Data*, ed. A. Zellner, U.S. Bureau of the census, 74–100.

[31] Hillmer, S.C. and Tiao, G.C. (1982), "An ARIMA-based approach to seasonal adjustment", *Journal of the American Statistical Association*, Vol. 77, No. 377, 63-70.

[32] Hodges P.E. and Hale, D. F. (1993), "A computational method for estimating densities of non-Gaussian nonstationary univariate time series", *Journal of Time Series Analysis*, Vol. 14, No.2, 163–178.

[33] Hull, J. and White, A. (1987), "Hedging the Risks from Writing Foreign Currency Options", *Journal of International Money and Finance*, 6, 131-152.

[34] Jacquier, E., Polson, N. G., and Rossi, P. E. (1994), "Bayesian Analysis of Stochastic Volatility Models", *Journal of Business and Economic Statistics*, Vol. 12, No. 4, 371-389.

[35] Johnson, N. L. and Kotz, S. (1970), *continuous univariate distributions-1, 2*, Houghton Mifflin.

[36] J.P. Morgan, (1995a), *RiskMetrics - Technical Document*, May, Third Edition.

[37] J.P. Morgan, (1995b), *Five questions about RiskMetrics*, Feb.

[38] Kalman, R. E. (1960), "A New Approach to Linear Filtering and Prediction Problems", *Journal of Basic Engineering*, 82, 1-19.

[39] Kalman, R. E. and Bucy, R. S. (1961), "New Results in Linear Filtering and Prediction Problems", *Trans. ASME, Journal of Basic Engineering*, 82, 35-45.

[40] Kariya, T., Tsukuda, Y., Maru, J., Matsue, Y. and Omaki, K. (1995), "An extensive analysis on the Japanese markets via S. Taylor's model", *Financial Engineering and the Japanese Markets*, Vol. 2, No. 1, 15-86.

[41] Kariya, T. (1993), *Quantitative Methods for Portfolio Analysis*, Kluwer.

[42] King, B. F. (1966), "Market and Industry Factors in Stock Price Behavior", *Journal of Business*, January Supplement.

[43] Kitagawa, G. (1987), "Non-Gaussian State-Space Modeling of Nonstationary Time Series", *Journal of the American Statistical Association*, 82, 1032-1063.

[44] Kitagawa, G. (1989), "Non-Gaussian seasonal adjustment". *Computers & Mathematics with Applications*, Vol. 18, No. 617, 503–514.

[45] Kitagawa, G. (1991), "A Nonlinear Smoothing Method for Time Series Analysis", *Statistica Sinica*, Vol. 1, No. 2, 371-388.

[46] Kitagawa, G. (1993), "A Monte Carlo filtering and smoothing method for non-Gaussian nonlinear state space models", *Proceedings of the 2nd U.S.-Japan Joint Seminor on Statistical Time Series Analysis*, Honolulu, January 25–29, 110–131.

[47] Kitagawa, G. (1994), "The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother", *Annals of the Institute of Statistical Mathematics*, Vol. 46, No. 4, 605–623.

[48] Kitagawa, G. (1993), *Fortran77 Programming for Time Series Analysis*, Iwanami Publishing Company, (in Japanese).

[49] Kitagawa, G. and Gersch, W. (1984), "A smoothness priors-state space modeling of time series with trend and seasonality", *Journal of the American Statistical Association*, Vol. 79, No. 386, 378–389.

[50] Kitagawa, G. and Nagahara, Y. (1995), "Monte Carlo Smoothing Method for Seasonal Adjustment", *Research Memorandum*, No.579, The Institute of Statistical Mathematics.

[51] Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices", *Journal of Business*, 36, 394-419.

[52] Merrington, M. and Pearson, E. S. (1958), "An Approximation to the distribution of Noncentral t", *Biometrika*, 45, 484-491.

[53] Mori, M. (1987), *Fortran77 Programming for Numerical Caluculation*, Iwanami Publishing Company, (in Japanese).

[54] Mortensen, R. E. (1979), "Mathematical Problems of Modeling Stochastic Dynamic Systems", *J. Statist. Physi.* , 1, 271-296.

[55] Mossin, J. (1966), "Equilibrium in a Capital Asset Market", *Econometrica*, 34, 768-783.

[56] Nagahara, Y. (1992), "Time Series Analysis of Industry Returns", Master Thesis, Tsukuba University, (in Japanese).

[57] Nagahara, Y. (1994), "Recursive Formula for Evaluating the Normalizing Constant and Cumulative Function of Type IV Family of Pearson System of Distributions", *Research Memorandum*, No.515, The Institute of Statistical Mathematics.

[58] Nagahara, Y. (1995a), "Cross-Sectional-Skew-Dependent Distribution Models for Industry Returns in Japanese Stock Market", *Financial Engineering and the Japanese Markets*, Vol. 2, No. 2, 139-154.

[59] Nagahara, Y. (1995b), "Non-Gaussian Distribution for Stock Returns and Related Stochastic Differential Equation", *Research Memorandum*, No.568, The Institute of Statistical Mathematics.

[60] Nagahara, Y. (1995c), "Note on a Comparison of Parameter Estimation Methods for a Stochastic Differential Equation", *Research Memorandum*, No.569, The Institute of Statistical Mathematics.

[61] Nagahara, Y. and Kitagawa, G. (1995d) "Non-Gaussian Stochastic Volatility Model", *Research Memorandum*, No.576, The Institute of Statistical Mathematics.

[62] Ozaki, T. (1985a), "Non-Linear Time Series Models and Dynamical Systems", *Handbook of Statistics*, Vol. 5, 25-83, edited by E. J. Hannan, et al., North-Holland.

[63] Ozaki, T. (1985b), "Statistical Identification of Storage Models with Application to Stochastic Hydrology", *Water Resources Bulletin*, 21, 663-675.

[64] Ozaki, T. (1989), "Statistical Identification of Nonlinear Random Vibration Systems", *Journal of Applied Mechanics*, 56, 186-191.

[65] Ozaki, T. (1992a), "A Bridge between Nonlinear Time Series Models and Nonlinear Stochastic Dynamical Systems: A Local Linearization Approach", *Statistica Sinica*, 2, 113-135.

[66] Ozaki, T. (1992b), "Identification of Nonlinearities and Non-Gaussianities in Time Series", in D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt, and M. S. Taqqu, eds, *New Directions in Time Series Analysis, Part I* , 227-264 Springer-Verlag.

[67] Ozaki, T. (1993), "A Local Linearization Approach to Nonlinear Filtering", *International Journal of Control*, 57, 75-96.

[68] Pearson, E. S. and Hartrey, H. O. (1954), *Biometrika Tables for statisticians, 1.*, Cambridge University Press.

[69] Pearson, E. S. (1963), "Some Problems Arising in Approximating to Probability Distributions by Using moments", *Biometrika*, 50, 95-111.

[70] Pearson, K. (1914), *Tables for Statisticians and Biometricians, 1.*, Cambridge University Press.

[71] Sage, A. P. and Mersa, J. L. (1971), *Estimation Theory with Application to Communication and Control*, McGraw-Hill, New York.

[72] Sakamoto, Y., Ishiguro, M. and Kitagawa, G. (1986), *Akaike Information Criterion Statistics*, D.Reidel.

[73] Schlicht, E. (1981), "Seasonal adjustment principle and a seasonal adjustment method derived from this principle", *Journal of the American Statistical Association*, Vol. 76, No. 374, 374–378.

[74] Schnatter, S. (1992), "Integration-based Kalman-filtering for a dynamic generalized linear trend model", *Computational Statistics & Data Analysis*, Vol. 13, 447–459.

[75] Sharpe, W. (1964), "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk", *Journal of Finance*, 19, 425-442.

[76] Shenton, L.R. and Carpenter, J.A. (1964), "The Mills Ratio and the Probability Integral for a Pearson Type IV Distribution", *Biometrika*, 52, 119-126.

[77] Shoji, I. and Ozaki, T. (1994), "Estimation for a Continuous Time Stochastic Process: A New Local Linearization Approach", *Research Memorandum*, No.524, The Institute of Statistical Mathematics.

[78] Smith, A. F. M. and West, M. (1983), "Monitoring renal transplants: an applications of the multiprocess Kalman filter", *Biometrics*, Vol. 39, 867–878.

[79] Tanizaki, H. (1993), "Nonlinear Filters", *Lecture Notes in Economics and Mathematical systems*, No. 400, Springer-Verlag.

[80] Tsay, R. S. (1986), "Time series model specification in the presence of outliers", *Journal of the American Statistical Association*, Vol. 81, 132–141.

[81] West, M., Harrison, P.J., and Migon, H.S. (1985), "Dynamic Generalized Linear Models and Bayesian Forecasting(with discussion), *Journal of the American Statistical Association*, 80, 73-97.

[82] Wong, E. (1963), "The Construction of a Class of Stationary Markoff process", *Proc. Amer. Math. Soc. Symp. Appl. Math.*, 16, 264-276.