MODELLING AND ANALYSIS OF SOME TIME SERIES

THESIS SUBMITTED TO THE

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

UNDER THE FACULTY OF SCIENCE

By

KESAVAN NAMPOOTHIRI C.

Department of Statistics
Cochin University of Science and Technology
Cochin 682 022.

MARCH 2001

CERTIFICATE

*Certified that the thesis entitled **"MODELLING AND ANALYSIS OF SOME TIME SERIES"** is a bonafide record of work done by Shri. Kesavan Nampoothiri, C under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included any where previously for award of any degree or title.*

Dr.N.Balakrishna
Lecturer in Statistics

Cochin University of Science
and Technology
March 2001

CONTENTS

CHAPTER - 1

INTRODUCTION

1.1 Introduction

A statistical data is useful only when we extract its important features. We can use those features to understand what lies behind the real data. The quantitative indicators such as mean, mode and standard deviation etc. capture useful information about the data. But usually we want more detailed information and that we must set up a statistical model for the data. This may be something like a mathematical formula that describes the probabilities of observing various data values or it may be a more complicated stochastic process, which is a mathematical system that models the actual physical process, which generates those values.

A time series is a set of observations generated sequentially in time. The primary objective of time series modelling is to develop sample models capable of forecasting, interpreting and testing hypothesis regarding data. Examples of time series are annual yield of a crop for a particular period, the population of a country during a specified time, the number of births of babies in a hospital according to the hour at which they were born. The time series has an important place in the field of economics and business statistics. The time series relating to prices, consumptions, money in circulation, bank deposits and bank clearing, sales and profit in a departmental store, national income and foreign exchange reserves, prices and dividend of shares in a stock exchange etc. are examples of economic and business time series.

The original use of time series analysis is to provide an aid to forecasting. As such methodology was developed to decompose a time series into trend, seasonal, cyclical and irregular components. An important feature of time series is that the successive observations are usually dependent. When successive observations are dependent future values may be predicted from the past observations. If the future values of time series can be predicted exactly it is said to be a deterministic time series. But in most of time series the future is only partially determined by its past values. Such a time series is known as stochastic time series. In that case exact prediction is not possible and therefore the future values have a probability distribution, which is conditional by knowledge of past values. In that case the model can be written as

$$X_n = f(n) + \varepsilon n \qquad , n=1,2,3,\ldots p,$$

where X_n, n=1,2,3, . . . p are observations on time series made at p equally distant time points, f (n) is called the systematic part and $\{\varepsilon_n\}$ is the random or stochastic sequence , it obeys a probability law and is called the innovation process. There are five sections in this chapter and the details of each section are as follows. The Section 1.2 gives some basic definitions, 1.3 is a brief description of the non-linear time series models, 1.4 describes the non-Gausssian time serics and Section 1.5 is a summary of the thesis.

1.2 SOME BASIC DEFINITIONS

1.2.1 Stochastic Process

A stochastic phenomenon that evolves in time according to some probabilistic law is called a stochastic process. That is , a stochastic process is a family of random variables $\{X_n , n \in T\}$ defined on the probability space (Ω, f, P).

The time series can be regarded as a particular realization of stochastic process. Time series analysis is primarily an aid of specifying the most likely stochastic process that could have generated an observed time series. A model that can be used to calculate the probability of a future value is called the stochastic model or the probability model.

1.2.2 Stationary process

The estimation of the parameters of a stochastic process will not be possible if they change as time progresses. The most practical models will be those whose parameters are constant over time. This will happen when the finite dimensional distribution of $\{X_n\}$ does not depend on the time. A stochastic process $\{X_n\}$ is said to be strictly stationary if the joint distribution of $X_{n1}, X_{n2}, \ldots, X_{np}$ made at time points $n_1, n_2, \ldots, n_p$ is same as that associated with n observations at $X_{n1+k}, X_{n2+k}, \ldots, X_{np+k}$ made at time points $n_{1+k}, n_{2+k}, \ldots, n_{p+k}$ for every k. A stochastic process whose mean is constant and variance is finite and covariance between X_n and X_s

is a function of |n-s| is said to be second order stationary or weakly stationary.

1.2.3 Autocovariance and Autocorrelation functions

Let $\{X_n\}$ be a stochastic process, the covariance between X_n and X_{n+k} is known as the autocovariance at lag k and is defined by

$$\gamma_k = Cov(X_n X_{n+k})$$
$$= E(X_n X_{n+k}) - E(X_n)E(X_{n+k}).$$

The correlation coefficient between two random variables Xn and Xn_{+k} obtained from a stationary process {Xn} is called autocorrelation function (ACF) at lag k and is given by

$$\rho_k = \frac{Cov(X_n X_{n+k})}{\sqrt{Var(X_n).Var(X_{n+k})}} = \frac{\gamma_k}{\gamma_0} \quad ,$$

therefore $\rho_0 = 1$, $\rho_k = \rho_{-k}$ and $-1 \le \rho_k \le 1$.

1.2.4 White Noise Process

A sequence of uncorrelated random variables with zero mean and constant variance is called a white noise process.

1.2.5 Gaussian Process

A sequence of random variables $\{X_n\}$ defining a stationary process can have any probability distribution. A stationary process $\{X_n\}$ is called a Gaussian process if the joint distribution of $(X_{n+1}, X_{n+2, \ldots} X_{n+k})$ is a k-variate normal for every positive integer k. Now we consider some standard time series models, which are frequently used.

1.2.6 Autoregressive Process

One of the most useful and simplest models used in time series modelling is the autoregressive model of appropriate order. A sequence of random variables $\{X_n, n \geq 0\}$ is said to follow an autoregressive process of order p or AR (p) if it can be written in the form

$$X_n = \phi_1 X_{n-1} + \phi_2 X_{tn2} + \ldots + \phi_p X_{n-p} + \varepsilon_n, \qquad (1.2.1)$$

where $\{\varepsilon_n\}$ is a sequence of independent and identically distributed random variables with zero mean and variance $\sigma_e^{\,2}$. The AR (1) sequence $\{X_n\}$ is stationary if it satisfies the condition that $|\phi_1| < 1$ (for details refer Box and Jenkins, 1970).

1.2.7 Yule-Walker equation

An important recurrence relation for estimating the parameters of an AR (p) model is due to Yule-Walker (see Box and Jenkins, (1970)). They give the autoregressive parameters in terms of autocorrelations. Multiplying both sides of equation (1.2.1) by X_{n-k} and taking expectations we get a difference equation in autocovariance. Also, the autocorrelation function satisfies the same form of difference equation.

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p} \qquad , k=1,2,3,\ldots.$$

The variance of an AR(p) process is

$$\sigma^2 = \frac{\sigma_\varepsilon^{\ 2}}{1-\rho_1\phi_1-\rho_2\phi_2-\ldots-\rho_p\phi_p}.$$

1.2.8 Moving Average Process

A q[th] order moving average process {Xn} is defined by

$$X_n = \varepsilon_n - \theta_1\varepsilon_{n-1} - \theta_2\varepsilon_{n-2} - \ldots - \theta_q\varepsilon_{n-q},$$

where $\{\varepsilon_n\}$ is a sequence of independent and identically distributed random variables with zero mean and variance $\sigma^2{}_\varepsilon$.

The AR model can be generalized to the integrated AR models as follows.

1.2.9 Autoregressive Integrated Moving Average (ARIMA) process

Many empirical time series do not have homogeneous stationary behaviour. In such cases the stationary behaviour can be obtained by taking suitable differences. The models for which d[th] difference is stationary is a mixed autoregressive integrated moving average process and is given by

$$\phi(B)\nabla^d X_n = \theta(B)\varepsilon_n$$

where ,

$$\theta(B) = 1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q,$$

$B^p X_n = X_{n-p}$, $B^q\varepsilon_n = \varepsilon_{n-q}$ and

$$\nabla^1 X_n = X_n - X_{n-1},$$

$$\nabla^2 X_n = (X_n - X_{n-1}) - (X_{n-1} - X_{m2}) = X_n - 2X_{n-1} + X_{n-2}$$

1.2.10 Box-Jenkins Modelling Techniques

Box and Jenkins (1970) proposed a three-stage method for selecting an appropriate model for the purpose of estimating and forecasting a univariate time series. We can describe the stages as follows.

Identification stage

In this stage we visually examine the plots of the series, the autocorrelation function and the partial autocorrelation function (PACF). If the variance of the series changes along with the time, a logarithmic transformation will often suitable for the changes in the variance. If the series or its appropriate transformation is not stationary then the next step is to determine the proper degree of differencing. For that we can use the plot of the time series, plot of sample ACF, sample variance of the successive differences etc. The last step in the identification stage is to determine the values of the order 'p' and 'q'. It can be obtained by comparing the sample ACF and PACF with theoretical patterns of known models. The values of 'p' and 'q' are usually small. After identifying a tentative model it is necessary to estimate its parameters by suitable methods.

Estimation Stage

A detailed discussion on this is given in Box and Jenkins (1970).

Diagnostic Stage

Once we identify a model and its parameters are estimated, the next step is the diagnostic stage. That is, to verify whether the selected model satisfies the assumptions. If the assumptions are not satisfied, continue the above steps till a good model is obtained. After identification of a good model for a given set of data, it can be used for forecasting.

Forecasting

There are various forecasting methods available depending on the structure of the time series model. A good reference is Box and Jenkins (1970).

In practise, some of the basic assumptions, especially the linearity and the normality of the series, of standard Box Jenkins methodology are not satisfied. Therefore, recently there has been a growing interest in studying non-linear and non-normal time series models. The following sections provide an introduction to those time series models and a detailed study on some of these models are presented in the subsequent chapters.

1.3 Non-Linear Time Series Models

A linear time series model is often adequate in one step ahead prediction. However, a linear differential equation is totally inadequate as a tool to analyse more intricate phenomena such as limit cycles, time reversibility, amplitude frequency dependency etc (Tong, 1980). The non-linear time series modelling gives a more detailed understanding of the data. Tong has given a detailed discussion of the merits and demerits of the linear Gaussian models. Here we describe some of the non-linear models and later we use these models to analyse a set of data.

1.3.1 Threshold Autoregressive (TAR) models

The concept of a threshold is the local approximation over the state, that is the introduction of regimes namely thresholds. The thresholds allow the analysis of complex system by decomposing it into simpler sub systems. A time series $\{X_n\}$ is said to follow TAR process if,

$$X_n = \phi_0^{(j)} + \sum_{i=1}^{n} \phi_i^{(j)} X_{n-i} + \varepsilon_n^{(j)}, \qquad \text{if } r_{j-1} < X_{n-d} < r_j \quad ,$$

where j=1,2...k and d is a positive integer, k is the number of regimes and d is the delay parameter. These models allow the autoregressive coefficients to change over time and the changes are determined by comparing the previous values back shifted by a time lag d.

1.3.2 Random Coefficient autoregressive (RCA) models

The idea of multiplicative noise may be further extended to the class of RCA models. A time series $\{X_n\}$ is said to follow a RCA model of order k if X_n has the form

$$X_n = \sum_{i=1}^{k} (\beta_i + B(n)) X_{n-1} + \varepsilon_n$$

where $\{\varepsilon_n\}$ is a sequence of independent and identically distributed (iid) random variables with zero means and variance σ^2_i , β_i, $i=1,2,\ldots k$ are constants, $\{B(t)\}$ is a sequence of 1 x k vectors with zero mean and

$E[B^T(n)B(n)] = C$,

the term $B^T(n)$ is the transpose of the vector $B(n)$.

1.3.3 Bilinear Models

Bilinear models lie somewhere between fixed coefficients autoregressive models and random coefficient models. A time series $\{X_n\}$ is said to follow a bilinear model if it satisfies the equation

$$X_n + \sum_{i=1}^{p} a_i X_{n-i} = \alpha + \sum_{j=1}^{r}\sum_{k=1}^{s} b_{jk} X_{n-j}\varepsilon_{n-k} + \varepsilon_n$$

where $\{\varepsilon_n\}$ is a sequence of iid random variables usually but not always with zero mean and a, b and α are real constants.

1.3.4 Autoregressive Models with Conditional Heteroscedasticity (ARCH)

A sequence $\{X_n\}$ is said to follow an ARCH model if X_n is of the form

$$X_n = \varepsilon_n \sqrt{h_n}$$

where $\{\varepsilon_n\}$ are iid random variables with standard normal distribution and

$$h_n = \gamma + \phi_1 X_{n-1}^2 + \phi_2 X_{n-2}^2 + \ldots + \phi_p X_{n-p}^2 ,$$

where $\gamma_i \geq 0$, $\phi_i \geq 0$ for all i. We can see that $\{Xn^2\}$ follows a bilinear model if {Xn} follow an ARCH model. If we write the above as

$$h_n = \gamma + \sum_{i=1}^{q} \phi_i X_{n-i}^2 + \sum_{i=1}^{p} \psi_i h_{n-i} ,$$

where $\varphi \geq 0$ for all i, then $\{X_n\}$ is said to follow a generalized ARCH model or GARCH model.

In chapter 2 we consider the applications of two non-linear models viz. TAR and ARCH models to analyse a set of real data.

1.3.5 Heteroscedasticity

The assumption of constant variance of the disturbance term of a regression equation is not always valid. For example, the variance of food expenditure among families may increase as family income increases. Similarly the variance of public spending may increase with city size. Heteroscedasticity is the formal name for the case of non-constant variance of the disturbance term. In applied research, heteroscedasticity is usually associated with data. Consider a regression model

$$Y_n = \alpha_0 + \alpha_1 X_n + \varepsilon_n , \qquad n = 1,2,3 \ldots N$$

Then the heteroscedasticity assumption is

$$E(\varepsilon_n^2) = k_i^2 \sigma^2 ,$$

for all i

1.3.6 Financial Time Series

The fluctuations in financial markets attract our attention frequently . Daily reports on news papers, television and radio inform us the variation in the stock markets, currency exchange rates and gold prices etc. It is often desirable to monitor the price behaviour frequently and try to understand the probable development of price fluctuations. Suppose we planned a holiday abroad and we need to purchase some currency, we may look at the latest exchange rates from time to time and try to forecast them. We call the series of prices thus obtained as financial time series.

The first objective of the price studies is to understand how prices behave. That is such a complex subject , for that we have to look into the distribution of the actual prices. Tomorrow's price is uncertain and it must therefore be described by a probability distribution. The second objective is to use our knowledge of price behaviour to take better decisions. Decisions based on better forecasts are profitable in trading commodities. Forecasts of the variance of the future price changes are very helpful in assessing prices at the relatively new option markets. This innovation leads to the development of suitable methods for analysing financial time series. Here in this thesis we consider the applications of ARCH models.

1.4 Non-Gaussian Time Series Models

Recently a considerable amount of work appears in non-Gaussian time series models. The search for such time series models arises form the fact that many of the naturally occurring time series are clearly non-

Gaussian. The method for analysing time series proposed by Box and Jenkins (1976) assume Gaussianity. Similarly the basic assumption in the non-linear models proposed by Tong(1983) also use Gaussian assumptions. However, most of the empirical time series are far from Gaussian. Some of the non-Gaussian time series models introduced in the literature are by Graver and Lewis (1980), Lawrence and Lewis (1985) and Tavares (1980). A bibliography on non- Gaussian time series is given by Balakrishna (1999). The rest of this section gives a small review of non-Gaussian time series.

The non-Gaussian time series provides stationary sequences having non-normal marginal random variables. One of the basic problems in non-Gaussian time series is to identify the innovation distribution for a specified marginal (Balakrishana, 1999). In the case of Gaussian models both Xn and ε_t have normal distributions whereas it is not the case in non-Gaussian models. Adke and Balakrishna (1992) have studied non-negative random variables having exponential and Gamma distributions. They studied the properties such as mixing , time reversibility and estimation problems for EAR(1) and NRAR(1) processes. Jayakumar and Pillai (1993) introduced Mittag-Leffler process; Abhraham and Balakrishna (1999) introduced inverse Gaussian AR process. Similarly other AR(1) models are available with marginal distributions such as Logistic, exponential and Laplace. The AR models with infinite variance innovation is studied by Cline and Brockwell (1985) and Brockwell and Davis (1987).

Generating functions such as Laplace transform and characteristic functions are the tools used for finding solutions for AR models. But if the generating functions do not have closed forms it is difficult to find these solutions. Another important non-Gaussian process is the autoregressive minification process. This process with marginal distributions such Weibull, logistic etc. are studied by many authors. Balakrishna and Jayakumar (1997, 1997a) have studied multivariate versions of non-Gaussian models for certain distributions like Pareto, semi-Pareto and exponential. An important problem involved is the estimation of the parameters. Now we explain the definitions of some of the probabilistic properties of a time series which are useful in studying the properties of the estimators. This is followed by a summary of the thesis.

1.4.1 Ergodic Sequences

A sequence $\{X_n\}$ of r.vs is stationary and ergodic if $Pr\{(X_0,X_1,X_2, \ldots)\in A\}$ is either zero or one whenever A is a shift invariant event.

1.4.2 Mixing Properties

The strong mixing properties for a sequence of random variables is useful as a tool in establishing central limit theorems. In the context of time series, the asymptotic normality of various estimators can be established by assuming the strong mixing properties of the series. We can define the strong mixing property as follows.

Let $\{X_n\}$ be a sequence of random variables in the probability space $(\Omega,\mathcal{B},P)$. Then $\{X_n\}$ is said to be strong mixing if

$\alpha(m) = \sup | P(A \cap B) - P(A).P(B) |, \rightarrow 0 \;\; n \rightarrow \infty$, when the supreme is taken over all $A \in F_0^n$, $B \in F_{n+m}^{\infty}$, where F_0^n and F_{n+m}^{∞} are the are the minimal sigma fields induced by $(X_0, X_1, \ldots, X_n)$ and $(X_{n+m}, X_{n+m+1}, \ldots)$ respectively.

1.4.3 Harris Recurrent Markov Cahin

A Markov chain $\{X_n\}$ is Harris recurrent if there exits a non-trival σ-finite measure $\varphi(.)$ on (S,δ) such that $\varphi(E)>0$ implies that $P_x[X_n \in E$, for some $n \geq 1]=1$ for all x in S where P_x refers to the probability measure corresponding to the initial condition $X_0=x$.

1.4.4 Time Reversibility

A stationary time series $\{X_n\}$ is said to be time reversible if for every k and every $n_1, n_2, \ldots, n_k$, $\{X_{n1}, X_{n2}, \ldots, X_{nk}\}$ and $\{X_{-n1}, X_{-n2}, \ldots, X_{-nk}\}$ have the same joint probability distributions. Otherwise $\{X_n\}$ is said to be time irreversible.

1.5 Summary of the Thesis

In this thesis we consider some of the non-linear Gaussian and non-Gaussian time series models and mainly concentrated in studying the properties and application of a first order autoregressive process with Cauchy marginal distribution. The major part of the thesis is devoted to Cauchy AR (1) process. The main objective here is to identify an appropriate model to a given set of data. The data considered are the daily coconut oil prices for a period of three years. Since it is a price data the consecutive prices may not be independent and hence a time series based model is more appropriate. It is well known that the price data usually follow heavy tailed distributions. One of the important distributions to study the price behaviour is the Cauchy distribution. The chapter-wise summary is as follows.

The second chapter discusses mainly the non-linear Gaussian time series models. There are three main sections in this chapter. The first section discusses the application of a threshold autoregressive (TAR) model. Here we try to fit a TAR model to a time series data. This model was introduced by Tong(1980). Because of the complexities of the method proposed by Tong , it is not widely used in practice. Tsay (1989) proposed a simultaneous method for testing the non-linearity and identification of the delay parameter. Here we essentially follow the method proposed by Tsay(1989). This Section explains the methodology used for the analysis followed by a detailed analysis of the data. The fitted model is compared with simple autoregressive model. The results are in favour TAR process. Another important non-linear Gaussian model discussed in this chapter is the ARCH model introduced by Engle(1982). This Section discusses the importance of this model followed by the definition and the modelling technique. Here also we mainly concentrate on the applications of the ARCH model. A discussion of the an empirical data analysis is also included here. The third important non-linear model we discussed here is the TARCH models, that is threshold models with ARCH effect. This threshold plus ARCH effect has many applications in modelling financial time series. Here we discusses the definition of the model followed by a real data analysis.

The chapter 3 is the most important part of the thesis, where we define a first order autoregressive process with one-dimensional Cauchy marginal distribution. The first Section contains an introduction to the chapter, while the second Section gives the definition, the innovation distribution and the joint distribution of 'n' consecutive random variables of the process. This Section also discusses the properties like ergodicity, mixing property and time reversibility. The rest of this chapter discusses various estimation procedures used to estimate the unknown parameters of the process. The maximum likelihood estimation is discussed in section 3.3. Since the likelihood equations do not have closed form for their solutions, we obtained mle by Newton-Raphson method. The estimators are consistent and asymptotically normal under certain regularity conditions. Therefore this is followed by the verifications of the regularity conditions. Since some of the regularity conditions do not hold when both of the model parameters are unknown, we assume that one is known and verify the conditions. Here also we find some problems when AR coefficient is unknown. Therefore we go for an alternative method of estimation. The alternative method of estimation is discussed in the Section 3.4. Here we use the method proposed by Brockwell and Davis (1987) for estimating the AR coefficient. The scale parameter is estimated using an empirical distribution function method. The asymptotical properties of the estimators are also discussed in this Section

The chapter 4 discusses the application of the Cauchy AR(1) model introduced in the previous chapter. The first section is a simulation study to investigate the performance of the estimators and the second section is a real data analysis. This section explains how we arrive at this model. The daily coconut oil prices at Cochin market for period of three years is used for the analysis. The importance of this commodity, its characters, nature etc. are discussed followed by the estimation of the parameters using different methods.

CHAPTER-2

SOME NON-LINEAR GAUSSIAN TIME SERIES MODELS

2.1 INTRODUCTION

Linearity is one of the basic assumptions in the classical analysis of time series by Box-Jenkins methodology. But non-linearity can often be detected in time series. There are several types of non-linear time series models proposed by Tong (1990), among those we studied the applications of some of these models. In this chapter we consider some of the non-linear Gaussian time series models. Section 2.2 discusses the definition, properties along with an empirical analysis of a Threshold Autoregressive (TAR) model, Section 2.3 gives the application of Autoregressive Conditional Heteroscedastic models (ARCH) and the Section 2.4 discusses Threshold Autoregressive Conditional Heteroscedastic (TARCH) models.

2.2 THRESHOLD AUTORESSIVE MODELS

The idea of threshold autoregressive models (TAR) was introduced by Tong (1980a). The essential idea underlying the class of threshold AR models is the piece-wise liberalization of non-linear models over the state space by the introduction of the thresholds. These models are locally linear. Similar ideas were used by Priestely(1965). Priestely and Tong(1978) and Ozaki and Tong(1975) in the analysis of non-stationary time series and time dependent systems, in which local stationary was the counterpart of the local linearity. The local linearity has an important role in practical situations. For example, Tong (1980a) has adopted piece-wise linear models in the analysis of Canadian Iynx data and Wolf Sunspot numbers.

Motivated by the complex behaviour of the solutions of non-linear systems, Tong(1990) has introduced a class of time series models which could reproduce some of the features of these solutions. In threshold autoregressive models, different autoregressive models may operate and the changes between the various autoregressions is governed by threshold values and a time lag. These models have been reviewed by many researchers and compared with classical time series models with respect to data sets such as Wolf's Sunspot numbers and Canadian lynx data. Tsay (1989) proposed a simultaneous method for testing the non-linearity and the identification of the delay parameter. Here we essentially follow the steps proposed by Tsay (1989) and compare it with a simple autoregressive (AR) model. In the following sections the TAR modelling technique is briefly described followed by the results and discussions.

2.2.1 Definition

A time series $\{X_n\}$ is TAR process if it follows a model of the form

$$X_n = \phi_0^{(j)} + \sum_{i=1}^{l} \phi_i^{(j)} X_{n-i} + \varepsilon_n^{(j)} \qquad \text{if } r_{j-1} < X_{n-d} < r_j, \qquad (2.2.1)$$

where j=1,2...k, k is the number of regimes, with the regimes being separated by k-1 threshold values $r_j (r_0 = -\infty; r_k = +\infty)$, $d \in N^+$ is the delay parameter ($d \leq p$), $\{a_0^{(j)}, a_2^{(j)}\}$, i=1,2....p, j=1,2...k are the model parameters regime j and $\{\varepsilon_n^{(j)}\}$, j=1,2...k are sequences of independent normal variates with zero mean and variance $\sigma^2_{\varepsilon j}$.

The procedure proposed by Tong(1980) is complex. It involves several computing stages and there was no diagnostic statistic available to assess the need for a threshold model to a given set of data. Tsay(1989) proposed a procedure for testing the threshold non linearity and building, if necessary, a TAR model. The procedure consists of the following steps.

Step1- Select the order 'p' of the autoregression and the set of possible threshold lags 's'.

Step2- Fit an arranged autoregressive model for a given 'p' and perform the threshold non-linearity test. If non-linearity of the process is detected, select the delay parameter d_p.

Step3- For a given 'p' and 'd_p' locate the threshold values using the scatter plots.

Step4- Refine the AR order and threshold values, if necessary, in each regime by using linear autoregression techniques.

The AR order 'p' in 'step 1' may be selected by considering the autocorrelation function (ACF) and partial autocorrelation function

(PACF) or some information criteria like Akaike information Criteria (AIC) as described in Enders (1995).

2.2.2 Tests for non-linearity.

Before estimating a TAR model, it is necessary to detect specific non-linear behavior in the series by using an appropriate test. Classical non-linearity tests based on maximum likelihood are complicated as the likelihood function is not differentiable with respect to the unknown threshold values r_j (Tong, 1990). Several researchers have proposed methods for testing these types of non-linearity. For example see Tong and Lim 1980, Kennen 1985, Tsay 1986, Petrueelli and Davis, 1986. Here we prefer the test proposed by Tsay (1989) for the reasons stated above. It is fairly simple and widely applicable. Its asymptotic distribution under the linear manipulation is the classical F-distribution. The procedure is as follows:

Consider an example of TAR (2,p,d), which consists of two regimes and one threshold value r_1. Assume the order of the autoregression is 'p' in each regime and the delay parameter is equal to 'd'. Then the model can be written as:

$$X_n = \phi_0^{(1)} + \sum_{v=1}^{p} \phi_v^{(1)} X_{n-v} + \varepsilon_n^{(1)} \qquad \text{if } X_{n-d} \le r_1$$

$$= \phi_0^{(2)} + \sum_{v=1}^{p} \phi_v^{(2)} X_{n-v} + \varepsilon_n^{(2)} \qquad \text{if } X_{n-d} > r_1. \qquad (2.2.2)$$

where $n \in \{p+1, \ldots, l\}$, l being the number of observations and other parameters are defined as before. Now arrange the observations in the ascending order. Let π_i be the time index of i[th] smallest observation, then the above model can be written equivalently as

$$X_{\pi_i+d} = \phi_0^{(1)} + \sum_{v=1}^{p} \phi_v^{(1)} X_{\pi_i+d-v} + \varepsilon_{\pi_i+d}^{(1)} \qquad \text{if } i \le s$$

$$= \phi_0^{(2)} + \sum_{v=1}^{p} \phi_v^{(2)} X_{\pi_i+d-v} + \varepsilon_{\pi_i+d}^{(2)} \qquad \text{if i>s}, \qquad (2.2.3)$$

with $i \in \{p+1,\ldots,m-d\}$ and s satisfying $X_{\pi_s} < r_1 \le X_{\pi_s+1}$. This is an arranged autoregression with the first 's' cases in the first regime and the rest in the second regime. The arranged autoregression provides a mean by which the data points are grouped so that all the observations in a group follow the same AR model. The separation does not require the precise value of r_1; it only requires that the number of observations, in each group depends on r_1.

Tsay described the motivation for the test as follows. If one knew the threshold value r_1, then the consistent estimator of the parameters could easily be obtained. Since the threshold values are unknown, one must proceed sequentially. The least squares estimates of the $\hat{\phi}_v^{(1)}$ of $\phi_v^{(1)}$ is consistent if there are large numbers of observations in the first regime (ie. many $i \le s$). In this case, the predictive residuals are white noise asymptotically and orthogonal to the regressors $\{X_{\pi_i+d-v} \mid v = 1,2,\ldots,p\}$. On the other hand, when 'i' arrives at or exceeds 's' the predictive residual for the observation with time index $\pi_s + d$ is biased because of the model changes at time $\pi_s + d$. That is, the predictive residual is a function of the regressors $\{X_{\pi_i+d-v} \mid v = 1,2,\ldots,p\}$. Consequently the orthogonality between the predictive residuals and the regressors is destroyed once the recursive autoregression goes on to the observation whose threshold value exceeds r_1. Based on the above, one way to test the non-linearity is to regress the predictive residuals of the arranged autoregression (2.1.3) on the regressors $\{X_{\pi_i+d-v} \mid v = 1,2,...,p\}$, and use the F-statistic of the resulting regression. The F-statistic is defined in (2.2.6) below.

Consider the arranged autoregression (2.1.3), let $\hat{\beta}_m$ be the vector of least squares estimates based on the first 'm' cases, P_m be the associated $X'X$ inverse matrix and X_{m+1}, the vector of regressors of the next observation to enter the autoregression, namely $X_{\pi_{m+1}+d}$. These vectors and matrices are given below. Then the recursive least squares estimation of the parameters can be done using the following algorithm given by Ertel and Fowlkes (1976).

Here

$$\hat{\beta} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ . \\ \phi_p \end{bmatrix} \qquad \text{and } X = \begin{bmatrix} 1 & X_{\pi_{p+1}+d-1} & X_{\pi_{p+1}+d-p} \\ 1 & X_{\pi_{p+2}+d-1} & X_{\pi_{p+2}+d-p} \\ . & & . \\ 1 & X_{\pi_{n-d}+d-1} & X_{\pi_{n-d}+d-p} \end{bmatrix}.$$

$$\hat{\beta}_{m+1} = \hat{\beta}_m + K_{m+1}[X_{\pi_{m+1}+d} - X'_{m+1}\hat{\beta}_m]$$

$$D_{m+1} = 1 + X'_{m+1}P_mX_{m+1}$$

$$K_{m+1} = \frac{P_m X_{m+1}}{D_{m+1}} \quad \text{and}$$

$$P_{m+1} = \left(I - P_m \frac{X_{m+1}X'_{m+1}}{D_{m+1}} \right) P_m$$

and the predictive residual is given by

$$e\hat{}_{d+\pi m+1} = \frac{X_{d+\pi_{m+1}} - X'\hat{\beta}_m}{\sqrt{D_{m+1}}} \qquad (2.2.4)$$

In the above equations 'I' denotes an identity matrix of appropriate order. The predictive residuals can also be used to locate the threshold values by using various scatter plots. For fixed 'p' and 'd' the effective number of observations in the arranged autoregression is l-p. Assume that the recursive estimation begins with $m = \frac{l}{10} + p$ observations so that there are (l-p-m) predictive residuals available. The test statistic proposed Tsay defines the classical F-.statistic as below. Corresponding to the regression of the predictive residuals (recursively estimated) of the arranged autoregression on the regressors (1, $X_{\pi i+d-1} \ldots X_{\pi i+d-p}$). That is, if

$$\hat{e}_{\pi_i+d} = \omega_0 + \sum_{v=1}^{p} \omega_v X_{\pi_i+d-v} + \varepsilon_{\pi_i+d} \qquad (2.2.5)$$

for i = b+1,...n-p and then compute the F-statistic as

$$\hat{F}(p,d) = \frac{(\sum e^2{}_n - \sum \varepsilon^2{}_n)/(p+1)}{\sum \varepsilon^2{}_n /(n-2p-b-1)} \qquad (2.2.6)$$

The summations are over the observations in (2.2.4) and $\varepsilon\hat{}_n$ is the least squares residual of (2.2.5). The above statistic follows approximately an F-distribution, which stated in the following Lemma proved by Tsay (1989).

Lemma 2.2.1: Suppose that Z_t is a linear stationary AR process of order 'p'. That is, X_n follows model (2.2.1) with k=1. Then for large n, the statistic F(p,d) defined in (2.2.6) follows approximately an F-distribution with p+1 and (l-2p-b-1) degrees of freedom (d.f). Further more, (p+1) F(p,d) is asymptotically a chi-squared random variable with (p+1) d.f

The relative power, feasibility and simplicity are the major considerations in proposing the above statistic. Also since it requires only a sorting routine and the linear regression method, it can be easily implemented. The next step is the identification of TAR model in the estimation of the delay parameter and threshold values.

2.2.3 Identification of the delay parameter.

The threshold variable plays a key role in the non-linear nature of the model. For model (2.2.1) the specification amounts to the selection of the delay parameter d. Tong and Lim (1980) used AIC for the selection of d after selecting all other parameters (threshold values and AIC coefficients). Tsay(1989) proposed a different method, that is to identify the delay parameter 'd' and then the threshold values. For a given 'p' the delay value d_p to be chosen from {1,2...p} as follows:

$$d_p = \max_{1 \le \delta \le p}\{\hat{F}(p,\delta)\},$$

where $\hat{F}(p,\delta)$ is the statistic defined by (2.2.6). That is, d_p is the value that maximizes $\hat{F}(p,\delta)$.

2.2.4 Identification of the threshold values

A graphical method is used to locate the threshold values. Two scatter plots are used for this purpose.

(1) The scatter plot of predictive residuals of (2.2.4) versus X_{n-d_p}. A non-random change will be observed at the threshold values, since the predictive residuals will be biased at the thresholds. It is closely related to the traditional on-line residual plot for quality control. It shows the locations of the threshold values directly.

(2) The scatter plot of the t-ratios of recursive estimate of an AR coefficient versus X_{n-d_p}, where the t-ratio is given by

$$t = \frac{\hat{\beta}_{m+1}}{\sqrt{RSS * R(I,I)}} ,$$

RSS denote the mean residual sum of squares and R(I,I) is the I^{th} diagonal element of $(X'X)^{-1}$. In this case, the t-ratios have two functions: (a) they show the significance of that particular AR coefficient, and (b) when the coefficient is significant the t-ratio gradually and smoothly converge to a fixed value as the recursion continues. To explain the use of the second scatter plot to identify the threshold values, consider a simple TAR models with a single threshold given by

$$\begin{aligned} X_n &= \phi^{(1)} X_{n-1} + \varepsilon_n^{(1)} \qquad \text{if } X_{n-d} \le r_1 \\ &= \phi^{(2)} X_{n-1} + \varepsilon_n^{(2)} \qquad \text{if } X_{n-d} > r_1 . \end{aligned}$$

The t-ratios behave exactly as that of a linear time series before the recursion reaches r_1. Once r_1 is reached the estimate of $\phi^{(1)}$ starts to change and t-ratio begins to deviate (see Tsay, 1989). The pattern of the gradual convergence of the t-ratio starts to turn and changes direction at the threshold value. This behaviour of the t-ratio is used to identify the value of the threshold .

2.2.5 Empirical Example

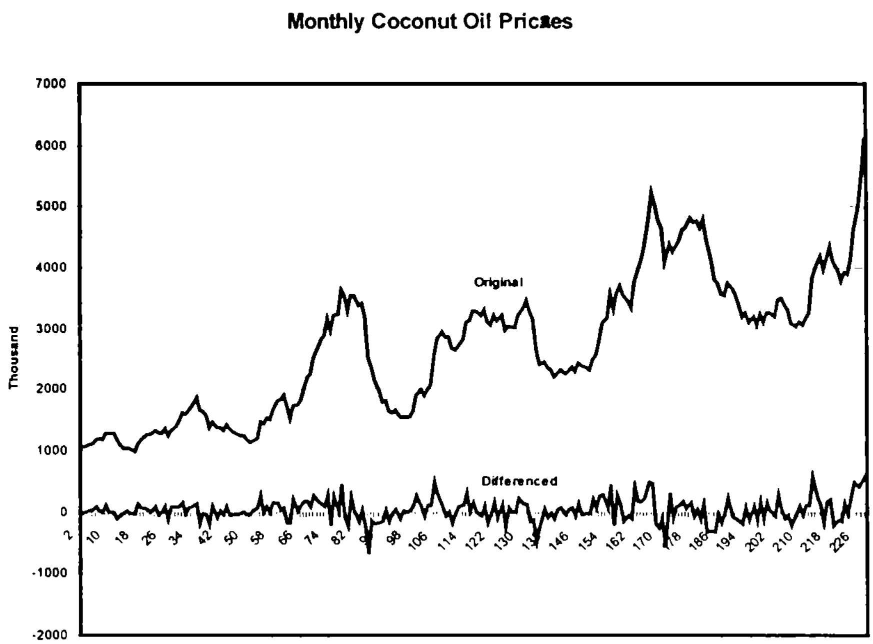
In this section we apply the above procedure to a set of real data. The data used in this study consists of the monthly coconut oil prices for a period from January 1978 to December 1996, which is presented in Appendix-I. The series consists of 228 observations. The Fig. 2.1 shows an upward trend in the process during the period. Apart from the sharp increase, fluctuations in prices within the year can also be seen.

Since the observed prices arise in a time sequence, it is possible that the consecutive observations are dependent. Therefore a time series model based approach has been tried to explain the fluctuations other than trend and seasonal variations. A fairly good estimate of the parameters of the series is obtained only if the series is stationary. Plotting of the original data shows that it is not stationary. Therefore we take a first order difference of the prices (that is, if X_n is the price sequence then their first order difference is $\nabla X_n = X_n - X_{n-1}$, n=1,2,...) for further analysis (Fig.2.1 , given below).

Firstly we try to model the prices using the Univariate Box Jenkin's (UBJ) method and then using the threshold AR method. In the

UBJ technique a model can be fitted to data by studying the behaviour of the characteristics such as ACF and PACF or by using some Information criteria like AIC. After identifying order and nature of the relationships, the model parameters are to be estimated. These models can be used for short term forecasting, because most of the autoregressive models place emphasis on recent past rather than its distant past. The ACF and PACF converge to zero reasonably quickly (Fig.2.2). The cutoff of the PACF after the lag two (Fig.2.2) recommends an autoregressive process of order two for the series. Also an examination of the AIC and residual sum of squares (RSS) for different orders of 'p' and 'q' (Table 2.1) suggests an AR(2) is more appropriate for the series. After fitting a UBJ model a non-linear model was also fixed to get a better representation.

Fig 2.1



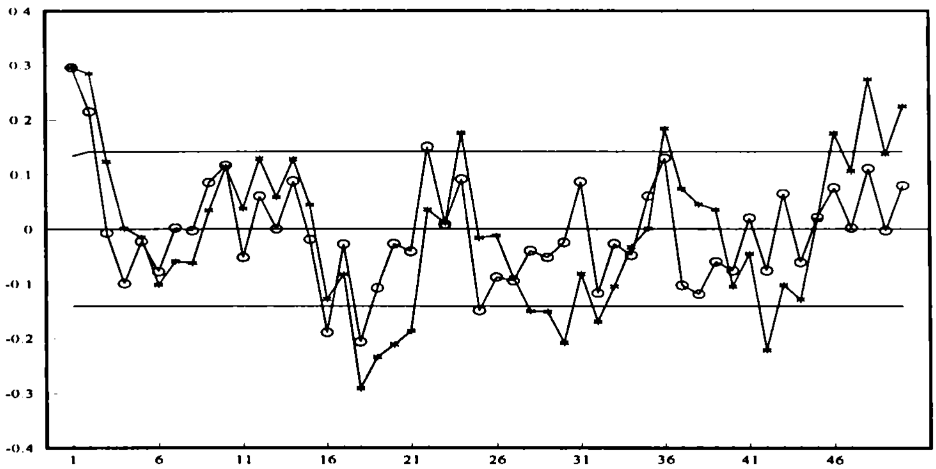


Fig 2.2

ACF and PACF of First Order Difference

Table 2.1 : Estimates of the parameters of ARMA(p,q)

	p=1 q=0	P=2 q=0	p=1 q=1	p=2 q=1	p=2 q=2
Constant	22.77 (16.37)	24.89 (21.16)	25.10 (21.31)	24.90 (21.23)	22.36 (16.92)
AR(1)	-0.308 (0.065)	0.243 (0.066)	0.688 (0.136)	0.247 (0.286)	1.446 (0.217)
AR(2)	-	0.236 (0.067)	-	0.235 (0.108)	-0.616 (0.163)
MA(1)	-	-	0.401 (0.166)	0.004 (0.293)	1.193 (0.217)
MA(2)	-	-	-	-	-0.575 (0.129)
AIC	2981	2306	2974	2973	2971
RSS	659277	6262937	6362139	6243104	6169479
ErrorVariance	4833	27844	28381	27699	27753

AIC – Akaike Information Criteria, AR(p)–Autoregressive process of order 'p',RSS–Residual Sum of Squares MA(q)–Moving Average process of order 'q'

The first step in the Tsay (1989) procedure is to identify the order of the AR process. From the above details we can choose the order as two. Therefore the possible values of the delay parameters are either d=1 or d=2. The next step is to test the non-linearity using the statistic (2.2.6). Recursion starts with 25 observations, so that there are 200 predictive residuals. The values of the F-statistic are given in Table 2.2. The p-value

Table 2.2 Estimates of the autoregressive parameters for TAR (3,2,1) and AR (2)

	Results for			
	TAR (3,2,1)			AR(2)
Constant	-45.11	10.88	67.67	24.89
1	.1088	0.0297	0.3419	0.243
	(0.145)	(0.118)	(0.092)	(0.066)
2	-0.00418	(-0.123)	0.0169	0.236
	(0.146)	(0.121)	(0.125)	(0.067)
AIC	528	702	1022	2306
RSS	2168589	1044825	3191737	6262937
Residual Variance	44257	14313	32569	27844

Values F-statistic

d	F(δ,d)
1	1.81
2	1.20

is maximum for d=1. After identifying the delay parameter the next step is to locate the threshold value using the t-ratios of the recursive estimates of an AR coefficient versus ∇X_{n-dp}. The scatter diagram reveals the threshold value directly. The t-ratios of the estimates behaves exactly as those of a linear time series before the recursion reaches the threshold value r_1. Once r_1 is reached, t-ratio begins to deviate. The pattern of the gradual convergence of the t-ratio is destroyed. In effect, the t-ratio starts to turn and, perhaps, changes direction at the threshold value. The scatter plot (Fig 2.3, below) of the t-ratios indicate the possible threshold values are around -100, 40 and 100. Since it needs a minimum of 50 observations for accurate parameter estimation, we choose the value as -100 and 40, that is, threshold model with three regimes. There are 51, 74 and 100 observations in the first, second and third regimes respectively. Since AIC is minimum for p=2 and r_1=-100 and r_2=40 we choose the order of AR as two and the threshold values as -100 and 40. The parameters are estimated for all the models (Table 2.2). The identified TAR (3,2,1) is as follows.

$$\begin{aligned}\nabla X_n &= -45.11 + 0.1088\nabla X_{n-1} - 0.00418\,\nabla X_{n-2} + \varepsilon^{(1)}{}_n \quad \text{if } \nabla X_{n-1} \leq -100\\ &= 10.88 + 0.0297\nabla X_{n-1} + 0.12370\nabla X_{n-2} + \varepsilon^{(2)}{}_n \quad \text{if } -100 < \nabla X_{n-1} \leq 40\\ &= 67.69 + 0.3419\nabla X_{n-1} - 0.3250\nabla X_{n-2} + \varepsilon^{(3)}{}_n \quad \text{if otherwise.}\end{aligned}$$

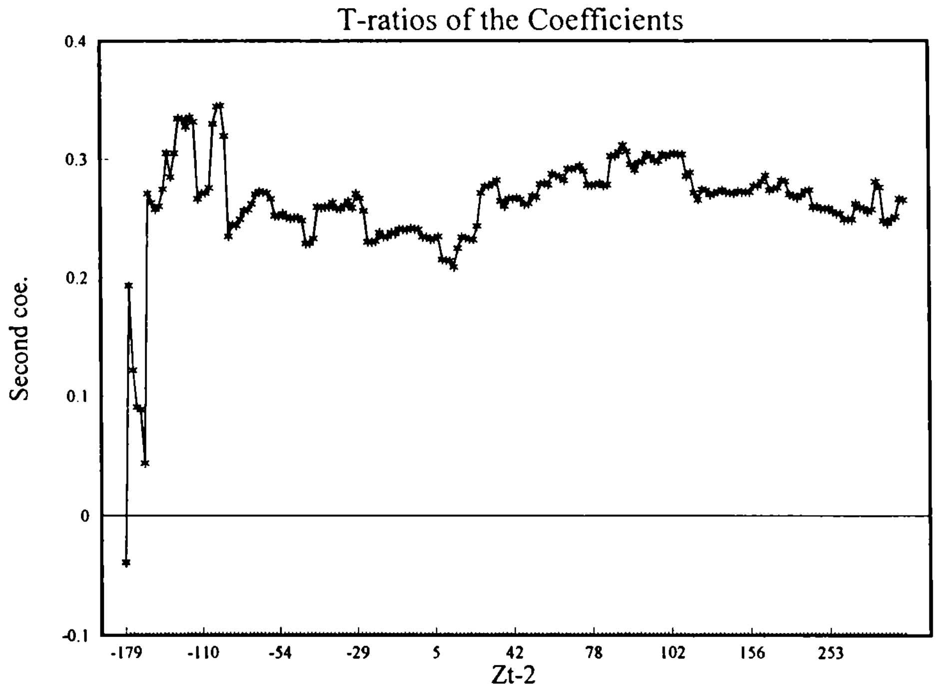
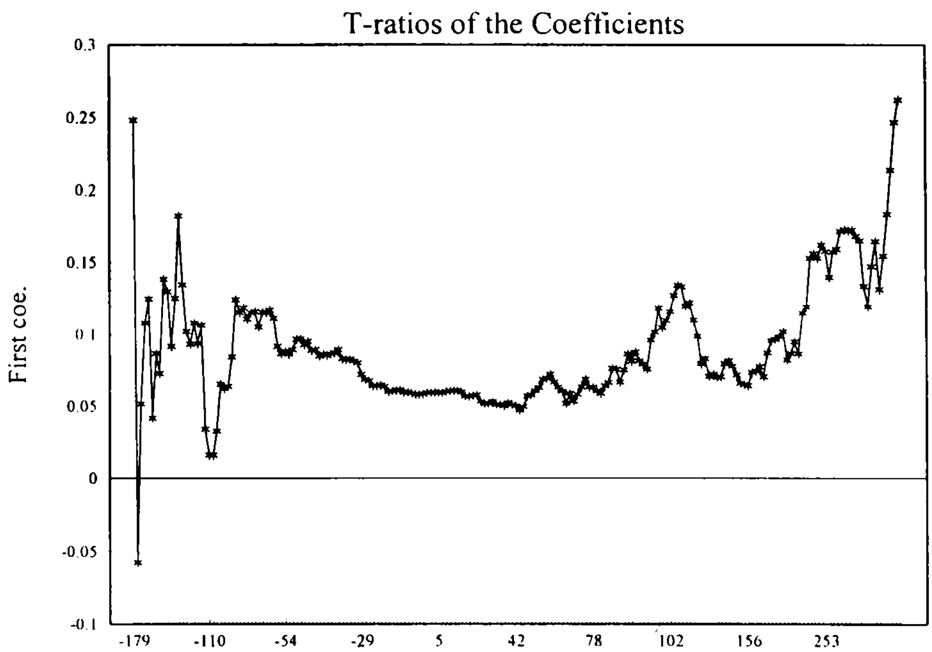

T-ratios of the Coefficients
Second coe.
0.4
0.3
0.2
0.1
0
-0.1
-179
-110
-54
-29
5
42
78
102
156
253
Zt-2

Fig 2.3


T-ratios of the Coefficients
First coe.
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-179
-110
-54
-29
5
42
78
102
156
253
Zt-1

In the diagnostic stage, we compute the ACF of the residuals for each of the models. Most of ACF are out of the 2σ limit shows that the residuals are independent. The sum of squared residuals and AIC values are less for TAR model than those for an AR model (see Table 2.2). The forecast percent error (Fig 2.4) is also minimum for the TAR model. These observations are in favour of modelling the series by a TAR process.

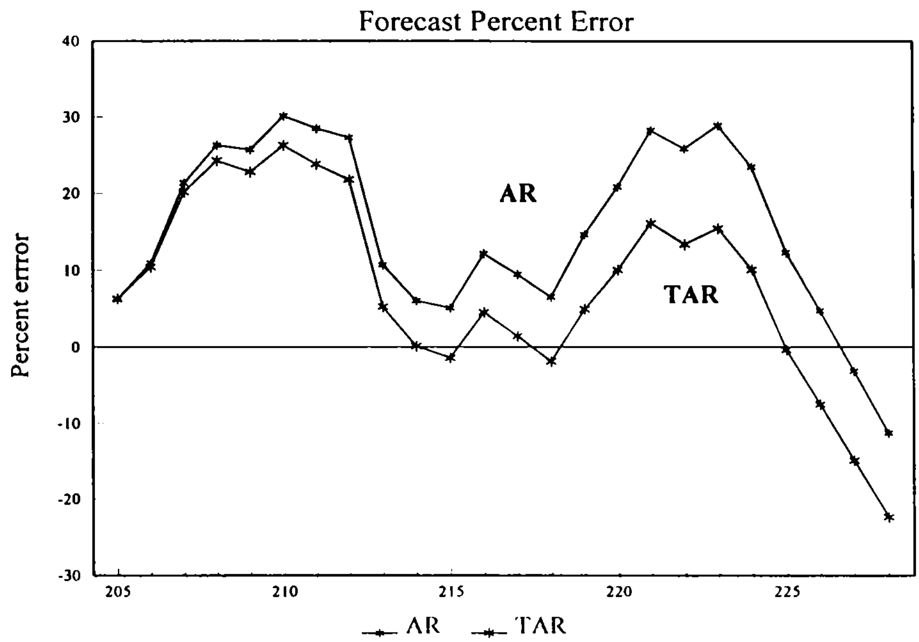


Fig 2.4

Here the values of the F-statistic do not show any non-linearity in the series. But no other factors like the RSS, AIC etc are in favour of TAR process. The percent forecast error for TAR process is lower than that of an AR process. Thus most of the factors are in favour of modeling the coconut oil prices using a TAR process. The TAR process gives a better fit for coconut oil prices. For an observation X_n, the model change is identified by using the difference of previous two observations.

2.3 AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTIC MODELS

One of the basic assumptions in the classical Box-Jenkins methodology is that, the variance of the error random variable is a constant. But most of the financial and economic time series usually exhibit the characteristic feature that the variance at time 'n' is some varying function of the variances at times (n-1), (n-2), . . . Recently, most of the economic research is concerned with extending the Box-Jenkins methodology to analyze this type of time series behavior. One of the most important tool in characterizing such changes in variance is the autoregressive conditional heteroscedastic (ARCH) model introduced by Engle(1982). A stochastic variable with constant variance is called homoscedastic and varying variance is heteroscedastic. A brief description of homoscedasticity and heteroscedasticity is given in Chapter 1.

The prices of commodities, stock market indices, stock returns etc. appear to vary through time according to some probabilistic laws. If the time series of stock returns consists of independent and identically distributed random variables, the process is called random walk process. In the early 1960's, the random walk model was favorite for modelling financial data. Since then, the independent and identically distributed nature has been challenged by many researches, for example see the references in Li (1995) in the field of finance. In random walk models, it is difficult to predict the direction of the future return by using the past return, the future magnitude is more predictable (see Li (1995)). This invalidates the assumptions of the random walk hypothesis and points out the necessity of modelling time series, which have changing conditional variances.

The ARCH model proposed by Engle (1982) becomes widely acceptable for financial time series with conditional heteroscedasticity. If the series exhibits periods of very large volatility followed by periods of relative tranquility (Enders, 1995), the assumption of constant variance becomes inappropriate. Also forecasting will be meaningful, if we can forecast the future prices along with their variances. If ARCH effect is present, ordinary method of fitting an ARMA model to the time series lead to ineffective estimates and sub optimal inference (Bollerslev et al., 1992). Here in this part , the autoregressive nature of the monthly coconut prices is studied by taking into account the ARCH effect present in the series.

2.3.1 Description of ARCH model

There are several models for changing variances and covariances. One approach to forecast the variance is to introduce an independent (exogenous) variable that helps to predict the volatility. Consider the simplest case in which $X_n=\varepsilon_n Y_{n-1}$, where X_n is the variable of interest , ε_n is a white noise process with $E(\varepsilon_n)=0$ and $Var(\varepsilon_n) = \sigma^2$ and Y_{n-1} is the independent variable observed at time 'n-1'. The conditional variance of X_n is $\sigma^2 Y^2_{n-1}$, which depends on the realized values of Y_n. If the magnitude of Y^2_{n-1} is large (small) the variance of X_n will be large (small) as well. Further more, if the successive values of $\{Y_n\}$ exhibit positive serial correlation, the conditional variances of X_n also follows positive serial correlation. In this way the introduction of the independent variable can explain the periods of volatility. The procedure is also simple to implement. A major difficulty in this strategy is that it requires the specification of the changing variance. Also we may not have theoretical reason for selecting one candidate for the Y_n sequence over the other reasonable choices. The bilinear model given in (1.3.3) also allows conditional variance to depend on the past realization of the series. The model is $X_n = \varepsilon_n X_{n-1}$ and the conditional variance is $\sigma^2 X^2_{n-1}$. A similar model, not exactly the same but very close to the bilinear model was introduced by Engle(1982). He showed that it is possible simultaneously to model the mean and variance of the series. Before getting into the details of the model, we shall explain some of the importance of the conditional forecasts. To explain this, let us consider a AR(1) model defined by

$$X_n = \phi_0 + \phi_1 X_{n-1} + \varepsilon_n \qquad |\phi_1|<1,$$

and suppose that the parameters are already estimated.

A forecast of X_{n+1} is given by

$$E(X_{n+1} \mid X_n) = \phi_0 + \phi_1 X_t .$$

If we use the conditional mean to forecast X_{n+1}, the forecast error variance is

$$Var(X_{n+1} \mid X_n) = E(X_{n+1} - \phi_0 - \phi_1 X_n)^2 = E(\varepsilon_{n+1}^2) = \sigma^2$$

Instead, if unconditional forecast are used, then

$$E(X_n) = \frac{\phi_0}{1-\phi_1}$$

and the unconditional forecast error variance is

$$Var(X_n) = E(X_n - \frac{\phi_0}{1-\phi_1})^2$$

$$= E(\varepsilon_n + \phi_1 \varepsilon_{n-1} + \phi_1^2 \varepsilon_{n-2} + \ldots)^2$$

$$= \frac{\sigma^2}{1-\phi_1^2}.$$

Since $\frac{1}{1-\phi_1^2} > 1$, the unconditional forecast has a greater variance than that of a conditional forecast. Thus conditional forecasts are superior to unconditional forecasts in terms of their variances. The model proposed by Engle (1982) is

$$X_n = \varepsilon_n \sqrt{h_n} , \tag{2.3.1}$$

where $\{\varepsilon_n\}$ is a white noise process with $E(\varepsilon_n)=0$ and $Var(\varepsilon_n)=1$ and

$$h_n = \alpha_0 + \alpha_1 X^2_{n-1} , \tag{2.3.2}$$

α_0 and α_1 are constants such that $\alpha_0>0$ and $0< \alpha_1$. Also assumes that ε_n and X_{n-1} are independent of each other and ε_n follows a standard normal distribution for each t. Let $\varphi_n = \{X_j, j \le n\}$ be the past history of $\{X_n\}$ up to time n. It is referred as the information set up to n. Also assumes the conditional distribution of X_n given φ_{n-1} is normal with mean zero and variance h_n. This is an ARCH process of order one. The properties of an ARCH process are discussed by Engle (1982). The conditional variance follows an autoregressive process. In order to ensure that the conditional variance is positive it is necessary to assume that the unknown parameters α_0 and α_1 are positive. Thus $\{X_n\}$ is a zero mean serially uncorrelated process with non-constant unconditional variance and constant conditional variance. Also it generates a data with fatter tails than the normal density as it has the coefficient of kurtosis given by

$$\gamma = \frac{E(Xn^4)}{[E(X_n^{\ 2})]^2}$$

$$= \frac{3(1-\alpha_1^{\ 2})}{(1-3\alpha_1^{\ 2})}, \quad \text{if } 3\alpha_1<1 .$$

Note that $\gamma>3$.

The simplest and often useful ARCH model is the first order linear model given by (2.3.1) and (2.3.2). The generalization of the first order linear ARCH model is given by Engle(1982). The model is defines as

$$X_n = \varepsilon_n \sqrt{h_n} \ ,$$

where $\{\varepsilon_t\}$ is a standard normal variable and

$$h_n = \alpha_0 + \alpha_1 X^2{}_{n-1} + \alpha_2 X^2{}_{n-2} + \ldots + \alpha_p X^2{}_{n-p},$$

assume that the conditional distribution of X_n is normal with mean zero and variance h_n. This is an ARCH process of order 'p' or ARCH (p) process. The following theorem gives a set of conditions for stationarity of an ARCH(p) process.

Lemma 2.3.1 : The p^{th} order linear ARCH process with $\alpha_0 > 0, \alpha_1, \ldots \alpha_p \geq 0$, is covariance stationary if and only if all the roots of the associated characteristic equation lie outside the unit circle. The stationary variance is given by

$$E(X_n^{\,2}) = \frac{\alpha_0}{1-\sum_{j=1}^{p}\alpha_j}.$$

Proof : See Engle (1982).

The technique of constructing an ARCH process explained by Enders (1995) is as follows.

Estimate the best fitting ARMA model to the sequence $\{Y_n\}$ and obtain the squares of the residuals $\hat{\varepsilon}_n^{\,2}$. Calculate the sample variance of the residuals $\hat{\sigma}^2$ as

$$\hat{\sigma}^2 = \frac{\sum_{1}^{T}\hat{\varepsilon}_n^{\,2}}{T},$$

where T = number of residuals. Obtain the sample autocorrelations of the squared residuals as

$$\rho(\,i\,) = \frac{\sum_{n=i+1}^{T}(\hat{\varepsilon}_n^{\,2}-\hat{\sigma}^2)(\hat{\varepsilon}_{n-i}^{\,2}-\hat{\sigma}^2)}{\sum_{n=i+1}^{T}\hat{\varepsilon}_n^{\,2}-\hat{\sigma}^2)^2}.$$

In large samples, the standard deviation of ρ(i) can be approximated by $\frac{1}{\sqrt{T}}$. Individual values of ρ(i) with a value that is significantly different form zero is an indicative of ARCH effect. The Lung-Box Q statistic, given by

$$Q = T(T+2)\sum_{i=1}^{n}\frac{\rho(i)}{(T-i)}, \qquad (2.3.3)$$

can be used to test for groups of significant coefficients (Enders, 1995). This statistic Q has an approximate chi-square distribution with n (total number of observations) degree of freedom if $\hat{\varepsilon}^2{}_n$ are uncorrelated. Rejecting the null hypothesis that the $\hat{\varepsilon}^2{}_n$ are uncorrelated is equivalent to rejecting the null hypothesis of no ARCH effects. In practice consider the value of up to $n = \frac{T}{4}$.

The Lagrange multiplier test procedure proposed by Engle (1982) may be described as follows. Consider an AR(p) model defined by

$$X_n = a_0 + a_1 X_{n-1} + \ldots\ldots\ldots\ldots + a_p X_{n-p} + \varepsilon_n \ .$$

Obtain the squares of the residuals of the error and denote it by $\hat{\varepsilon}^2{}_t$. Regress these squared residuals on a constant α_0 and on the p lagged values, $\hat{\varepsilon}^2{}_{n-1},\ldots\hat{\varepsilon}^2{}_{n-p}$. That is, obtain the estimate as

$$\hat{\varepsilon}^2{}_n = \alpha_0 + \alpha_1 \hat{\varepsilon}^2{}_{n-1} + \ldots + \alpha_p \hat{\varepsilon}^2{}_{n-p}\,.$$

If there is no ARCH effect then $\alpha_1 = \alpha_2 \ldots \alpha_p = 0$. Obtain the statistic TR^2 where R^2 is the usual coefficient determination. With a sample of T residuals, under the null hypothesis of no ARCH effect, the test statistic TR^2 converges to a chi-square distribution with p degrees of freedom. Therefore, rejection of H_0 is equivalent to say that there is no ARCH effect. Or if TR^2 is sufficiently small, it is possible to conclude that there is no ARCH effect.

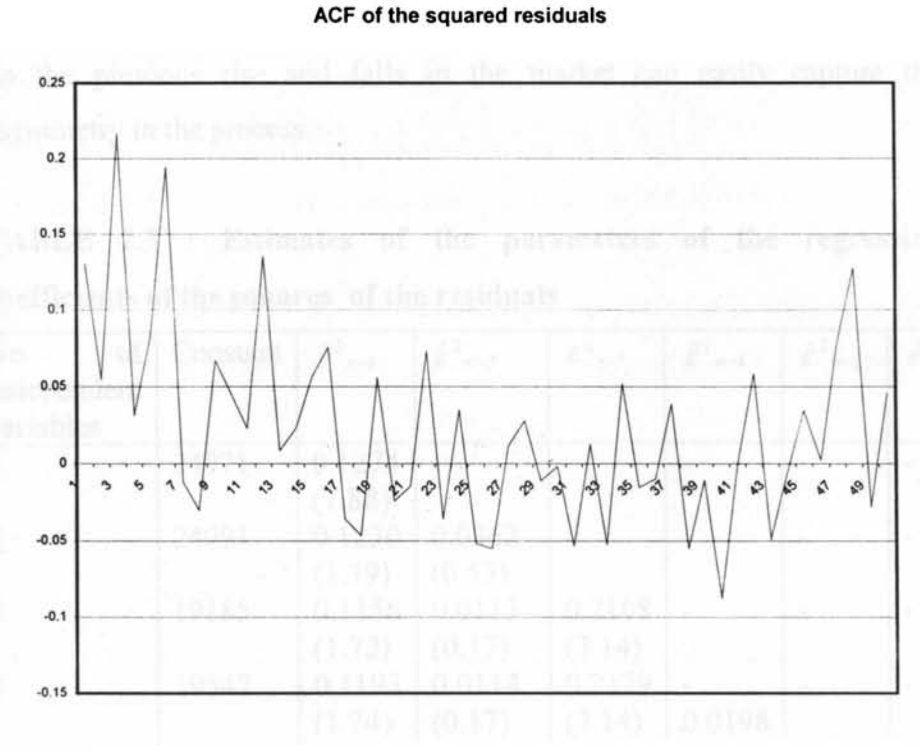
Empirical Analysis

The time series data of the monthly coconut oil prices at Cochin Market, described in the previous section is used for the analysis. The data shows that the process undergoes wide and violent fluctuations (Fig 2.1). Also we can observe periods of high variability followed by relatively smaller ones.

The prices have increased nearly four times during the period (1978-96). It is below the average up to 1987 while it fluctuates around the means form December 1987 to July 1990 and after that it never comes down below the average. The actual data set is provided in the Appendix I. Similarly the variance also undergoes fluctuations. These variations in the means and variances of the process lead to test the presence of ARCH effect in the series. Using the Box-Jenkins procedure, an autoregressive process of order two, (that is, AR (2) is found suitable for the series (details are given in the previous section).

The above modelling procedure is based on the assumption that the error variance is a constant. This may not true always. Therefore, the next step is to check whether there is any ARCH effect present in the series. The significant autocorrelation coefficient of the squared residuals is shown in Fig 2.5 given below. Since the calculated value of the Q-statistic (Q=180.75 for n=50) is greater than the table value (76.15 at 1% level of significance) we reject the hypothesis of no ARCH effect. The Table 2.3 given below gives the values of the regression coefficients and the corresponding TR^2 values. We estimated the values up to lag six. The significant value of TR^2 is obtained when the number of independent variables is six. Since the regression coefficients corresponding to $\hat{\varepsilon}^2_{n-2}, \hat{\varepsilon}^2_{n-4}$ and $\hat{\varepsilon}^2_{n-5}$ are very low

Fig 2.5



(also the value is negative for $\hat{\varepsilon}^2_{n-4}$) we omit those squared errors and continued the procedure. The value of the TR^2 has not much reduced even when there are three independent variables and the t-value of those coefficients are significant also. These coefficients satisfy the stationarity conditions (Lemma 2.3.1) of a p^{th}-order ARCH process. Thus finally the model is

$$\nabla X_n = 24.89 + 0.243 \nabla X_{n-1} + 0.236 \nabla X_{n-2} + \varepsilon_n ,$$

(21.16) (0.066) (0.067)

where ε_n is assumed to be normally distributed with means zero and conditional variance given by

$$h_n = 16384 + 0.10431\hat{\varepsilon}^2{}_{n-1} + 0.1811\hat{\varepsilon}^2{}_{n-3} + 0.1529\hat{\varepsilon}^2{}_{n-6}.$$

The numbers in the parentheses are the standard errors. By conditioning on the previous rise and falls in the market can easily capture the asymmetry in the process.

TABLE 2.3 : Estimates of the parameters of the regression coefficients of the squares of the residuals

No of independent variables	Constant	$\hat{\varepsilon}^2{}_{n-1}$	$\hat{\varepsilon}^2{}_{n-2}$	$\hat{\varepsilon}^2{}_{n-3}$	$\hat{\varepsilon}^2{}_{n-4}$	$\hat{\varepsilon}^2{}_{n-5}$	$\hat{\varepsilon}^2{}_{n-6}$	TR^2
1	24971	0.1274 (1.88)	-	-	-	-	-	3.60
2	24091	0.1230 (1.79)	0.0362 (0.53)	-	-	-	-	3.80
3	19185	0.1156 (1.72)	0.0113 (0.17)	0.2108 (3.14)	-	-	-	13.50
4	19547	0.1193 (1.74)	0.0114 (0.17)	0.2129 (3.14)	- 0.0198	-	-	14.85
5	18120	0.1206 (1.76)	0.3462 (0.05)	0.2125 (3.14)	(-0.28) -	0.077 (1.11)	-	19.58
6	15572	0.1086 (1.59)	0.00016 (0.002)	0.1831 (2.68)	0.0288 (-0.41)	0.059 (0.85)	0.1529 (2.21)	19.36
3	16384	0.1043 (1.58)	-	0.1811 (2.69)	- 0.0286 (- 0.411) -	-	0.1596 (2.33)	18.68

Figures in parenthesis denote the student t-value.

2.4 THRESHOLD AUTOREGRESSIVE CONDITIONAL HETEROSCADASTIC MODEL (TARCH)

This section deals with threshold models having ARCH effect. That is, extends the usual threshold specification to take care of the ARCH effect. This threshold specification was briefly mentioned in Tong (1990). The threshold plus ARCH specification has many application is modelling financial time series. The threshold ARCH models and Asymmetries in Volatility introduced by Rabemananjara and Zakoian (1993) was used to model the asymmetry in the conditional variance. This model allows relaxing the possible constraints in the parameters of the conditional variance and also this unconstrained model allows nonlinearity in the volatility. The threshold model with conditional heteroscedasticity, for X_t is defined as

$$X_n = \phi^{(j)}{}_0 + \sum_{v=1}^{p} \phi^{(j)}{}_{n-v} X_{n-v} + \varepsilon_n^{(j)} \qquad \text{if } r_{j-1} < X_{n-d} < r_j$$

$$j=1,2,\ldots k,\ v=1,2,\ldots,p,$$

where ε_n is assumed to be normally distributed with mean zero and conditional variance

$$h_n^{(j)} = \alpha_0^{(j)} + \alpha_1^{(j)} \varepsilon_{n-1}^{\ 2} + \ldots + \alpha_r \varepsilon_{n-r}^{\ 2}$$

where $\alpha_i \geq 0$, i =0, 1, 2.... r. The above model is a TARCH (p_1, p2, r). If $p_1=p_2=p_3\ldots\ p_t$ and $\phi_2^{(1)}=\phi_i^{(2)}$ for all i, then we have the usual autoregressive model with conditional heteroscedasticity that is TARCH (p,r). Here h_n is changing in each regime according to the previous ε_n's. The stationarity and ergodicity of TARCH model is discussed by many authors like Chan (1990) and Nelson and Cao (1992). Li and Lam (1995) used TARCH model to study the asymmetry in stock returns. They studied the asymmetric behavior of stock prices using threshold type non-linear time series model with conditional variance. Here we explain the method of estimation with the help of an empirical data. The rest of the Section explains details of the TARCH modelling technique for the monthly coconut oil price data.

Empirical Results

The empirical example here is the same as that we considered in the previous sections. Here we combined the TAR and ARCH models together. The method used here is as follows. First we analyse the data using the method given in the Section 2.2. That first we try to fit a TAR model to the data. The TAR model to monthly coconut oil prices is (cf Section 2.2.5):

$$\nabla X_n = -45.11 + 0.1088\nabla X_{n-1} - 0.00418\,\nabla X_{n-2} + \varepsilon^{(1)}_n \quad \text{if } \nabla X_{n-1} \leq -100$$
$$= 10.88 + 0.0297\nabla X_{n-1} + 0.12370\nabla X_{n-2} + \varepsilon^{(2)}_n \quad \text{if } -100 Z \nabla X_{n-1} \leq 40$$
$$= 67.69 + 0.3419\nabla X_{n-1} - 0.3250\nabla X_{n-2} + \varepsilon^{(3)}_n \quad \text{if otherwise.}$$

That is, a TAR model with three regimes. The model is piece-wise linear. That is an autoregressive model of order two is fitted in each regime. Now the next step is to test the ARCH effect in each regime separately using the method explained in Section 2.2. Estimate the squares of the residuals and the Q statistic (2.3.3) in each regime.

Regimes	Q - statistic
1	-18.46 (21.3,n=12)
2	8.36 (28.87,n=18)
3	114.60 (36.42,n=24)

Figures in parenthesis are the table values

By comparing the calculated values with the table values we can see that the ARCH effect is present only in the third regime. Thus the procedure continues only for the third one. By regressing the residual sum of squares on the lagged values we get the conditional variance. Here also we calculated the TR^2 statistic for different regression equations (see the previous section). By comparing the values of TR^2 values and the regression coefficients finally we arrived at the following equation for the conditional variance in the third regime. That is the TARCH model is

$$\nabla X_n = -45.11 + 0.1088\nabla X_{n-1} - 0.00418\,\nabla X_{n-2} + \varepsilon^{(1)}_n \quad \text{if } \nabla X_{n-1} \leq -100$$

$$= 10.88 + 0.0297\nabla X_{n-1} + 0.12370\nabla X_{n-2} + \varepsilon^{(2)}_n \quad \text{if } -100Z\nabla X_{n-1} \leq 40$$

$$= 67.69 + 0.3419\nabla X_{n-1} - 0.3250\nabla X_{n-2} + \varepsilon^{(3)}_n \quad \text{if otherwise,}$$

where $\varepsilon^{(3)}_n$ follows a normal distribution with zero mean and conditional variance

$$h_n = 21268 + 0.15091\hat{\varepsilon}^2_{n-1} + 0.2086\hat{\varepsilon}^2_{n-2} + 0.13062\hat{\varepsilon}^2_{n-5} \quad (TR^2 = 12.14).$$

Finally we arrived at a conclusion that there is an ARCH effect in the TAR model for the monthly coconut oil prices.

A portion of this chapter is already published in Nampoothiri and Balakrishna (2000).

CHAPTER – 3

CAUCHY AUTOREGRESSIVE MODELS

3.1 INTRODUCTION

Recently there has been a growing interest in studying non-Gaussian time series models. The need for such models arises from the fact that many of the naturally occurring time series are non-Gaussian. In the classical Box-Jenkins method, it is assumed that the observed time series is realization from a Gaussian process. A variety of data, especially in the field of economics, tended to jump around too much and involve outliers that contained important information. That is, such data have tendency to follow distributions with heavy tails. As a consequence a number of non-Gaussian time series models are developed to study time series data, which do not fit into the standard Gaussian linear models. One such class is the class of exponential time series models characterized by a set of observations distributed as exponential. This class was first introduced by Gaver and Lewis (1980) and later extended by Lawrence and Lewis (1981,1985). Other important studies in the non-Gaussian time series models are by Sim (1990) and Adke and Balakrishna (1992). They studied autoregressive models (AR) with gamma marginal. Abraham and Balakrishna (1999) introduced an inverse-Gaussian AR model. In their study, they introduced a first order autoregressive process with one-dimensional inverse Gaussian marginal. They obtained innovation distribution and estimated the unknown parameters by different methods.

One of the important non-normal distributions suitable for studying the behavior of price data is the Cauchy distribution with probability density function

$$f(x,\mu,\delta)=\frac{\delta}{\delta^2+(x-\mu)^2}, -\infty<x<\infty, 0<\delta<\infty, -\infty<\mu<\infty \qquad (3.1.1)$$

If a random variable X has the probability density function (3.1.1), then we say that X has $C(\delta,\mu)$ distribution, where δ and μ are the scale and the location parameters respectively. The distribution is symmetric about x= μ and hence the median is μ. The distribution does not possess finite moments of order greater than or equal to one, however μ and δ may be regarded as being analogous to mean and standard deviation. (Johnson and Kotz, 1994 pp. 299). A major difference between Normal and Cauchy distributions is that the latter has a longer tail than the former. This facilitates a better modeling of the price data using Cauchy distribution. The observations on a characteristic collected at different time points need not be independent. For example, the price of a commodity on a particular day depend on the previous day price. In such cases a Markov dependent sequence will be a better model to describe the data. So in this chapter we study the properties of a first order autoregressive model, which generates a sequence of Cauchy random variables.

A summary of this chapter is as follows. In Section 3.2 we define the model and study its properties. The Section 3.3 discusses the maximum likelihood estimation of the model parameters of a first order Cauchy autoregressive model while the Section 3.4 suggests some alternative method of estimation.

3.2 THE MODEL AND ITS PROPERTIES

Let $\{\varepsilon_n, n\geq 1\}$ be a sequence of independently and identically distributed (i.i.d) random variables and X_0 be a random variable with distribution function (d.f) F independent of ε_1. For $n\geq 1$, define

$$X_n = \rho X_{n-1} + \varepsilon_n \qquad n=1,2,\ldots \quad 0<|\rho|<1 \quad (3.2.1)$$

Now $\{X_n , n\geq 1\}$ defines a strictly stationary sequence of random variables . Sometimes $\{\varepsilon_n\}$ is referred to as an innovation sequence. Let $\phi_X(t)$ be the characteristic function (c.f.) of X_0 and $\phi_\varepsilon(t)$ be that of ε_t . The c.f. of a $C(\delta,\mu)$ is given by

$$\phi_X(t) = \exp(i\mu t - \delta|t| \quad , -\infty < t < \infty \qquad (3.2.2)$$

If $X_n \sim C(\delta,\mu)$ for all n then from (3.2.1) we can write

$$\phi_X(t) = \phi_X(\rho t) \,.\, \phi_\varepsilon(t)$$

and hence $$\phi_\varepsilon(t) = \frac{\phi_X(t)}{\phi_X(\rho t)}.$$

Now substituting for $\phi_\varepsilon(t)$ we get

$$\phi_\varepsilon(t) = e^{it\mu(1-\rho) - \delta|t|(1-|\rho|)}$$

That is ,ε_n has a Cauchy distribution with parameters $\mu(1-\rho)$ and $\delta(1-|\rho|)$. There fore, the probability density function of ε_n is given by

$$g(x) = \frac{\delta(1-|\rho|)}{\pi} \cdot \frac{1}{\delta^2(1-|\rho|)^2+[X-\mu(1-\rho)]^2} \quad -\infty < x < \infty$$

$$0< \delta < \infty, \ -\infty < \mu < \infty, 0<|\rho|<1 \qquad (3.2.3)$$

Clearly $\{X_n\}$ defined in (3.2.1) is a stationary Markov sequence and the transition probability distribution of X_n at x given X_{n-1}=y becomes ,

$$P[X_n \le x \mid X_{n-1} = y] = P[\varepsilon_n \le x - \rho y]$$

$$= \int_{-\infty}^{x-\rho y} g(u)du \ .$$

The corresponding transition density is

$$h(x \mid y; \delta, \mu, \rho) = \frac{\delta(1-|\rho|)}{\pi} \cdot \frac{1}{\delta^2(1-|\rho|)^2 + [(x-\rho y)-\mu(1-\rho)]^2}$$

$$-\infty < x < \infty \qquad (3.2.4)$$

we will use this density function for estimating the parameters in the next section.

Remark 3.2.1 : In AR(1) model (3.2.1) , the parameters ρ is usually interpreted as correlation coefficient. But here we have a Cauchy marginal distribution for X_n and the correlation coefficient does not exist. Hence we refer ρ as a parameter of the model and not as the correlation coefficient.

The transition distribution of X_n at x given X_0=y is

$$P[X_n \le x \mid X_0 = y] = P[\rho^n X_0 + \rho^{n-1}\varepsilon_1 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n \le x \mid X_0 = y], \qquad (3.2.5)$$

since repeatedly using (3.2.1) we can write

$$
\begin{aligned}
X_n &= \rho X_{n-1} + \varepsilon_n \\
&= \rho(\rho X_{n-2} + \varepsilon_{n-1}) + \varepsilon_n \\
&= \rho^2 X_{n-2} + \rho\varepsilon_{n-1}) + \varepsilon_n \\
&= \rho^3 X_{n-3} + \rho^2 \varepsilon_{n-2} + \rho\, \varepsilon_{n-1} + \rho\, \varepsilon_n \\
&\dots \\
&= \rho^n X_0 + \rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \dots + \rho\, \varepsilon_{n-1} + \varepsilon_n.
\end{aligned}
$$

There fore equation (3.2.5) becomes

$$
\begin{aligned}
\mathrm{P}[X_n \le x \mid X_0 = y] &= \mathrm{P}[\rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \dots + \rho\varepsilon_{n-1} + \varepsilon_n \le x - \rho^n y] \\
&= \mathrm{P}[Z_n \le x - \rho^n y]
\end{aligned}
$$

where

$$Z_n = \rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \dots + \rho\varepsilon_{n-1} + \varepsilon_n$$

and $\{\varepsilon_n\}$ is a sequence of i.i.d. $C(\delta(1-|\rho|), \mu(1-\rho))$.

Therefore, the characteristic function of Z_n is

$$
\begin{aligned}
\phi_{Z_n}(\mathrm{t}) &= \mathrm{E}(e^{itZ_n}) \\
&= \mathrm{E}\left(e^{it(\rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \dots + \rho\varepsilon_{n-1} + \varepsilon_n)}\right) \\
&= \prod_{j=1}^{n} \phi_{\varepsilon}\left(\rho^{n-j} t\right) \\
&= \prod_{j=1}^{n} \left[e^{i\mu(1-\rho)t\rho^{n-j} - \delta(1-|\rho|)|t\rho^{n-j}|} \right] \\
&= e^{i\mu(1-\rho)t\sum_{j=1}^{n}\rho^{n-j} - \delta(1-|\rho|)|t|\sum_{j=1}^{n}|\rho|^{n-j}}
\end{aligned}
$$

$$= e^{i\mu(1-\rho)t\left(\frac{1-\rho^n}{1-\rho}\right)-\delta(1-|\rho|)|t|\left(\frac{1-|\rho|^n}{1-|\rho|}\right)}$$

$$= e^{i\mu t(1-\rho^n)-\delta|t|(1-|\rho|^n)} \,. \tag{3.2.6}$$

Thus we have the following result.

Result 3.2.1 : If $\{X_n\}$ is a sequence of Cauchy AR(1) sequence defined in (3.2.1), then $Z_n = \rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n$ follows a Cauchy distribution with parameters $\delta(1-|\rho|^n)$ and $\mu(1-\rho^n)$.

It is well known that if $\{X_n\}$ is a sequence of i.i.d $C(\delta,\mu)$ then $\frac{S_n}{n}$ also has $C(\delta,\mu)$ distribution for every n where $S_n=X_1+X_2+ \ldots + X_n$. In the case of Cauchy AR (1) sequence we have the following theorem.

Theorem 3.2.1: Let $\{X_n\}$ be a Cauchy AR (1) sequence defined by (3.2.1). Then $\frac{S_n}{n}$ converges in law to $C\left(\frac{1-|\rho|}{1-\rho}\delta,\mu\right)$ distribution as $n \to \infty$. In other words, the Cauchy AR (1) sequence belongs to the domain of attraction of a Cauchy distribution.

Proof:

We have $X_n = \rho X_{n-1} + \varepsilon_n$

and

$$S_n = X_1 + X_2 + \ldots + X_n.$$

That is

$$S_n = X_1 + (\rho X_1 + \varepsilon_2) + (\rho^2 X_1 + \rho\varepsilon_2 + \varepsilon_3) + \ldots + (\rho^{n-1} X_1 + \rho^{n-2}\varepsilon_2 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n)$$

$$= (1 + \rho + \rho^2 + \ldots + \rho^{n-1})X_1 + (1 + \rho + \rho^2 + \ldots + \rho^{n-2})\varepsilon_2 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n$$

$$= \frac{1-\rho^n}{1-\rho} X_1 + \frac{1-\rho^{n-1}}{1-\rho}\varepsilon_2 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n$$

$$= \frac{1-\rho^n}{1-\rho} X_1 + \sum_{j=2}^{n} \frac{1-\rho^{n-j+1}}{1-\rho}\varepsilon_j .$$

The characteristic function of S_n becomes

$$\phi_{S_n}(t) = e^{i\mu t\left(\frac{1-\rho^n}{1-\rho}\right) - \delta\left|\frac{t(1-\rho^n)}{1-\rho}\right|} \cdot \prod_{j=2}^{n} \phi_\varepsilon\left(t.\frac{1-\rho^{n-j+1}}{1-\rho} \right)$$

$$= e^{i\mu t\left(\frac{1-\rho^n}{1-\rho}\right) - \delta\left|\frac{t(1-\rho^n)}{1-\rho}\right|} .\exp\{\sum_{j=2}^{n} i\mu t\left(\frac{(1-\rho^{n-j+1})(1-\rho)}{1-\rho}\right) - \left(\delta(1-|\rho|)\left|\frac{t(1-\rho^{n-j+1})}{1-\rho}\right|\right)\}$$

$$= e^{i\mu t\left(\frac{1-\rho^n}{1-\rho}\right) - \delta\left|\frac{t(1-\rho^n)}{1-\rho}\right|} .\exp\{\sum_{j=2}^{n} i\mu t(1-\rho^{n-j+1}) - \left(\delta(1-|\rho|)\left|\frac{t(1-\rho^{n-j+1})}{1-\rho}\right| \right)\} .$$

There fore, the c.f. of $\frac{S_n}{n}$ is

$$\phi_{\frac{S_n}{n}}(t) =$$

$$e^{i\mu(t/n)\left(\frac{1-\rho^n}{1-\rho}\right)-\delta\left|\frac{(t/n)(1-\rho^n)}{1-\rho}\right|}.\exp\left\{\sum_{j=2}^{n} i(t/n)\mu(1-\rho^{n-j+1})-\left(\delta(1-|\rho|)\left|\frac{(t/n)(1-\rho^{n-j+1})}{1-\rho}\right|\right)\right\}.$$

As $n \to \infty$ we get

$$\phi_{\frac{S_n}{n}}(t) \to e^{i\mu t-\delta|t|\frac{(1-|\rho|)}{1-\rho}}.$$

If $0 \le \rho \le 1$ then

$$\phi_{\frac{S_n}{n}}(t) \to e^{i\mu t-\delta|t|} \quad \text{as} \quad n \to \infty.$$

This completes the proof.

Joint Distribution of X_0 and X_n

Lemma 3.2.1:

If $\{X_n, n \ge 0\}$ is a Cauchy AR (1) sequence defined as (3.2.1), then the joint distribution of X_0 and X_n is a bivariate Cauchy with characteristic function

$$\phi_{X_0,X_n}(t_1,t_2) = e^{i\mu(t_1+t_2)}\ e^{-\{|t_1+t_2\rho^n|-|t_2|(1-|\rho|^n)\}\delta}$$

Proof

$$X_n = \rho^n X_0 + \rho^{n-1}\varepsilon_1 + \rho^{n-2}\varepsilon_2 + \ldots + \rho\ \varepsilon_{n-1} + \varepsilon_n$$

$$\phi_{X_0,X_n}(t_1,t_2) = E\left(e^{it_1X_0+it_2X_n}\right)$$

$$=\mathrm{E}\left(e^{it_1X_0+it_2(\rho^nX_0+\rho^{n-1}\varepsilon_1+\ldots+\rho\varepsilon_{n-1}+\varepsilon_n)}\right)$$

$$=\mathrm{E}\left(e^{i(t_1+t_2\rho^n)X_0}\right).\,\mathrm{E}\left(e^{it_2(\rho^{n-1}\varepsilon_1+\ldots+\rho\varepsilon_{n-1}+\varepsilon_n)}\right)$$

$$=\phi_{X_0}(t_1+t_2\rho^n).\prod_{j=1}^{n}\phi_\varepsilon(t_2\rho^{n-j})$$

$$=e^{i\mu(t_1+t_2\rho^n)-\delta|t_1+t_2\rho^n|}\,.\,e^{i\mu t_2(1-\rho^n)-\delta|t_2|(1-|\rho|^n)}$$

$$=e^{i\mu(t_1+t_2\rho^n)-\delta|t_1+t_2\rho^n|+i\mu t_2(1-\rho^n)}.e^{i\mu t_2}.e^{-i\mu t_2\rho^n}.e^{-\delta|t_2|(1-|\rho|^n)}$$

$$=e^{i\mu(t_1+t_2)}.e^{-\{|t_1+t_2\rho^n|+|t_2|(1-|\rho|^n)\}\delta}\,.$$

When $\mu=0$ this becomes

$$\phi_{X_0X_n}(t_1,t_2)=e^{-\{|t_1+t_2\rho^n|+|t_2|(1-|\rho|^n)\}\delta},$$

which is the required result. In the next theorem, we establish some properties of $\{X_n\}$, which are useful in studying the properties of the estimators.

In the rest of this chapter we assume that $\mu=0$.

Lemma 3.2.2 : Let $\{X_n, n\geq 0\}$ be an AR(1) process defined by (3.2.1) with marginal p.d.f. (3.1.1). Assume that

i. $E[\{\log|\varepsilon_1|\}^+]<\infty$ and

ii. ε_1 has a non-trivial absolutely continuous component.

Then for any initial distribution of X_0, the Markov sequence $\{X_n\}$ is Haris recurrent and strong mixing.

A proof of this Lemma is given in Arthreya and Pantula (1986a) .

Next we show that the conditions of Lemma 3.2.2 hold for the Cauchy AR(1) sequence.

Verification of the conditions of Lemma 3.2.2.

Take $\mu=0$ and $\delta=1$, then the $\{\varepsilon_n\}$ follow a $C(1-|\rho|, 0)$. Then the probability distribution function is

$$f_\varepsilon(y) = \frac{(1-|\rho|}{\pi} \cdot \frac{1}{(1-|\rho|)^2 + y^2} \qquad -\infty < y < \infty$$

$$[[\log|\varepsilon_1|]^+ = 0 \qquad \text{if } |\varepsilon_1| \le 1$$

$$= |\varepsilon_1| \quad \text{if } |\varepsilon_1| > 1$$

$$\text{and } E[[\log|\varepsilon_1|]^+ = \int \log|\varepsilon_1| f_\varepsilon(y)dy$$

The density function becomes

$$F_\varepsilon(y) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x}{1-|\rho|}\right)$$

$$P(|\varepsilon_1| \le x) = P[-x \le \varepsilon_1 \le x]$$

$$= \frac{1}{\pi}\tan^{-1}\left(\frac{x}{(1-|\rho|)}\right) - \frac{1}{\pi}\tan^{-1}\left(\frac{-x}{1-|\rho|}\right)$$

$$= \frac{2}{\pi}\tan^{-1}\left(\frac{x}{1-|\rho|}\right)$$

$$= 2F(x) - 1$$

$$= \frac{2}{\pi}\tan^{-1}\left(\frac{x}{1-|\rho|}\right)$$

Therefore, probability density function is

$$f_{\varepsilon_1}(x) = \frac{2}{\pi}\left[\frac{1}{1+\frac{x^2}{(1-|\rho|)^2}} \cdot \frac{1}{1-|\rho|}\right]$$

$$= \frac{2}{\pi}\left[\frac{(1-|\rho|)}{(1-|\rho|)^2 + x^2}\right] \qquad 0 \le x \le \infty$$

put $y = |\varepsilon_1|$. There fore

$$\text{E}[\log|\varepsilon_1|]^+ = \int_1^\infty \log y . \frac{2(1-|\rho|}{\pi} . \frac{1}{(1-|\rho|)^2 + y^2} dy$$

(Since $\log|\varepsilon_1| < 0$ for $0 \le |\varepsilon_1| \le 1$) .

$$= \frac{2(1-|\rho|)}{\pi}\int_1^\infty \frac{\log y}{(1-|\rho|)^2 + y^2} dy$$

$$= \frac{2(1-|\rho|)}{\pi}\int_0^\infty \frac{te^t}{(1-|\rho|)^2 + e^{2t}} dt$$

$$< \frac{2(1-|\rho|)}{\pi} \int_0^\infty \frac{te^t}{e^{2t}} dt$$

$$= \frac{2(1-|\rho|)}{\pi} \int_0^\infty te^{-t} dt$$

$$= \frac{2(1-|\rho|)}{\pi} < \infty .$$

Further ε_t is absolutely continuous and (ii) automatically holds. There fore, the conditions of Lemma 3.2.2 are satisfied.

The following Lemma proved by Arthreya and Pantula (1986a) helps us to obtain the mixing coefficients of a strongly mixing sequence.

Lemma 3.2.3: For a Harris recurrent Markov sequence $\{X_n]$,

$$\sup_{A \in F^n{}_0, B \in F^\infty{}_{m+n}} |P(A \cap B) - P(A)P(B)| = \alpha'(m) \leq 2 \sup_n E[K_{m-1}(X_{n+1})]$$

where $F^n{}_0$ and $F^\infty{}_{m+n}$ are the minimal sigma fields induced by $(X_0, X_1, ..., X_n)$ and $(X_{n+m}, X_{n+m+1}, ...)$ respectively, and

$$K_{m-1}(X_{n+1}) = \|P(X_{n+m+1} \in A \mid X_{n+1}) - \pi(A)\|,$$

$P(X_{n+m+1} \in A \mid X_{n+1})$ denotes the m-step transition function of $\{X_n\}$, $\|\mu - \nu\|$ is the total variation norm of the signed measure $\mu - \nu$ for the probability measures μ and ν and $\pi(.)$ is the stationary measure.

Theorem 3.2.2: The stationary Cauchy AR (1) process $\{X_n\}$ defined by (3.2.1) is strong mixing with mixing coefficients :

$$\alpha(m) = 2|\rho|^{m-1} \text{ , m=1,2,3,...}$$

Proof

By Lemma 3.2.3, the strong mixing coefficients

$$\alpha(m) = 2\sup_n E(K_{m-1}(X_{n+1}))$$

where

$$K_{m-1}(X_{n+1}) = \|P(X_{n+m+1} \in A \mid X_{n+1} - \pi(A)\|$$

where $P((X_{m+n+1} \in A \mid X_{n+1}))$, $\|.\|$ are as defined above and $\pi(.)$ is the stationary measure given by

$\pi(A) = \int_A f(x)dx$, f(x) is the probability density function of X_n given by (3.1.1).

Now $$E[K_{m-1}(X_{n+1})] = \int_{\infty}^{\infty} K_{m-1}(x)f(x)dx$$

$$= \int_{-\infty}^{\infty} \|P_x(X_{m-1} \in A) - \pi(A)\| f(x)dx,$$

where $P_x(X_{m-1} \in A) = \Pr[X_{m-1} \in A \mid X_0 = x]$ and A is an arbitrary event.

Now $$P_x(X_{m-1} \in A) = P(X_{m-1} \in A / X_0 = x)$$

$$= \quad P[\rho^{m-1}X_0 + \rho^{m-2}\varepsilon_1 + \ldots + \varepsilon_{m-1} \in A \mid X_0 = x)$$

$$= P[\rho^{m-2}\varepsilon_1 + \ldots + \varepsilon_{m-1} \in A - \rho^{m-1}x]$$

$$= P[Z_{m-1} \in A - \rho^{m-1}x].$$

From the Result (3.2.6) the p.d.f of Z_{m-1} with $\mu=0$ can be written as

$$h(z) = \frac{`\delta(1-|\rho|^{m-1})}{\pi} \cdot \frac{1}{\delta^2(1-|\rho|^{m-1})^2 + z^2}.$$

Therefore,

$$P_x(X_{m-1} \in A) = \int_{A-\rho^{m-1}x} h(z)dz \le \int_A h(z)dz.$$

Hence for any arbitrary A,

$$\|P_x(X_{m-1} \in A) - \pi(A)\| \le \left\| \int_A [h(z) - f(z)]dz \right\|.$$

Note that

$$\|P_x(X_{m-1} \in A) - \pi(A)\| = \|\pi(A) - P_x(X_x \in A)\|$$

Therefore consider

$$\int_A [f(z) - h(z)]dz = \int_A \left[\frac{\delta}{\pi(\delta^2 + z^2)} - \frac{\delta(1-|\rho|^{m-1})}{\pi} \cdot \frac{1}{\delta^2(1-|\rho|^{m-1})^2 + z^2} \right] dz$$

$$= \frac{\delta}{\pi} \int_A \left[\frac{1}{\delta^2 + z^2} - \frac{(1-|\rho|^{m-1})}{\delta^2(1-|\rho|^{m-1})^2 + z^2} \right] dz$$

Now we simplify

$$\|P_x(X_{m-1} \in A) - \pi(A)\| = \left\| \int_{A-\rho^{m-1}x} h(z)dz - \int_A f(z)dz \right\|.$$

If $\rho^{m-1}x > 0$ then

$$\int_{A-\rho^{m-1}x} h(z)dz \le \int_A h(z)dz$$

Hence

$$\|P_x(X_{m-1} \in A) - \pi(A)\| \le \left\| \int_A [h(z) - f(z)]dz \right\|$$

$$= \left\| \int_A \left[\frac{\delta(1-|\rho|^{m-1})}{\pi} . \frac{1}{\delta^2(1-|\rho|^{m-1})^2 + z^2} - \frac{\delta}{\pi(\delta^2 + z^2)} \right] dz \right\|$$

$$= \left\| \frac{\delta}{\pi} \int_A \left[\frac{1}{\delta^2 + z^2} - \frac{(1-|\rho|^{m-1})}{\delta^2(1-|\rho|^{m-1})^2 + z^2} \right] dz \right\|$$

$$= \left\| \frac{\delta}{\pi} \int_A \left[\frac{1}{\delta^2 + z^2} - \frac{1-|\rho|^{m-1}}{\delta^2 + z^2} \right] dz \right\|,$$

$\left(\text{since } z^2 + \delta^2(1-|\rho|^{m-1})^2 \le z^2 + \delta^2\right)$

$$= \left\| \int_A \frac{\delta}{\pi} . \frac{1}{\delta^2 + z^2} . |\rho|^{m-1} dz \right\|$$

$$= |\rho|^{m-1} \int_A f(z)dz$$

$$\le |\rho|^{m-1}$$

If $\rho^{m-1}x < 0$ then we can use the inequality

$$\int_A f(z)dz \le \int_{A-\rho^{m-1}x} f(z)dz$$

Then

$$\left\|P_x(X_{m-1}\in A)-\pi(A)\right\|=\left\|\int_{A-\rho^{m-1}x}h(z)dz-\int_A f(z)dz\right\|$$

$$=\left\|\int_A f(z)dz-\int_{A-\rho^{m-1}x}h(z)dz\right\|$$

$$\le\left\|\int_{A-\rho^{m-1}x}f(z)dz-\int_{A-\rho^{m-1}x}h(z)dz\right\|$$

$$=\left\|\int_{A-\rho^{m-1}x}\left[\frac{\delta}{\pi}\frac{1}{(\delta^2+z^2)}-\frac{\delta(1-|\rho|^{m-1})}{\pi(\delta^2(1-|\rho|^{m-1})^2+z^2)}\right]dz\right\|$$

$$\le\left\|\int_{A-\rho^{m-1}x}\left[\frac{\delta}{\pi}\left(\frac{1}{(\delta^2+z^2)}-\frac{(1-|\rho|^{m-1})}{\delta^2+z^2}\right)\right]dz\right\|$$

$$=\left\|\frac{\delta}{\pi}\int_{A-\rho^{m-1}x}\frac{1}{(\delta^2+z^2)}.|\rho|^{m-1}dz\right\|$$

$$=\left\|\int_{A-\rho^{m-1}x}\frac{\delta}{\pi}.\frac{1}{\delta^2+z^2}dz\right\|.|\rho|^{m-1}$$

$$=|\rho|^{m-1}\int_{A-\rho^{m-1}x}f(z)dz$$

$$\le|\rho|^{m-1}.$$

Therefore, in any case (if μ=0)

$$\left\|P_x(X_{m-1}\in A)-\pi(A)\right\|\le|\rho|^{m-1}.$$

Therefore,

$$\alpha(m)=2\sup E[K_{m-1}X_{n+1})]$$

$$=2\sup\int_{-\infty}^{\infty}\left\|P_x\in A-\pi(A)\right\|f(x)dx$$

$$\le2\sup\int_{-\infty}^{\infty}|\rho|^{m-1}f(z)dz$$

$$=2|\rho|^{m-1}$$

Therefore $\alpha(m) = 2|\rho|^{m-1}$ can be taken as a sequence of mixing parameters.

Hence the theorem.

Theorem 3.2.3 : Let $\{X_n\}$ be a sequence of Cauchy AR(1) sequence defined as before, then $\{X_n\}$ is ergodic.

Proof: Note that X_n can be written as a sum of independent random variables as

$$X_n = \rho^n X_0 + \rho^{n-1}\varepsilon_1 + \ldots + \rho\varepsilon_{n-1} + \varepsilon_n$$

Let $F_n = \sigma(X_1, \ldots X_n)$ and $G_n = \sigma(X_0, \varepsilon_1, \ldots, \varepsilon_n)$ n=1,2,... be the σ-fields generated by $(X_1, \ldots X_n)$ and $(X_0, \varepsilon_1, \ldots, \varepsilon_n)$ respectively.

Then $F_n \subseteq G_n$, that the sigma field of X_n is contained in the sigma field of independently and identically random variables $\{\varepsilon_n\}$. By Kolmogrov 0-1 law each event of the sigma field of independently and identically distributed random variables has probability zero or one. Hence each tail event of sigma field of $\{X_n\}$ has probability zero or one. This is a sufficient condition for $\{X_n\}$ to be ergodic. (Nicholls and Quinn (1981), pp37).

Note 3.2.1 : By Lemma 3.2.1 the joint characteristic function of X_0 and X_n is

$$\varphi_{X_0, X_n}(t_1, t_2) = e^{i\mu(t_1+t_2)} e^{-\{|t_1+t_2\rho^n| - |t_2|(1-|\rho|^n)\}\delta}.$$

Similarly

$$\phi_{X_n, X_0}(t_1, t_2) = e^{i\mu(t_1+t_2)}.e^{-\{|t_2+t_1\rho^n|-|t_1|(1-|\rho|^n)\}\delta}$$

That is,

$$\phi_{x_0, X_n}(t_1, t_2) \neq \phi_{X_n, X_0}(t_1, t_2).$$

There fore, Cauchy AR(1) sequence is not time irreversible, unlike a Gaussian AR(1) sequence.

3.3 MAXIMUM LIKELIHOOD ESTIMATION

In this section we obtain the maximum likelihood estimator of the parameters δ and ρ . Let $X = \{X_0, X_1, ... X_n\}$ be a sample from a Cauchy AR (1) sequence. The likelihood function of θ = (δ,ρ) based on X is given by

$$\begin{aligned} L &= \{(\delta, \rho) \mid X_0, X_1, ..., X_n\} \\ &= f(X_0).\prod_{j=1}^{n} h(X_j \mid X_{j-1}), \end{aligned}$$

where $h(X_j \mid X_{j-1})$ is the one-step transition density of $\{X_n\}$ given by (3.2.4). The parameter space here is given by

$$\Theta = \{(\delta, \rho) \mid \delta > 0, 0 < |\rho| < 1\}.$$

Now consider

$$\log L = \log(f(X_0)) + \sum_{j=1}^{n} \log h(X_j \mid X_{j-1})$$

$$=\log\frac{\delta}{\pi(\delta^2+X_0^{\ 2})}+\sum\log\left[\frac{\delta(1-|\rho|}{\pi[\delta^2(1-|\rho|)^2+(X_j-\rho X_{j-1}\ )^2]}\right].\quad (3.3.1)$$

Differentiating with respect to δ and ρ and ignoring the terms

corresponding to X_0 we get,

$$\frac{\partial\log L}{\partial\delta}=\frac{n}{\delta}-\sum_{j=1}^{n}\frac{2\delta(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+(X_j-\rho X_{j-1})^2]}$$

and

$$\frac{\partial\log L}{\partial\rho}=\frac{n(-\alpha)}{(1-|\rho|)}+2\sum\left[\frac{\delta^2(1-|\rho|)\alpha+(X_j\rho X_{j-1})X_{j-1}}{\delta^2(1-|\rho|)^2+(X_j-\rho X_{j-1})^2}\right]$$

where $\alpha=\pm1=\frac{\partial|\rho|}{\partial\rho}$, the sign of ρ. Here $\alpha=+1$ if $\rho>0$ and $\alpha=-1$ if $\rho<0,\quad \rho\neq0$.

These likelihood equations do not have closed form expressions for their solutions and hence an iterative procedure of Newton – Raphson type is used to estimate the parameters. The procedure is as follows. The estimator after k iterations is given by

$$\begin{bmatrix}\delta_{k+1}\\ \rho_{k+1}\end{bmatrix}=\begin{bmatrix}\delta_k\\ \rho_k\end{bmatrix}+\frac{1}{n}\begin{bmatrix}\frac{\partial^2\log L}{\partial\delta^2} & \frac{\partial^2\log L}{\partial\delta\partial\rho}\\ \frac{\partial^2\log L}{\partial\delta\partial\rho} & \frac{\partial^2\log L}{\partial\rho^2}\end{bmatrix}\begin{bmatrix}\frac{\partial\log L}{\partial\delta}\\ \frac{\partial\log L}{\partial\rho}\end{bmatrix},\ \text{k=1,2, 3, . . .}\quad (3.3.2)$$

The second term on the right hand side is evaluated at (δ_k,ρ_k) and (δ_0,ρ_0) is an initial value. In this case

$$\frac{\partial^2 \log L}{\partial \delta^2} = \frac{-n}{\delta^2} - 2(1-|\rho|)^2 \sum_{j=1}^{n} \left[\frac{(X_j - \rho X_{j-1})^2 - [\delta(1-|\rho|)]^2}{[\delta^2(1-|\rho|)^2 + (X_j - \rho X)^2} \right]$$

$$\frac{\partial^2 \log L}{\partial \rho^2} = \frac{-n\alpha^2}{(1-|\rho|)^2} +$$

$$2\sum_{j=1}^{n} \left[\frac{\left[\delta^2(1-|\rho|)^2 + (X_j - \rho X_{j-1})^2\right]\left[-\left(\delta^2\alpha^2 + X_{j-1}{}^2\right)\right] + 2\left[\delta^2\alpha(1-|\rho|) + (X_j - \rho X_{j-1})X_{j-1}\right]^2}{\left[\delta^2(1-|\rho|)^2 + (X_j - \rho X_{j-1})^2\right]^2} \right]$$

and

$$\frac{\partial^2 \log L}{\partial \rho \partial \delta} =$$

$$-2\sum_{j=1}^{n} \left[\frac{\left[\delta^2(1-|\rho|)^2 + (X_j \rho X_{j-1})^2\right]\left[-2\delta(1-|\rho|)\alpha\right] + \delta(1-|\rho|)^2\left[2\left(\delta^2(1-|\rho|)\right)\alpha + (X_j \rho X_{j-1})X_{j-1}\right]}{\left[\delta^2(1-|\rho|)^2 + (X_j - \rho X_{j-1})^2\right]^2} \right]$$

According to Billingsley (1961), under certain regularity conditions (which are stated below) the maximum likelihood estimate (mle) of (δ, ρ) is consistent and asymptotically normal. But some of the regularity conditions of Billingsley, in particular the moment conditions do not hold if both δ and ρ are unknown. However, if we assume that ρ is known then the maximum likelihood estimator of δ has optimal properties. The mle of δ can be obtained by solving the equations $\frac{\partial \log L}{\partial \delta} = 0$ by Newton-Raphson method. In this case it is readily verified that all the regularity conditions of Billinsley (1961) hold and the maximum likelihood estimates $\hat{\delta}$ of δ is consistent and asymptotically normal

(CAN) estimate with mean δ and asymptotic variance $2\delta^2/n$. The performance of the mle based on simulation experiment is discussed in Chapter 4. We will state the regularity conditions given by Billingsley(1961) before their verification.

Let $\{X_n\}$ be a Markov process, $f(\varepsilon;\theta)$ be the density of the initial distribution and $f(\varepsilon,\eta;\theta)$ be the densities of all transition measures. Let $\theta=(\theta_1,\ldots,\theta_r)$ be the unknown parameter vector and Θ be the parameter space.

Condition 3.3.1 : For any ε, the set of η for which $f(\varepsilon,\eta;\theta)>0$ does not depend on θ. For any ε and η, $f_u(\varepsilon,\eta;\theta), f_v(\varepsilon,\eta;\theta), f_{uv}(\varepsilon,\eta;\theta)$ and $f_{uvw}(\varepsilon,\eta;\theta)$ exist and are continuous through out Θ (where $f_u=\dfrac{\partial f}{\partial\theta_u}$, $f_v=\dfrac{\partial f}{\partial\theta_v}$, $f_{uv}=\dfrac{\partial^2 f}{\partial\theta_u\partial\theta_v}$ and $f_{uvw}=\dfrac{\partial^3 f}{\partial\theta_u\partial\theta_v\partial\theta_w}$ are the partial derivatives). For any $\theta\in\Theta$ there exists a neighbourhood N of θ such that for any u,v,w,ε,

$$\int_X \sup_{\theta'\in N}\left|f_u(\varepsilon,\eta;\theta'\right|\lambda d\eta<\infty$$

$$\int_X \sup_{\theta'\in N}\left|f_{uv}(\varepsilon,\eta;\theta'\right|\lambda d\eta<\infty \qquad \text{and}$$

$$E_\theta\{\sup_{\theta'\in N}|\,g_{uvw}(x_1,x_2;\theta')\,|\}<\infty.$$

Finally, for u=1,2,…,r

$$E_\theta\{|g_u(x_1,x_2,;\theta'^{\,2}\}<\infty$$

and if $\sigma_{uv}(\theta)$ is defined by

$$\sigma_{uv}(\theta) = E\{g_u(x_1, x_2; \theta) g_v(x_1, x_2; \theta)\}$$

then the rxr matrix $\sigma(\theta) = (\sigma_{uv}(\theta))$ is nonsingular.

Condition 3.2.2 : (i) For each $\theta \in \Theta$, the stationary distribution, for (which by assumption exists and is unique) has the property that for each ε in the state space, is absolutely ontinuous with respect to $p_\theta(.)$. That is,

$$p_\theta(\varepsilon,.) << p_\theta(.) .$$

(ii) There is some $\delta > 0$ such that for u=1,2,...r,

$$E_\theta\{|g_u(x_1, x_2; \theta)|^{2+\delta}\} < \infty .$$

The δ in the above equation may depend on θ.

Lemma 3.3.1: If part (i) of the second condition holds, then, for any $\theta \in \Theta$, the process $\{X_n\}$ is metrically transitive if the initial distribution is the stationary one. No matter what the initial distribution is, if φ is measurable f_X X f_X and if $E\{|\varphi(x_1, x_2)|\} < \infty$, then

(i) $\lim_{n \to \infty} n^{-1} \sum_{k=1}^{n} \varphi(x_k, x_{k+1}) = E_\theta\{\varphi(x_1, x_2)\}$,with probability one.

(ii) The second condition implies that for every $\theta \in \Theta$ and for any initial distribution , the random vector $n^{-1/2} \sum_{k=1}^{n} g_u(x_k, x_{k+1}; \theta)$ converges in law to $N(0, \sigma(\theta))$.

Proof - See Billingsley (1961).

Verification of Regularity conditions when δ and ρ are unknown

Consider the conditional covariance matrix

$$D_{x_{j-1}} = E\begin{bmatrix} \left[\dfrac{\partial \log h(x_j / x_{j-1})}{\partial \delta}\right]^2 & \dfrac{\partial^2 \log h(x_j / x_{j-1})}{\partial \delta \partial \rho} \\ \dfrac{\partial^2 \log h(x_j / x_{j-1})}{\partial \delta \partial \rho} & \left[\dfrac{\partial \log h(x_j / x_{j-1})}{\partial \rho}\right]^2 \end{bmatrix}$$

$$= E\begin{bmatrix} g^2{}_\delta & g_{\delta\rho} \\ g_{\delta\rho} & g^2{}_\rho \end{bmatrix} \text{ say,}$$

where

$$g_\delta = \frac{\partial h(x_j / x_{j-1})}{\partial \delta}$$

$$= \frac{1}{\delta} - \frac{2\delta(1-|\rho|)^2}{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2}$$

$$g_\rho = \frac{-\alpha}{(1-|\rho|)} + 2\left\{\frac{\delta^2(1-|\rho|)\alpha + (x_j - \rho x_{j-1})x_{j-1}}{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2}\right\},$$

where $\alpha = \pm 1$.

and

$$g_{\rho\delta} = \frac{\partial^2 h(x_j \mid x_{j-1})}{\partial \rho \partial \delta}$$

=

$$\frac{4\alpha\delta(1-|\rho|)}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}-\frac{4\delta(1-|\rho|)^2\{\delta^2\alpha(1-|\rho|)+(x_j-\rho x_{j-1})x_{j-1}\}}{[\delta^2(1-|\rho|)^2+(x_j\rho x_{j-1})^2]^2}$$

$$E[g^2{}_\delta(X_j\mid X_{j-1})]=\int_\infty^\infty g^2{}_\delta(x_j\mid x_{j-1})h(x_j\mid x_{j-1})dx_j$$

=

$$\int_\infty^\infty\left[\frac{1}{\delta}-\frac{2\delta(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}\right]^2\left[\frac{\delta(1-|\rho|)}{\pi}\left\{\frac{1}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right\}\right]dx_j$$

$$=\int_\infty^\infty\left\{\frac{1}{\delta^2}+\frac{4\delta^2(1-|\rho|)^4}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}-\frac{4(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}\right\}\cdot$$

$$\left\{\frac{\delta(1-|\rho|)}{\pi}\frac{1}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right\}\mathrm{dx_j}$$

$$=\int_\infty^\infty\left[\frac{1}{\delta^2}+\frac{4\delta^2(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+u^2\delta^2(1-|\rho|)^2]^2}-\frac{4(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+u^2\delta^2(1-|\rho|)^2}\right]$$

$$\cdot\left\{\frac{\delta(1-|\rho|)}{\pi}x\frac{\delta(1-|\rho|)du}{[\delta^2(1-|\rho|)^2+\delta^2(1-|\rho|)^2u^2]}\right\}$$

$$=\int_\infty^\infty\left[\frac{1}{\delta^2}+\frac{4}{\delta^2(1+u^2)^2}-\frac{4}{\delta^2(1+u^2)}\right]\frac{1}{\pi(1+u^2)}du$$

$$=\frac{1}{\delta^2}\int_\infty^\infty\frac{1}{\pi(1+u^2)}du+\frac{4}{\delta^2}\int_\infty^\infty\frac{1}{\pi}\frac{1}{(1+u^2)^3}du-\frac{4}{\delta}\int_\infty^\infty\frac{1}{\pi}\frac{1}{(1+u^2)^2}du$$

$$= \frac{1}{\delta^2} + \frac{8}{\delta^2}\int_0^\infty \frac{1}{\pi}\frac{1}{(1+u^2)^3}\,du - \frac{8}{\delta^2}\int_0^\infty \frac{1}{\pi(1+u^2)^2}\,du$$

$$= \frac{1}{\delta^2} + \frac{4}{\pi\delta^2}\int_0^\infty \frac{dt}{\sqrt{t}(1+t)^3} - \frac{4}{\pi\delta^2}\int_0^\infty \frac{1}{\sqrt{t}(1+t)^2}.$$

Next we will use a beta integral defined by

$$\text{B(x,y)}= \int_0^1 t^{x-1}(1-t)^{y-1}dt = 2\int_0^1 t^{2x-1}(1-t^2)^{y-1}dt$$

$$\int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}}dt = 2\int_0^\infty \frac{t^{2x-1}}{(1+t^2)^{x+y}}dt$$

$$= \int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}}dt$$

Now the above equation becomes ,

$$= \frac{1}{\delta^2} + \frac{4}{\delta^2}B\left(\frac{1}{2},\frac{5}{2}\right) - \frac{4}{\delta^2}B\left(\frac{1}{2},\frac{3}{2}\right)$$

$$= \frac{1}{\delta^2}\left[1 + \frac{4}{\pi}\frac{\Gamma(1/2).\Gamma(5/2)}{\Gamma 3} - \frac{4}{8\pi}.\frac{\Gamma(1/2)\Gamma(3/2)}{\Gamma 2}\right]$$

after simplification we get

$$E(g^2{}_\delta(X_j \mid X_{j-1})) = \frac{1}{\delta^2}\left[1 + \frac{3}{2} - 2\right] = \frac{1}{2\delta^2}$$

Consider the conditional expectation of $\frac{\partial g}{\partial \rho}$ given X_{j-1}.

$$E\left(\frac{\partial g(X_j \mid X_{j-1})}{\partial \rho}\right) = \int_{-\infty}^{\infty}\left\{\frac{-\alpha}{(1-|\rho|} + \frac{2\delta^2(1-|\rho|)\alpha + 2(x_j - \rho x_{j-1})x_{j-1}}{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2}\right\}.$$

$$\frac{\delta(1-|\rho|)}{\pi}\frac{dx_j}{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2}$$

$$=\frac{-\alpha}{(1-|\rho|)}+\frac{2}{\pi}\int_{-\infty}^{\infty}\left\{\frac{\alpha\delta^2(1-|\rho|)+x_{j-1}(x_j-\rho x_{j-1})}{(\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2)^2}\right\}\delta(1-|\rho|)dx_j$$

$$=\frac{-\alpha}{(1-|\rho|)}+\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{\alpha\delta^2(1-|\rho|)+x_{j-1}u\delta(1-|\rho|)}{[\delta^2(1-|\rho|)^2+\delta^2(1-|\rho|)^2u^2]^2}\delta^2(1-|\rho|)^2du$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{\delta^3(1-|\rho|)^3[\alpha\delta+ux_{j-1}]du}{\delta^4(1-|\rho|)^4[1+u^2]^2}$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2}{\pi}\frac{1}{\delta(1-|\rho|)}\int_{-\infty}^{\infty}\frac{(\alpha\delta+ux_{j-1})}{(1+u^2)^2}du$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2}{\pi\delta(1-|\rho|)}\int_{0}^{\infty}\frac{\alpha\delta+ux_{j-1}}{(1+u^2)^2}du$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2}{\pi\delta(1-|\rho|)}\left[2\int_{0}^{\infty}\frac{\alpha\delta}{(1+u^2)}du+x_{j-1}\int_{0}^{\infty}\frac{udu}{(1+u^2)}\right]$$

$\int_{-\infty}^{\infty}\frac{udu}{(1+u^2)^2}=0$, since it is an odd function.

Therefore, the expectation is equal to

$$=\frac{-\alpha}{1-|\rho|}+\frac{2\alpha\delta}{\pi\delta(1-|\rho|)}\int_{0}^{\infty}\frac{du}{(1+u^2)^2}$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2\alpha}{\pi(1-|\rho|)}B(1/2,3/2)$$

$$=\frac{-\alpha}{1-|\rho|}+\frac{2\alpha}{\pi(1-|\rho|)}\frac{\sqrt{\pi}.\frac{1}{2}.\sqrt{\pi}}{\Gamma 2}$$

$$=0$$

Now consider

$$E\left[\frac{\partial g(X_j \mid X_{j-1}}{\partial \rho r}\right]^2 = E\left[\frac{-\alpha}{1-|\rho|}+\frac{2\delta^2(1-|\rho|)\alpha+2(x_j-\rho x_{j-1})x_{j-1}}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right]^2$$

$$= E\left[\frac{\alpha^2}{(1-|\rho|)^2}\right]+E\left[\frac{2\delta^2(1-|\rho|)\alpha+2(x_j-\rho x_{j-1})x_{j-1}}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right]^2 -$$

$$\frac{-2\alpha}{(1-|\rho|)}E\left[\frac{2\delta^2(1-|\rho|)\alpha+2(x_j-\rho x_{j-1})x_{j-1}}{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right] \qquad (3.3.4)$$

Now the second term in (3.3.4) is

$$E\frac{4\delta^4(1-|\rho|)^2\alpha^2+8\delta^2(1-|\rho|)\alpha(x_j-\rho x_{j-1})x_{j-1}+4(x_j-\rho x_{j-1})^2 x_{j-1}{}^2}{\{\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2\}^2}$$

$$= E\left[\frac{4\delta^4(1-|\rho|)^2\alpha^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]+E\left[\frac{8\delta^2(1-|\rho|)\alpha(x_j-\rho x_{j-1})x_{j-1}}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]+$$

$$E\left[\frac{4(x_j-\rho x_{j-1})^2 x_{j-1}{}^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

Now

$$E\left[\frac{4\delta^4(1-|\rho|)^2\alpha^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

$$=4\delta^4(1-|\rho|)^2\alpha^2\int_{-\infty}^{\infty}\frac{\delta(1-|\rho|)}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^3}\frac{1}{\pi}dx_j$$

$$=\frac{4\delta^4(1-|\rho|)^2\alpha^2}{\delta^6(1-|\rho|)^6}\int_{-\infty}^{\infty}\frac{\delta^2(1-|\rho|)^2\,du}{\pi(1+u^2)^3}$$

$$=\frac{4\alpha^2}{(1-|\rho|)^2}.\frac{1}{\pi}.2\int_0^{\infty}\frac{1}{(1+u^2)^3}du$$

$$=\frac{4}{(1-|\rho|)^2}.\frac{1}{\pi}B(1/2,5/2)$$

$$=\frac{4}{(1-|\rho|)^2}.\frac{\Gamma 1/2.\Gamma 1/2.3/4}{\pi.2}$$

$$=\frac{3}{2}.\frac{1}{(1-|\rho|)^2}$$

$$\text{E}\left[\frac{8\delta^2(1-|\rho|).(x_j-\rho x_{j-1})x_{j-1}}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

$$=4.\text{E}_{x_j}\left[\frac{2\delta^2(1-|\rho|)\alpha(x_j\rho x_{j-1})x_{j-1}}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

$$=4.\frac{\alpha}{(1-|\rho|)} \qquad \text{by (3.3.3)}$$

and

$$\text{E}\left[\frac{4(x_j-\rho x_{j-1})^2 x_{j-}{}^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

$$=4\int_{-\infty}^{\infty}\frac{(x_j-\rho x_{j-1})^2 x_{j-1}{}^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}.\frac{\delta(1-|\rho|)}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}dx_j$$

$$= \frac{4x_{j-1}^{\;2}}{\pi}\int_{-\infty}^{\infty}\frac{u^2\delta^2(1-|\rho|)^2\delta^2(1-|\rho|)^2}{\delta^6(1-|\rho|)^6(1+u^2)^3}du$$

$$= \frac{8x_{j-1}^{\;2}}{\pi\delta^2(1-|\rho|)^2}\int\frac{u^2du}{(1+u^2)^3}$$

$$= \frac{4x_{j-1}^{\;2}}{\delta^2(1-|\rho|)^2}\frac{1}{\pi}B(3/2,3/2)$$

$$= \frac{4x_{j-1}^{\;2}}{\delta^2(1-|\rho|)^2}\frac{1}{\pi}\frac{\Gamma 1/2.1/2.\Gamma 1/2}{\Gamma 3}$$

$$= \frac{1}{2\delta^2(1-|\rho|)^2}x_{j-1}^{\;2}$$

The third term in equation (3.3.4)

$$= \frac{-2\alpha}{(1-|\rho|)}E\left[\frac{2\delta^2(1-|\rho|)\alpha+2(x_j-\rho x_{j-1})x_{j-1}}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2}\right]$$

$$=$$

$$\frac{-2\alpha}{(1-|\rho|)}\left\{\int_{-\infty}^{\infty}\frac{2\delta^2(1-|\rho|)\alpha}{[\delta^2(1-|\rho|)^2+(x_j-\rho r x_{j-1})^2]}\frac{\delta(1-|\rho|)}{\pi[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}dx_j\right.$$

$$+$$

$$\left.\int_{-\infty}^{\infty}\frac{2x_{j-1}(x_j-\rho x_{j-1})}{[\delta^2(1-|\rho|)^2+(x_j\rho x_{j-1})^2]}\frac{\delta(1-|\rho|)}{\pi[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}dx_j\right.$$

$$= \frac{-2\alpha}{(1-|\rho|)}\left\{\int\frac{2\alpha du}{(1-|\rho|)^4(1+u^2)^2}+\int\frac{2x_{j-1}\delta^3(1-|\rho|)^3udu}{\pi\delta^4(1-|\rho|)^4(1+u^2)^2}\right\}$$

$$= \frac{-4\alpha^2}{\pi(1-|\rho|)^2}B(1/2,3/2)$$

After simplification we get

The Third Term $= \dfrac{-2}{(1-|\rho|)^2}$

Now the conditional expectation becomes

$$E\left[\frac{\partial g(X_j / X_{j-1})}{\partial \rho}\right]^2$$

$$= \frac{1}{(1-|\rho|)^2} + \frac{3}{2}\frac{1}{(1-|\rho|)^2} + \frac{4\alpha}{(1-|\rho|)} + \frac{x_{j-1^2}}{2\delta^2(1-|\rho|)^2} - \frac{2\alpha^2}{(1-|\rho|)^2}$$

$$= \frac{3}{2(1-|\rho|)^2} + \frac{1}{2\delta^2(1-|\rho|)^2}X_{j-1}{}^2 + \frac{4\alpha}{(1-|\rho|)} - \frac{\alpha^2}{(1-|\rho|)^2} \qquad (3.3.5)$$

Thus we have the elements of the conditional information matrix D_{Xj-1}. Under stationarity, X_j has $C(\delta, 0)$ distribution for every j. Hence the information matrix can be obtained by considering the unconditional expectation of $D_{X_{j-1}}$ with respect to X_{j-1}. But unconditional expectation does not exist (not finite). That is, the regularity conditions are not satisfied for Cauchy AR (1) process. Thus some of the regularity conditions of Billingsley (1961) are violated. Hence we cannot use the theory by Billingsley to study the asymptotic properties of the Maximum Likelihood Estimators for stationary Markov sequence. But MLE of δ and ρ can be obtained by solving the likelihood equations.

It is also readily verified that some of the regularity conditions are violated when δ is known and ρ is unknown.

Verification of the regularity conditions with respect to δ when ρ is known.

(1) R = $\{x_j : f(x_j \mid x_{j-1}) > 0\}$ is real line does not dependent on δ .

(2) Conditions on partial derivatives of $f(x_j \mid x_{j-1}) = h(x_j \mid x_{j-1})$ with respect to δ.

$$h(x_j \mid x_{j-1}) = \left\{\frac{\delta(1-|\rho|)}{\pi}\right\} \cdot \frac{1}{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2}$$

$$\frac{dh}{d\delta} = \frac{(1-|\rho|)}{\pi} \cdot \left\{\frac{\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2 - 2\delta^2(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2]^2}\right\}$$

$$= \frac{(1-|\rho|)}{\pi}\left[\frac{(x_j - \rho xj - 1)^2 - \delta^2(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2]^2}\right]$$

$$\frac{d^2h}{d\delta^2} = \frac{2\delta(1-|\rho|)^3}{\pi}\left[\frac{-[\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2 - 2(x_j - \rho x_{j-1})^2 - \delta^2(1-|\rho|)^2]}{[\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2]^3}\right]$$

These derivatives exists and are continuous. Thus the conditions on derivatives of h with respect to δ hold.

Now let

$$g(x_j \mid x_{j-1}) = \log h(x_j \mid x_{j-1})$$

That is

$$g(x_j \mid x_{j-1}) = \log\delta + \log(1-|\rho|) - \log\pi - \log[[\delta^2(1-|\rho|)^2 + (x_j - \rho x_{j-1})^2]$$

$$\frac{dg}{d\delta}=\frac{1}{\delta}-\frac{2\delta(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]}$$

$$\frac{d^2g}{d\delta^2}=\frac{-1}{\delta^2}-\left[\frac{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]2(1-|\rho|)^2-2\delta(1-|\rho|)^2 2\delta(1-|\rho|)^2}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

$$=\frac{-1}{\delta^2}-\left[\frac{2(1-|\rho|)^2(x_j-\rho x_{j-1})^2-2\delta^2(1-|\rho|)^4}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\right]$$

and

$$\frac{d^3g}{d\delta^3}=\frac{2}{\delta^3}-\left\{\frac{2(1-|\rho|)^2[[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2[-2\delta(1-|\rho|)^2]]-}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^4}\right.$$
$$\left.\frac{[(x_j-\rho x_{j-1})^2-\delta^2(1-|\rho|^2].2\delta^2(1-|\rho|)^2[\delta^2(1-|\rho|)^2+(x_j\rho x_{j-1})^2 2\delta(1-|\rho|)^2}{}\right\}$$

Here also it follows that expectation of modulus of all derivatives are finite.

Now Fisher information is given by

$$\mathrm{I}((\delta)=-\mathrm{E}\left[\frac{\partial^2 g(X_j\mid X_{j-1})}{\partial\delta^2}\right]$$

$$=\frac{1}{\delta^2}+\int_{-\infty}^{\infty}\frac{2(1-|\rho|)^2(x_j\rho x_{j-})^2-2\delta^2(1-|\rho|)^4}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}f(x_j/x_{j-1})dx_j$$

$$=$$

$$\frac{1}{\delta^2}+\int_{-\infty}^{\infty}\frac{2(1-|\rho|)^2(x_j\rho x_{j-})^2-2\delta^2(1-|\rho|)^4}{[\delta^2(1-|\rho|)^2+(x_j-\rho x_{j-1})^2]^2}\,\frac{\delta(1-|\rho|)}{\pi(\delta^2(1-|\rho|)^2+(x_j\rho x_{j-1})^2)}dx_j$$

$$= \frac{1}{\delta^2} + \frac{2}{\pi} \int \frac{[(1-|\rho|)^4 \delta^2 u^2 - \delta^2 (1-|\rho|)^4] \delta^2 (1-|\rho|)^2}{[\delta^2 (1-|\rho|)^2 + \delta^2 (1-|\rho|)^2 u^2]^2 \pi [\delta^2 (1-|\rho|)^2 + \delta^2 (1-|\rho|)^2 u^2]}$$

$$= \frac{1}{\delta^2} + \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\delta^4 (1-|\rho|)^6)(u^2 - 1) du}{\delta^6 (1-|\rho|)^6 (1+u^2)^3}$$

$$= \frac{1}{\delta^2} + \frac{2}{\delta^2} \left[\frac{1}{\pi} B(3/2, 3/2) - \frac{1}{\pi} B(1/2, 5/2) \right]$$

$$= \frac{1}{\delta^2} + \frac{2}{\delta^2} \left[\frac{1}{8} - \frac{3}{8} \right]$$

$$= \frac{1}{2\delta^2}$$

The unconditional Fisher Information function is also $\frac{1}{2\delta^2}$.

Result : The mle of δ denoted by $\hat{\delta}$ is consitant and asymptotically normal for δ.

That is, $\hat{\delta} \xrightarrow{P} \delta$ as $n \to \infty$ and

$$\sqrt{n}(\hat{\delta}_n - \delta) \xrightarrow{L} N(0, 2\delta^2)$$

or $\hat{\delta}_n \sim \text{AN}(\delta, \frac{2\delta^2}{n})$, by (ii) of Lemma (3.3.1).

Where $\xrightarrow{L}$ means converges in law and $\xrightarrow{P}$ means converges in probability.

3.4 ALTERNATIVE ESTIMATORS FOR δ AND ρ

The discussion on the previous section shows that mles of δ and ρ do not have closed form expressions. Further numerical solutions also have some problems as we can see in the next chapter. So in this section we propose some alternative estimators for δ and ρ. For the estimation of ρ we used the method proposed by Brockwell and Davis (1987, p.480). They discussed the problem of estimation in the context of time series with infinite variance. The estimator of ρ is given by

$$\hat{\rho} = \frac{\sum X_t X_{t+1}}{\sum_{t=1}^{n} X_t^2} \tag{3.4.1}$$

This $\hat{\rho}$ resembles the sample autocorrelation function. But in our case the moments of X_t do not exist and hence it is not proper to call $\hat{\rho}$ as an autocorrelation function. Brockwell and Davis (1987) propose this estimator for ρ when the innovation has an infinite variance. The asymptotic properties of $\hat{\rho}$ are stated in the following theorem whose proof is given in Brockwell and Davis (1987, p.482).

Lemma 3.4.1 : Let $\{\varepsilon_n\}$ be a sequence of independently and identically distributed Cauchy random variables and let $\{X_n\}$ be a stationary process defined by (3.2.1) then

$$\left(\frac{n}{\log n}\right)(\hat{\rho}-\rho) \Rightarrow Y_1 \tag{3.4.2}$$

where $Y_1 = (1-2\rho^2)\frac{S_1}{S_0}$, S_1 and S_0 are independent random variables

having characteristic functions

$$E(e^{itS_0}) = \exp\{-\frac{2\delta}{\pi})\overline{1/2}\cos(\pi/4)|t|^{1/2}(1-isig(t)\tan(\pi/4)\} \tag{3. 4.3}$$

and

$$E(e^{itS_1}) = \exp\{-2\delta^2 \mid t \mid\} \quad , \tag{3.4.4}$$

then it follows that $\hat{\rho} \xrightarrow{P} \rho$.

This rate of convergence to zero compare favorably with the slower rate in the finite variance case. The asymptotic distribution of $\hat{\rho}$ is same as that of Y_1 and this is distributed like $(1-2\rho^2)\frac{U}{V}$ where V and U are independent random variables having characteristic functions given by (3.4.3) and (3.4.4) respectively with C=1. Percentiles of the distribution of U/V can be found either by simulation of independent copies of U/V or by numerical integration of the joint density (U,V) over an appropriate region . U is a Cauchy random variable with density

$$f_U(u) = \frac{1}{2}\left[\frac{\pi^2}{4} + u^2\right]^{-1},$$

and V is non-negative random variable with density

$$f_V(v) = \frac{1}{2} v^{-3/2} e^{-\pi/(4v)} \qquad v \geq 0.$$

The distribution function of U/V is given by

$$P\left(\frac{U}{V} \le x\right) = \int_0^\infty P(U \le xy) f_v(y) dy$$

$$= \int_0^\infty \left(\frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{xy}{\pi/2}\right)\right)\frac{1}{2} y^{-3/2} e^{-\pi/(4y)} dy$$

$$= \int_0^\infty \left(\frac{1}{2} + \frac{1}{\pi}\tan^{-1}(xz)\right)\frac{1}{2}\left(\frac{\pi z}{2}\right)^{-3/2} e^{-\left(\frac{1}{2z}\right)} dz$$

$$= \int_0^\infty 2^{-1/2}(\pi z)^{-3/2}\left(\frac{\pi}{2} + \tan^{-1}(xz)\right) e^{-\left(\frac{1}{2z}\right)} dz,$$

where U/V has the distribution same as that of the product of a standard Cauchy random variable and independent chi-square with one degree of freedom.

Lemma 3.4.2 : Let the stationary sequence $\{X_n\}$ be strongly mixing, with $\sum_{m=1}^{\infty} \alpha(m) < \infty$ and let X_j be bounded ; $P(|X_j|<\infty)=1$. then

$$\sigma^2 = E(X_0^{\ 2}) + 2\sum_{j=1}^{\infty} E(X_0 X_j) < \infty$$

and if σ#0

$$\lim_{n\to\infty} P[\sigma^{-1} n^{-1/2} \sum_{j=1}^{n} X_j < z] = \phi(z),$$

where $\phi(z)$ is the d.f.of a normal r.v. with mean $E(X_j)$ and variance 1.

For a proof see Ibragimov and Linnik (1971, pp.347).

For estimating δ we use the method of empirical distribution function (edf). Let $\{X_n\}$ be a Cauchy AR (1) sequence defined by (3.2.1) and define

$$U_j = 1 \text{ if } |X_j| \le t$$

$$= 0 \text{ otherwise,}$$

where $\{U_j\}$ is a sequence of Bernoulli r.vs. and

$$E(U_j) = P[|X_j| \le t]$$

$$= P[-t \le X_n \le t]$$

$$= F(t) - F(-t)$$

$$= \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) - \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left(\frac{-t}{\delta}\right)$$

That is

$$E(U_j) = \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) \qquad \text{since} \quad \tan(-x) = -\tan(x)$$

$$\mathrm{Var}(U_j) = E(U_j^2) - [E(U_j)^2]$$

$$= \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) - \frac{4}{\pi^2}\{\tan^{-1}\left(\frac{t}{\delta}\right)\}^2 .$$

Since $\{X_n\}$ is strictly stationary and strong mixing sequence with mixing parameter $\alpha(m)$ (see theorem 3.2.2) , it follows that $\{U_n\}$ is also strictly stationary and strong mixing sequence with mixing parameter

$\alpha(m) = 2|\rho|^{m-1}$, m=1,2,3,. . . , also $\{U_n\}$ is a sequence of random variables which is almost surely uniformly bounded by unity and $\sum_{m=1}^{\infty}\alpha(m) < \infty$.

Then by Lemma 3.4.2

we have

$$\frac{1}{n}\sum_{j=1}^{n}[U_j - E(U_j)] \xrightarrow{L} Z$$

where Z follows Normal distribution with mean zero and σ^2, where

$$\sigma^2 = Var(U_0) + 2\sum_{h=1}^{n} Cov(U_0, U_h) > 0$$

and $\xrightarrow{L}$ denotes converges in distribution. That is,

$$\overline{U}_n \sim AN(E(U_0), \sigma^2 / n)$$

means $\overline{U}_n$ is asymptotically normal with mean $E(U_0)$ and variance σ^2 / n.

The covariance is given by,

$$Cov(U_0, U_h) = E(U_0U_h) - E(U_0)E(U_h),$$

where

$$E(U_0U_h) = P[U_0 = 1, U_h = 1]$$

$$= P[|X_0| \le t, |X_h| \le t]|$$

$$= P[-t \le X_0 \le t, -t \le X_h \le t]$$

$$= \int_{-\infty}^{\infty} P[-t \le X_0 \le t, -t \le \rho^n X_0 + \rho^{n-1}\varepsilon_1 + ... + \rho\varepsilon_{n-1} + \varepsilon_n \le t] f_X(x)dx$$

$$= \int_{-\infty}^{\infty} P[-t \le X_0, \le t, -t \le \rho^n X_0 + Z_n \le t \mid X_0 = x] f_X(x)dx f_X(x)dx$$

$$= \int_{-t} P[-t - \rho^n x \le Z_n \le t - \rho^n x] \frac{\delta}{\pi(\delta^2 + x^2)} dx$$

$$= \int [F_\varepsilon(t - \rho^n x) - F_\varepsilon(-t - \rho^n x)] \frac{\delta}{\pi(\delta^2 + x^2)} dx$$

$$= \int_{t} \left[\frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{t - \rho^{n} x}{\delta(1-|\rho|^{n})} \right) - \frac{1}{2} - \frac{1}{\pi} \tan^{-1}\left(\frac{-t - \rho^{n} x}{\delta(1-|\rho|^{n})} \right) \right] \frac{\delta}{\pi(\delta^{2} + x^{2})} dx$$

$$= \frac{\delta}{\pi^{2}} \int_{t} \left[\tan^{-1}\left(\frac{t - \rho^{x} x}{\delta(1-|\rho|^{n})} \right) + \tan^{-1}\left(\frac{t + \rho^{n} x}{\delta(1-|\rho|^{n})} \right) \right] \frac{1}{\delta^{2} + x^{2}} dx$$

$$= \frac{\delta}{\pi^{2}} \int_{t} \tan^{-1}\left(\frac{\left(\frac{2t}{\delta(1-|\rho|^{n})} \right)}{1 - \left(\frac{t^{2} - \rho^{2n} x^{2}}{\delta^{2}(1-|\rho|^{n})^{2}} \right)} \right) \frac{1}{\delta^{2} + x^{2}} dx \qquad (3.4.5)$$

Since $\tan^{-1}(x) + \tan^{-1}(y) = \tan^{-1}\left(\frac{x+y}{1-xy} \right)$.

That is,

$$E(U_0U_h) = \frac{2\delta}{\pi^{2}} \int_{0} \tan^{-1}\left(\frac{2t\delta(1-|\rho|^{n})}{\delta^{2}(1-|\rho|^{n})^{2} - t^{2} + \rho^{2n} x^{2}} \right) \frac{1}{\delta^{2} + x^{2}} dx$$

$$= \frac{2\delta}{\pi^{2}} \int_{0}^{\rho^{n}} \tan^{-1}\left(\frac{2t\delta(1-|\rho|^{n})}{\delta^{2}(1-|\rho|^{n})^{2} - t^{2} + y^{2}} \right) \frac{\rho^{2n}}{\rho^{2n}\delta^{2} + y^{2}} \frac{dy}{\rho^{2n}} \qquad (3.4.6)$$

From (3.4.6) we see that $E(U_0U_h)$ can be approximated by

$$E(U_0, U_h) \approx \frac{2\delta}{\pi^{2}} \int_{0} \tan^{-1}\left(\frac{2t\delta}{\delta^{2} - t^{2}} \right) \frac{1}{\delta^{2} + x^{2}} dx$$

$$= \frac{2}{\pi}\tan^{-1}\left(\frac{2t\delta}{\delta^2 - t^2}\right)\frac{1}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right).$$

That is

$$\sigma^2 \approx \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) - \frac{4}{\pi^2}\left(\tan^{-1}\left(\frac{t}{\delta}\right)\right)^2 + \frac{4n}{\pi^2}\tan^{-1}\left(\frac{t}{\delta}\right)\tan^{-1}\left(\frac{2t\delta}{\delta^2 - t^2}\right) - 2n\left(\frac{4}{\pi^2}\right)\left[\tan^{-1}\left(\frac{t}{\delta}\right)\right]^2$$

(3.4.6a)

The scale parameter δ can be estimated using the method of moments as follows.

We have

$$\overline{U}_n = \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right),$$

implies that

$$\frac{\pi}{2}\overline{U}_n = \tan^{-1}\left(\frac{t}{\delta}\right)$$

or
$$\frac{t}{\delta} = \tan\left(\frac{\pi}{2}\overline{U}_n\right)$$

that is
$$\hat{\delta} = \frac{t}{\tan\left(\frac{\pi}{2}\overline{U}_n\right)} .=G_n(t), \text{ say.} \qquad (3.4.7)$$

Note that $0 < \overline{U}_n < 1 \Rightarrow 0 \le \frac{\pi}{2}\overline{U}_n < \frac{\pi}{2} \Rightarrow t > 0.$

$$E[\hat{\delta} - \delta]^2 = E\left[\frac{t}{\tan\left(\frac{\pi\overline{U}_n}{2}\right)} - \delta\right]^2.$$

We have

$$\overline{U}_n \to AN(E(U_0), \frac{\sigma^2}{n})$$

Let $$g(x) = \frac{t}{\tan\left(\frac{\pi}{2}\overline{U}_n\right)} \qquad ,0<x<1$$

Then $$g'(x) = \frac{-\pi t}{2}\operatorname{cosec}^2\left(\frac{\pi x}{2}\right)$$

Therefore

$$g(\overline{U}_n) \sim AN(\delta, \frac{\sigma^2}{n}\frac{\pi^2 t^2}{4}\frac{1}{\sin^4\left(\frac{\pi\mu}{2}\right)}) \quad , \mu = \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) \qquad (3.4.8)$$

where

$$\sigma^2 = \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right) - \frac{4}{\pi^2}\left(\tan^{-1}\left(\frac{t}{\delta}\right)\right)^2 + \frac{4n}{\pi^2}\tan^{-1}\left(\frac{t}{\delta}\right)\tan^{-1}\left(\frac{2t\delta}{\delta^2 - t^2}\right) - 2n\left(\frac{4}{\pi^2}\right)\left[\tan^{-1}\left(\frac{t}{\delta}\right)\right]^2.$$

Hence we have proved the following theorem.

Theorem 3.4.4 : The estimator $G_n(t)$ defined by (3.4.7) is consistent and asymptotically normal for δ and the asymptotic variance is specified in (3.4.8).

Thus $G_n(t)$ is CAN for any $t>0$. However, for a given situation we have to specify t for estimation. A method for choosing t is described in the next chapter.

CHAPTER – 4

APPLICATION OF CAUCHY AR (1) MODELS

4.1. Introduction

This chapter discusses the applications of Cauchy AR(1) models. The Section 2 is a simulation study to investigate the performance of the estimators discussed in Chapter 3 while the Section 3 suggests an alternative method for choosing t and Section 4 deals with real data analysis.

4.2 Simulation Study

In stochastic simulations, random numbers are used to generate random variables from a specified distribution in order to characterize the system behavior. Computer simulation methods are widely used to generate random variables. Most of the computers have built-in pseudo random number generator which produce a sequence of random numbers using a recursive formula. The user is required only to input an initial value X_0, and then it produces a realization of independent uniform (0,1) variates. The inverse-transform method for generating random variables is as follows.

Let X be a random variate with cumulative distribution function (cdf) F(x). Since F (x) is a non-decreasing function, the inverse function is defined as follows

$$F^{-1}(y) = \inf\{x; F(x) \geq y\} \qquad , \; 0 \leq y \leq 1.$$

Suppose that U has a Uniform distribution over (0,1,) and let U=F(x) then X=F^{-1} (U). In view of the fact that F is invertible,

$$P(X \leq x) = P(F^{-1}(U) \leq x)$$

$$= P(U \leq F(x)) = F(x).$$

Then to generate a value, say x, of a random variate X with cdf F (x), first generate the uniform variate U, compute $F^{-1}(u)$ and set equal to x. Therefore in order to generate Cauchy random variates, first generate a set of uniform random variable U and then the standard Cauchy variates with

$$C = \tan(\pi(U - 1/2)) .$$

After generating the standard Cauchy variables C(0,1), the Cauchy random variable X with parameters μ and δ is obtained by the formula

$$X = \mu + \delta . C(0,1)$$

We then simulate realization from an AR(1) sequence for specified values of the parameters. Let $\{X_n\}$ be a Cauchy AR (1) sequence defined as

$$X_n = \rho X_{n-1} + \varepsilon_n, n=1,2,3,... \; 0 < |\rho| < 1, \qquad (4.2.1)$$

where $\{\varepsilon_n\}$ is an independent and identically distributed sequence of $C(\delta(1-|\rho|),0)$ r.vs. and X_0 is a $C(\delta,0)$ random variable with distribution function

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x}{\delta}\right) \qquad (4.2.2)$$

We simulate a sample of size n from iid $C(\delta(1-|\rho|),0)$ distribution for specified values of ρ and δ, then generate a realization from $\{X_n\}$ iteratively using (4.2.1) for a given X_0. We use this realization for estimating ρ and δ by discarding the first 50 observations.

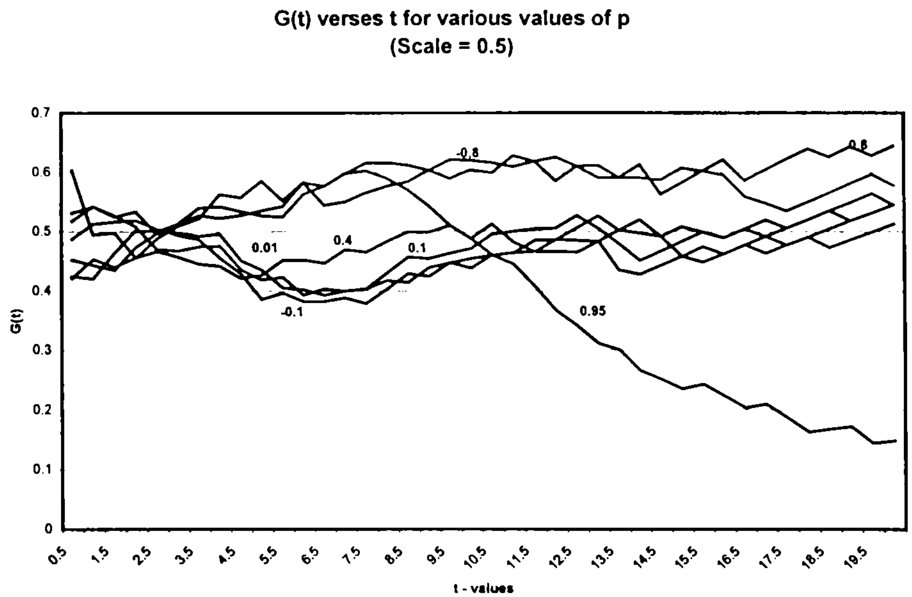
When both ρ and δ are unknown the iterative procedure for maximum likelihood estimation discussed in Section 3.3 of the previous chapter does not converge. If we assume that ρ is known we get good estimators of δ. The computation is summarized in Table 4.1. The table gives the maximum likelihood estimates of the scale parameter δ for different sample sizes n=20, 50, 100 and 500 when ρ is known. From the table it can be seen that the estimate $\hat{\delta}$ of δ performs well even for a small sample of size n=20.

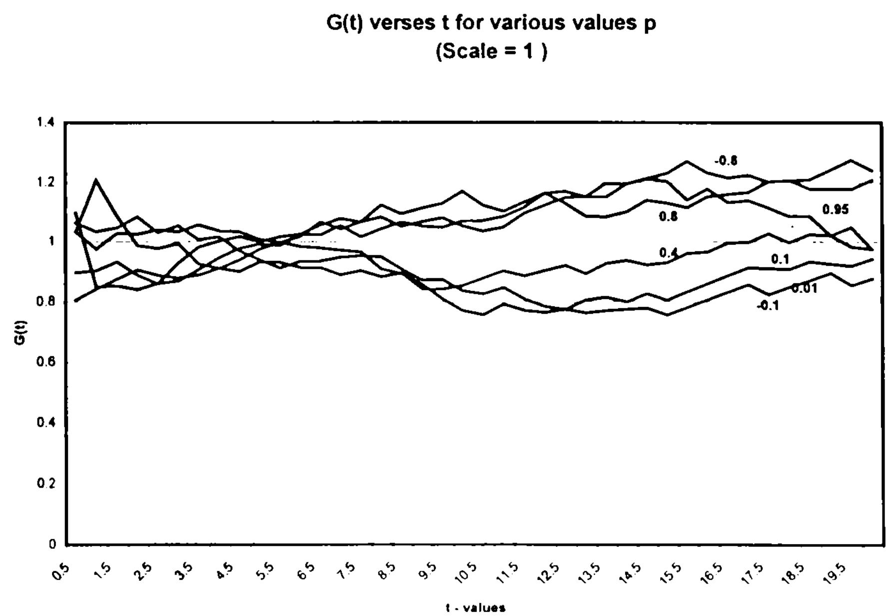
Table 4.1: Estimates of the scale parameter (mle method)

ρ	δ	N=20		N=50		N=100		N=500	
		Est. δ	A.V.	Est. δ	A.V.	Est. δ	A.V.	Est. δ	A.V.
0.1	0.1	.0912	0.0008	0.0963	0.00037	0.0979	0.0002	0.1041	0.00004
0.2	0.5	0.4560	0.0208	0.4815	0.0093	0.4896	0.0048	0.5204	0.00011
0.5	1.0	0.9119	0.08316	0.9791	0.0383	0.9791	0.0192	1.0407	0.0043
0.8	5.0	4.559	2.0784	4.895	0.9584	4.895	0.4792	5.236	0.1097

A.V. –Asymptotic variance of the estimator

Fig 4.1


G(t) verses t for various values of p
(Scale = 0.5)
G(t)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.8
0.8
0.01
0.4
0.1
-0.1
0.95
0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
8.5
9.5
10.5
11.5
12.5
13.5
14.5
15.5
16.5
17.5
18.5
19.5
t - values


G(t) verses t for various values p
(Scale = 1)
G(t)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.8
0.95
0.8
0.4
0.1
0.01
-0.1
0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
8.5
9.5
10.5
11.5
12.5
13.5
14.5
15.5
16.5
17.5
18.5
19.5
t - values

To study the behavior of the estimates discussed in Section 3.4, we simulate a sample of $\{X_n\}$ as discussed above and estimate ρ and δ using the formulae (3.4.1) and (3.4.7) respectively. Note that the estimator $G_n(t)$ of δ in (3.4.7) is a function of t (t>0). The estimator $G_n(t)$ of δ is consistent and asymptotically normally distributed (CAN) estimate for every t. However to estimate the values of δ for a given sample, we need to know the values of t. We followed a graphical method (proposed by Abraham and Balakrishna (1999)) by plotting $G_n(t)$ versus t for choosing the optimum values of t. The graphs for various values of ρ and δ are given in Fig 4.1. After analyzing these graphs for various values of ρ and δ, we arrived at the following formulae to estimate δ, that is, the estimate of δ is given by

$$\hat{\delta} = \underset{t>0}{Sup}(G_n(t)) \quad \text{if } 0<|\rho|<1.$$

The Table 4.2 shows the estimates of ρ and δ.

Table 4.2 Estimates of the scale parameter using the empirical d.f. (sample size n=950)

ρ	δ=0.1		δ=0.5		δ=1	
	t	Est. δ	t	Est. δ	t	Est. δ
-0.8	2.00	0.124	12.00	0.625	19.5	1.273
-0.1	19.00	0.141	20.00	0.545	3.50	1.057
0.01	20.00	0.148	20.00	0.512	1.50	1.095
0.1	19.00	0.141	19.25	0.564	2.00	1.074
0.4	12.5	0.155	5.00	0.526	1.00	1.200
0.8	4.00	0.129	20.00	0.644	20.00	1.217
0.95	1.50	0.120	7.50	0.602	14.50	1.215

4.3 Alternate method for choosing t

This Section gives an alternate method for choosing the value of t. Since most of the estimates of the scale parameter given in Table 4.2 are over estimates we need an alternate method for choosing the value of t. The rest of this section explains the theoretical importance of the method followed by the estimation of the parameter using simulated samples.

The asymptotic variance of $\overline{U}_n$ is approximated by (see 3.4.6a)

$$\sigma^2 \approx \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\delta}\right)-\frac{4}{\pi^2}\left(\tan^{-1}\left(\frac{t}{\delta}\right)\right)^2+\frac{4n}{\pi^2}\tan^{-1}\left(\frac{t}{\delta}\right)\tan^{-1}\left(\frac{2t\delta}{\delta^2-t^2}\right)-2n\left(\frac{4}{\pi^2}\right)\left[\tan^{-1}\left(\frac{t}{\delta}\right)\right]^2.$$

$$\frac{\sigma^2}{n}\to\frac{4}{\pi^2}\tan^{-1}\left(\frac{t}{\delta}\right)\tan^{-1}\left(\frac{2t\delta}{\delta^2-t^2}\right)-\left(\frac{8}{\pi^2}\right)\left[\tan^{-1}\left(\frac{t}{\delta}\right)\right]^2 = \text{AV(t), say.}$$

Note that as t→∞

$$\text{AV(t)}\to\frac{4}{\pi^2}\tan^{-1}(1)\tan^{-1}(\infty)-\left(\frac{8}{\pi^2}\right)\left[\tan^{-1}(1)\right]^2$$

$$=\frac{4}{\pi^2}.\frac{\pi}{4}.\frac{\pi}{2}-\frac{8}{\pi^2}\frac{\pi^2}{16}=0.$$

Hence AV (t)→0 as t→∞. That is the asymptotic variance of $\overline{U}_n$ is minimum when t=δ.We use this information to choose "t" iteratively . $G_n(t)$ is a function of $\overline{U}_n$ and hence the minimum of AV($G_n(t)$ and AV($\overline{U}_n$) are attained at the same point. Therefore, we compute the estimate of the scale parameter δ by choosing the value of t for which |t-δ| is very small.. We perform the estimation using the simulated sample described above and the values are given in the following Table 4.3.

Table 4.3 Estimates of the scale parameter using the empirical d.f.

(sample size n=950)

ρ	δ=0.1		δ=0.5		δ=1	
	t	Est. δ	t	Est. δ	t	Est. δ
-0.8	0.082	0.082	0.409	0.409	0.818	0.819
-0.1	0.095	0.095	0.472	0.472	0.943	0.944
0.01	0.104	0.103	0.519	0.519	1.038	1.039
0.1	0.104	0.104	0.522	0.522	1.042	1.043
0.4	0.117	0.115	0.588	0.588	1.175	1.176
0.8	0.091	0.091	0456	0.456	0.912	0.913
0.95	0.080	0.090	0.416	0.416	0.899	0.900

These estimates are better than those given in table 4.2.

4.4 PRACTICAL EXAMPLE

This section deals with application of the Cauchy AR (1) model to the daily coconut oil prices at Cochin market. In Chapter 2 we discussed the application of the Gaussian non-linear models to the monthly average coconut oil prices. Here we try to fit a non-Gaussian model to the actual daily prices of coconut oil. This Section gives a brief description of coconut, coconut oil its market situation, seasonal fluctuations etc. followed by the application of the Cauchy AR (1) process.

Coconut palm is one of the most useful trees in the world. The major producing regions of the coconuts are concentrated mainly in Asia and India ranks first (APCC Coconut Statistical Year Book, 1997) in the production of coconuts. In India, Kerala accounts for about 53 per cent of the area and 44 per cent (Directorate of Economics and Statistics, Kerala

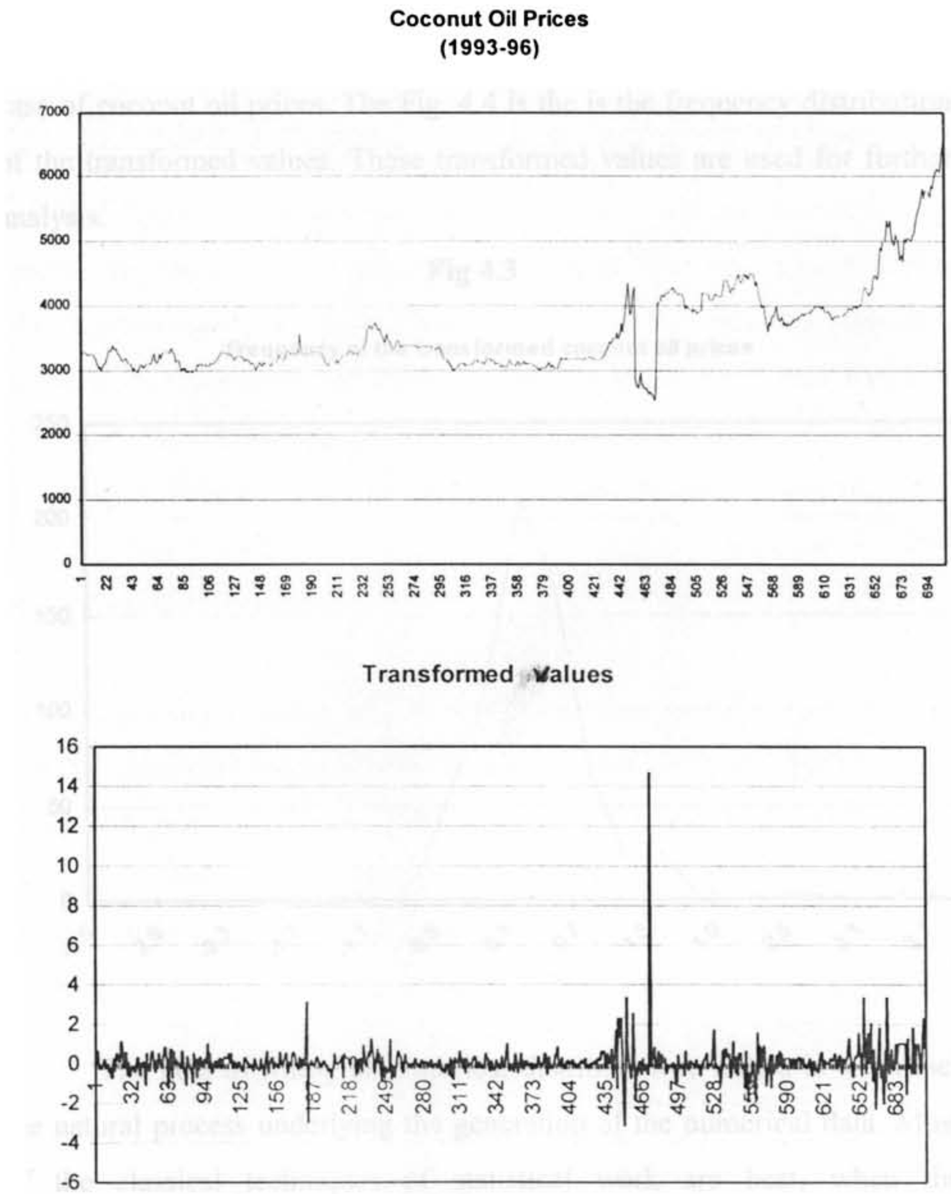
State) of the coconut production. Also coconut makes a significant contribution to the national economy. Coconut provides a variety of products. The nuts are consumed either as such or dried to produce copra which after crushing yields coconut oil and oil cake. Coconut oil is a major vegetable oil used in every household in Kerala for culinary and toiletry purposes. A major portion of the nuts produced are disposed in the form of nuts itself by cultivator. The growers generally sell these products to village merchants or to the agents of the wholesale merchants. Sales generally take place in the garden itself. The agents engaged in the distribution of coconut oil are oil millers, wholesale merchants, commission agents and brokers. Alleppey, Cochin and Calicut are the major important markets in Kerala, while Bombay and Calcutta are outside markets. The prices at Alleppey and Cochin markets are almost same while the prices at Calicut market is slightly higher (Jacob Mathew, 1978).

Since the production of coconuts involves large investment and long germination period, stability in coconut prices is necessary for the development of the crop. The markets of coconut, coconut oil and copra are well integrated and prices also found to move together very closely. Similarly, between markets also there is a strong association. In view of the close relationship between the prices of different coconut products and prices at different markets further analysis is restricted to the prices of coconut oil, which is the end product, at Cochin market alone.

Apart from the general variations seasonal fluctuations within the year can also seen in coconut oil prices. About sixty per cent of the nuts harvested during the first six months of the year, similarly a major portion of the oil is also produced during this season. Because of the abundant availability of the oil in the market, prices are generally low during the summer months and during the early periods of monsoon. During the later half, the production of coconuts and the availability of copra are low and this causes the market prices of coconut oil to rise. During June to October the prices are low because of the lack of facilities for conversion of nuts to copra during this period. The prices rise from October onwards because of the heavy demand of oil and low supply position. Though there is abundant supply of nuts during first half of the year, the prices have fallen below normal because of the heavy demand from the oil millers. Therefore, it is clear that the seasonal variation of the prices of coconut oil is more due to the demand factor than due to supply factors (Jacob Mathew, 1978). Detailed studies in connection with the behavior of coconut oil prices were done by Jacob Mathew (1978,1980 and 1984) and Das (1986,1990,1991). In this study a time series model based approach has been tried to explain the variations other than the trend and the seasonal fluctuations.

The data consist of the daily coconut oil prices at Cochin market from January 1994 to December 1996. The data is given in Appendix I. That is, altogether there are 707 observations. The analysis is concerned in the daily prices recorded on each trading day. Since consecutive prices are highly correlated a direct analysis is difficult. Also the series is not stationary. Consequently it is more convenient to analyze change in prices. Results for such changes can easily be used to give appropriate results for prices. From 1994 data, it appears that the prices increase almost in a constant rate. The following transformation (4.3.1) is used to make the series stationary. The Fig 4.3 shows the plots of the original and the transformed prices.

Fig 4.3



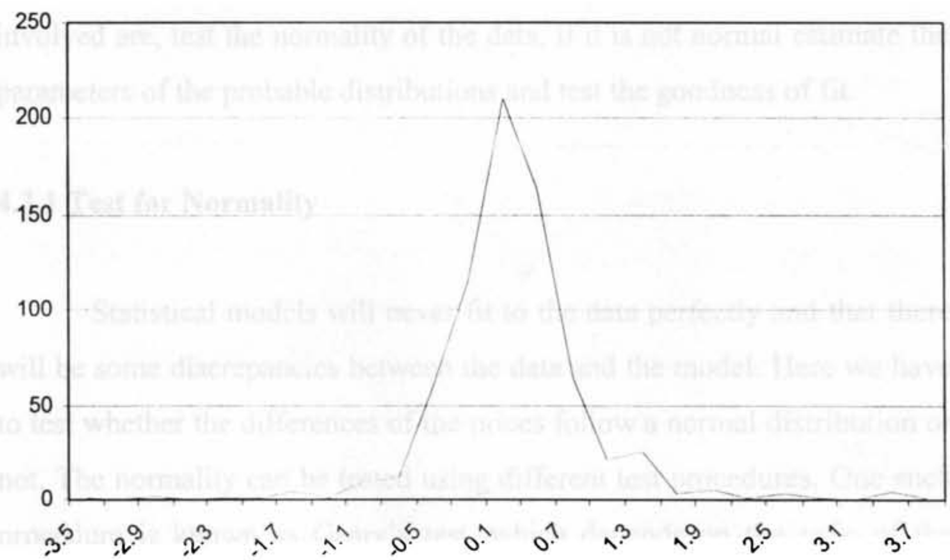
Let Z_n be the price on trading day n, then the price change may be defined by

$$X_n = \frac{Z_n - Z_{n-1}}{sd(Z_n - Z_{n-1})}, \qquad (4.3.1)$$

where sd $(Z_n - Z_{n-1})$ is the standard deviation which is equal to 90.4 in the case of coconut oil prices. The Fig. 4.4 is the is the frequency distribution of the transformed values. These transformed values are used for further analysis.

Fig 4.3

Frequency of the transformed coconut oil prices



The goal of fitting an empirical distribution to a data is to extract the natural process underlying the generation of the numerical data. Most of the classical techniques of statistical work are best, when the assumptions about the nature of the data are met. For example, it is often assumed that the observed series is a realization from a Gaussian sequence. The classical procedure gives quite misleading results if the assumptions are not satisfied by the data in hand. Therefore, in recent times researchers are used to check that the given set of data satisfies the assumptions of the classical procedure before go for further analysis.. The histogram is a display device to get an idea about the distribution of a set of data. Here, the frequency of transformed coconut oil prices (fig 4.3) shows the similarity of a symmetric distribution. The figure suggests that the distributions like normal, Cauchy or Laplace will be more suitable to the data. In order to fit a standard distribution to the given data, the steps involved are, test the normality of the data, if it is not normal estimate the parameters of the probable distributions and test the goodness of fit.

4.3.1 Test for Normality

Statistical models will never fit to the data perfectly and that there will be some discrepancies between the data and the model. Here we have to test whether the differences of the prices follow a normal distribution or not. The normality can be tested using different test procedures. One such procedure is known as Geary's test, which depends on the ratio of the mean deviation to the standard deviation. This ratio gives the test statistic as

$$G = \frac{\sum_{i=1}^{n} |x_i - \bar{x}| / n}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 / n}},$$

where $\bar{x}$ denotes the sample mean. The distribution of G is tabulated (Table A.5 of Appendix A, Cooper and Weeks, 1988) under the assumption that the data follows a normal distribution. The value of G is always positive, but the nature of the critical region is like that, for both extreme small values and extreme large values of G point to the hypothesis of normality being invalid. Extreme small values and extreme large values of G suggest that the data could be modelled by some other distributions. The distribution of G is tabulated under the assumption that the null hypothesis of normality is valid. If the null hypothesis is valid G is around 0.8. The normality can also be tested using the measure of skewness (Cooper and Weeks, 1988, pp.168)

$$g_1 = \frac{\sqrt{n}\sum_{i=1}^{n}(x_i - \bar{x})}{(\sum_{i=1}^{n}(x_i - \bar{x})^2)^{3/2}},$$

g_1 is zero for normal distribution and both large negative and large positive values constitute the critical region for rejecting the hypothesis of normality (Cooper and Weeks, 1988).

The next step is to test the goodness of fit. Two important test procedures are chi-squares test and Kolmogrov-Smirnov test. Here we use the Kolmogrov-Smirnov test. The chi-squares goodness of fit requires that the data values should be first be arranged in the form of a frequency distribution, but otherwise, its use is straight forward. The Kolmogrov-Smirnov test depends on the cumulative relative frequencies (crf) of the data set to the crf's of the theoretical distribution model. This test is based upon comparing empirical crf's and theoretical crf's having common variable values. It requires that the model distribution should be completely specified with numerical values given to all parameters.

The problem of fit is to test the hypothesis that the sample observations x_1, x_2. . . x_n is from a specified distribution against the alternative that it is from some other distributions. Then H_0: $X_I \sim F_0$ against H_1: $X_I \sim F$, where $F_0(x) \neq F(x)$. Let the sample be from the distribution function F and let F^*_n be the corresponding empirical distribution function. The statistic

$$D_n = \underset{x}{Sup}\left|F_n^*(x) - F(x)\right|$$

is called Kolmogrov-Smirnov statistic. The Kolmogrov-Smirnov test treats the individual observations directly, where as the chi-square discretizes the data and sometimes losses information through grouping. This test is applicable even in the case of small samples but chi-square test is essentially for large samples. It assumes the continuity of the distribution function means that the test provides a more refined analysis of the data.

4.3.2 Test of randomness

The simplest possible hypothesis that we can set up of a series which shows any chance of fluctuation is that it is random. In a random series, the observations are independent and could have occurred in any order. In practice, a mere inspection of the data is enough to discuss such a possibility, but there are cases where we need more accurate test. There are a number of such tests, here we explain only the difference-sign test.

This test may be conducted by counting the number of points where it increases. When there are n number of observations there are n-1 differences. Define

$$X_i = 1 \text{ if } U_{i+1} > U_i$$

$$= 0 \text{ if } U_{i+1} < U_i.$$

For a random series the expected number of points of increases, say c, is

$$E(c) = E[\sum_{i=1}^{n-1} X_i] = \frac{1}{2}(n-1).$$

4.3.3 Results and Discussions

The first step in our analysis is to test the normality using the methods described in the previous section. There are 707 transformed values of the coconut oil prices. The G-statistic for the transformed values of the coconut oil prices is G=0.4956. Since the value of G for a sample size of 707 (n=707) lie outside the acceptance (0.75 to 0.84 for n=101) region the normality assumption is invalid. The Fig 4.3 suggests that the other possible distributions may be Cauchy or Laplace. Since it is a price data first we go for a Cauchy distribution. To confirm possibility of the distribution we have to test the goodness of fit. The following paragraph gives a small description of the estimation of the parameters of the Cauchy distribution for the given set of transformed coconut oil prices.

For a Cauchy distribution, the median is an estimate of the location parameter, for the price differences the estimated value of location parameter is zero. The scale parameter is estimated using the maximum likelihood method (as given in Johnson and Kotz, 1994). The maximum likelihood equation to estimate the scale parameter is

$$\frac{\partial \log L}{\partial \delta} = \frac{n}{\delta} - \sum_{i=1}^{n} \frac{2\delta}{\delta^2 + x_i^2}$$

and

$$\frac{\partial^2 \log L}{\partial \delta^2} = \frac{-n}{\delta^2} - 2\sum_{i=1}^{n} \frac{(x_i - \delta^2)}{x_i^2 + \delta^2} \qquad (4.3.2)$$

Here x_1, x2, . . . ,x_n are the observations from the Cauchy density in (3.1.1). Since above equations (4.3.2) do not have a closed form expression for its solution, an iterative procedure is used to estimate the scale parameter. The procedure is as follows. The estimator after k iterations is given by

$$\hat{\delta}_{k+1} = \hat{\delta}_k - \frac{1}{n} \cdot \frac{\left(\partial \log L \big/ \partial \delta_k\right)}{\left(\partial^2 \log L \big/ \partial \delta_k^{\,2}\right)}$$

Table 4.5 gives the values of the empirical distribution function and the theoretical distribution functions. Since the calculated value of the Kolmogrov statistic (0.0488) is less than the corresponding tabulated value (0.0613) value, we can accept the hypothesis that the price differences follow a Cauchy distribution with scale parameter 0.243 and location parameter zero, that is C(0.243,0).

The next step is to test the randomness of the series. We proceed as above (see Section 4.3.2) and estimated the number of point where it increases. The estimated value is 325. If it is a random series the expected number is 353. Therefore, we can conclude that the observations are not independent. The scatter diagram (Fig. 4.4) of X_t and X_{t-1} also reject any chance of independence of successive observations. From above steps it is clear that the distribution of the transformed values of the coconut oil prices follows a Cauchy distribution and the consecutive prices are correlated. Thus a time series based modelling technique with Cauchy marginal distribution is an appropriate choice for the prices. The parameters are estimated using the methods described in the previous chapter. The Table 4.4 gives the estimated values of the parameters (see Fig.4.5).

Table 4.4 : Estimates of the different parameters of transformed Coconut oil prices

Estimates of the AR coefficient	0.033277
Estimate of the scale parameter mle method	0.2513
,, edf method of Section 4.2	0.3235 (t=0.155)
,, edf method of Section 4.3	0.2528 (t=0.208)

Finally the model identified is

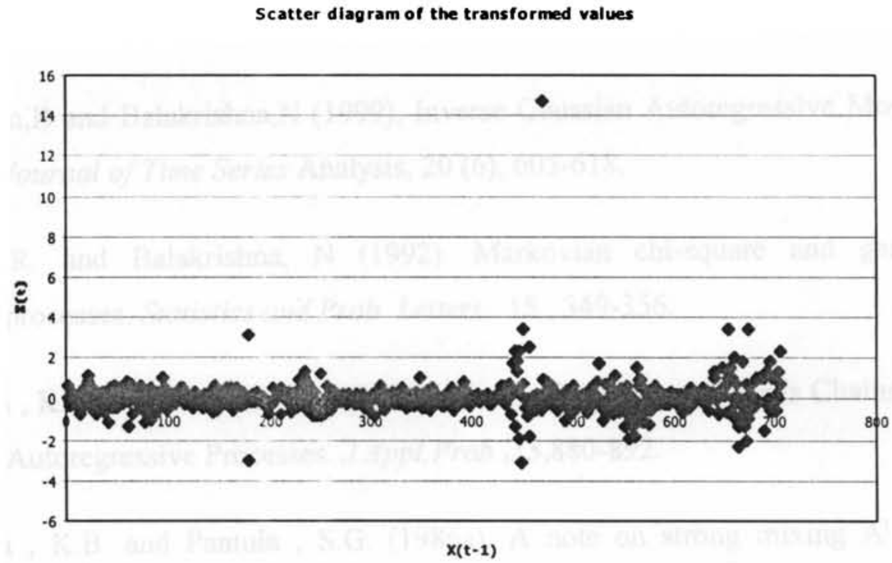
$$X_n = 0.033277 X_{n-1} + \varepsilon_n,$$

where $\{\varepsilon_n\}$ is a sequence of independently and identically distributed random variables and follow $C(\delta.(1-0.033277),0)$, where δ is either of the estimators obtained above. However, if we are following edf method then we recommend $\hat{\delta} = 0.2528$ as the estimate for δ.

Table 4.5: values of the empirical and theoretical distribution functions

Class	Frequency (f)	Cum. Frequency	Empirical d.f. F^*	Theoretical F	$\lvert F-F^*\rvert$
<-3.5	1	1	0.00141	0.0223	0.0208
-3.2	0	1	0.00141	0.0243	0.0229
-2.9	2	3	0.00424	0.0268	0.0226
-2.6	0	3	0.00424	0.0299	0.0256
-2.3	1	4	0.00566	0.0337	0.0280
-2.0	1	5	0.00707	0.0387	0.0316
-1.7	4	9	0.0127	0.0454	0.0326
-1.4	1	10	0.0141	0.0549	0.0407
-1.1	10	20	0.0283	0.0694	0.0411
-0.8	12	32	0.0453	0.0940	0.0488^{**}
-0.5	59	91	0.1287	0.1441	0.0154
-0.2	115	206	0.2914	0.2809	0.0105
0.1	210	416	0.5884	0.6242	0.0358
0.4	163	579	0.8189	0.8261	0.0071
0.7	64	643	0.9095	0.8935	0.0159
1.0	21	664	0.9392	0.9240	0.0152
1.3	25	689	0.9745	0.9410	0.0335
1.6	3	692	0.9859	0.9518	0.0269
1.9	5	697	0.9873	0.9593	0.0265
2.2	1	698	0.9915	0.9648	0.0225
2.5	3	701	0.9929	0.9690	0.0225
2.8	1	702	0.9929	0.9723	0.0206
3.1	0	702	0.9986	0.9749	0.0180
3.4	4	706	1.0000	0.9771	0.0215
3.7	1	707	1.0000	0.9789	0.0216

Fig 4.4


Scatter diagram of the transformed values
16
14
12
10
8
6
4
2
0
-2
-4
-6
100
200
300
400
600
700
800
X(t-1)

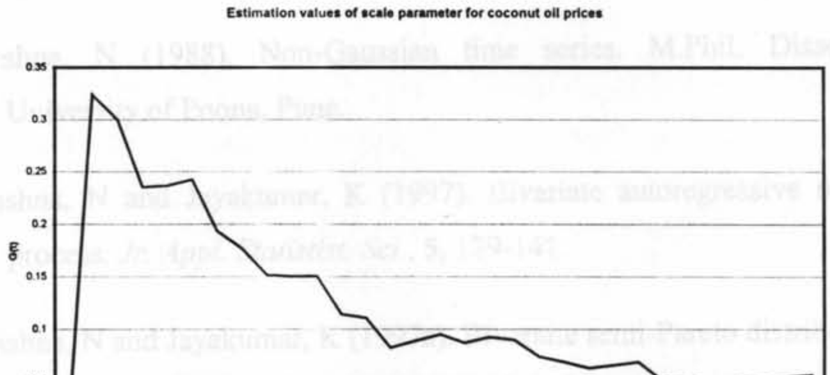

Estimation values of scale parameter for coconut oil prices
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
t-values

Fig 4.5

REFERENCES

Abraham,B and Balakrishna,N (1999). Inverse Gaussian Autoregressive Models. *Journal of Time Series* Analysis, 20 (6), 605-618.

Adke,S.R. and Balakrishna, N (1992). Markovian chi-square and gamma processes. *Statistics and Prob. Letters,* **15** , 349-356.

Athreya , K.B. and Pantula , S.G. (1986). Mixing Properties of Harris Chains and Autoregressive Processes. *J.Appl.Prob.*,**23**,880-892.

Athreya , K.B. and Pantula , S.G. (1986a). A note on strong mixing ARMA processes. *Statistics & Prob.* Letters, **4** , 187-190.

Balakrishna, N (1988). Non-Gaussian time series. M.Phil. Dissertation , University of Poona, Pune.

Balakrishna, N and Jayakumar, K (1997). Bivariate autoregressive minifiction process. *Jr. Appl. Statistist. Sci.*, **5**, 129-141.

Balakrishna, N and Jayakumar, K (1997a). Bivaraite semi-Pareto distributions and processes. *Statistical Papers*, **38**(2), 149-165.

Balakrishna, N. (1999). Non-Gaussian Time series – A Revie. *Statistical Methods*, **1**(1), 83-95.

Billingsley,P. (1961).*Statistical Inference For Markov Process.* University of Chicago Press.

Bollerslev, T, Chou, R.V. and Kroner, K.F. (1992). ARCH modeling in Finance. Areview of the theorey and numerical evidence. *Journal of Econometric* , **52** , 5-59.

Borah, D.K. and Bora, P.K. (1995). Predicting the rainfall around Guwahati using aseasonal ARIMA model. *Joyurnal of Indian Society of Agricultural Statistics*, 47(3), 278-287.

Box, G.E.P. and Jenkins, G.M. (1970) *Time Series Analysis : Forecasting and Control*. San Francisco : Holden-Day.

Brockwell, P.J. and Davis, R.A. (1987). *Time Series : Theorey and Methods* , Springer-Varlag, New York.

Chan, W.S. (1990). On test for non-linearity in Hong Kong stock returns. *Hong Kong J. Bus. Management*, **8**, 1-11.

Cline, D.B.H. and Brockwell, P.J. (1985). Linear Prediction of ARMA processes with infinte variance. *Stochastic Processes and their Appl.*, **19**, 281-296.

Cooper, R.A. and Weeks, A.J. (1988). Data, Models and Statistical Analysis. Heritage Publishers , New Delhi.

Das, P.K. (1986). Movements of whole sale prices of coconuts, copra and coconut oil in Kerala during the last two and half decades. *J. Plantn. Crops*, **14**(2), 1054-114.

Das, P.K. (1990). Sate of art of coconut oil production and utilization in India. *Coconuts Today*, 5 , 70-73.

Das, P.K. (1991). Price Behaviour of India's coconut sector. CORD, 7 (1) , 1-9.

Enders, W. (1995). *Applied Econometric Time Series.* John Wiley & Sons. New York.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econmometrica*, **50** (4), 987-

Ertel,J.E. and Fowlkes,E.B. (1976) Some algorithms for linear spline and piecewise multiple regressions. *J. Am. Statist. Ass.*, **71**, 640-648.

Gourieroux, C. and Monfort, A (1992). Quantitative threshold ARCH models. *Journal of Econometrics.* **52**, 159-199.

Graver, D.P. and Lewis, P.A.W. (1980). First order autoregressive gamma sequences and point processes. *Adv. Appl. Prob.* , **12** , 727-745.

Helfenstein, U. (1996). Box-Jenkins modeling in medical research., *Statistical Methods in Medical Research*, 5, 2-22.

Ibragimov , I.A. and Linnik, Yu.V. (1971). *Independent and stationary sequences of random variables.* Wolter-Noordhooff Publishing, Groningen.

Jacob Mathew (1978) Trend and Fluctuations in the prices of coconut and coconut oil. M.Phil Dissertation, Jawaharlal Nehru University, New Delhi.

Jacob Mathew (1980). Inter-district variations in the pattern of the prices coconuts, copra and coconut oil in Kerala. *PLACROSYSM III*, 362-373.

Jacob Mathew (1984). Seasonal fluctuations in the prices of coconut, copra and coconut oil in Kerala. In: *PLACROSYM V.* (eds. K.V.A. Bavappa and others), 24-31.

Jayakumar, K and Pillai, R.N. (1993). The first order autoregressive Mittag-Lelfler process. *J. Appl. Prob.*, **30**, 462-466.

John.N.L., Kotz,S. and Balakrishnan,N (1994). *Continuous Univariate Distributions, Vol.1*, 2nd ed., John Wiley & Sons, Inc., New Delhi.

Jose, C.T. (1996). A study of yield variuability in Cocoa. *Journal of Plantation Crop*, **24** (2), 126-129.

Keenan, D.M. (1985). A takey non-additivity-type test for time series non-linearity. Biometrika, **72**, 39-44.

Lawrance , A.J. and Lewis P.A.W. (1981). A new autoregressive time series model in exponential variables NEAR(1). *Adv. Appl. Prob.* , **13** , 826-845.

Lawrance , A.J. and Lewis P.A.W. (1985). Modeling and residual analysis of non-linear autoregressive time series in the exponential variables. *J.R.Statist.,B*,**47**,2,165-202.

Li, W.K. and Lam, K (1995). Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model. *The Statistician*, **44**(3), 333-341.

Mcgee, V.E. and Carleton, W.T. (1970). Piecewise regression. *J. Am. Statist. Ass.* , **65** (331), Application Section, 1109-1124.

Nampoothiri,C.K. and Balakrishna,N.(2000). Threshold Autoregressive Model for a Time Series Data. *Jr. of The Indian Society of Agrl. Statist.*, **53**(2), 151-160.

Nelson, D.B. and Cao, C.Q. (1992). Inequality constrains in the univariate GARCH model. *J. Bus. Econ. Statist.*, **10**, 229-235.

Nichols, D.F. and Quinn, B.G. (1981). Random coefficient autoregressive models, Lecturer Notes in Statistics, No.11, Springer-Waley, New York.

Ozaki, T. and Tong , H (1975). On fitting of non-stationary autoregressive models in time series analysis. In Proc. 8th Hawaii Int. Conf. on System Sciences, pp. 225-226., North Hollywood, Western Periodicals.

Petrucelli, J.D. and Davies, N. (1986). A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrica*, **73**(3), 687-94.

Priestly , M.B. (1965). Evolutionary spetra and non-stationary process. *J.R. Statist. Soc. B*, **27**, 204-237.

Priestly , M.B. and Tong , H (1973). On the analysis of bivariate non-stationary process. *J.R. Statist. Soc., B*, **35**, 155-166, 179-188.

Rabemananjara, R and Zakoian , J.M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, **8**, 31-49.

Sathar Abdul, E.I. (1999). On autoregressive conditional heterscedastic models. M.Phil. Dissertation, Cochin University of Science and Technology, Cochin.

Serfling , R.J.(1980). *Approximation theorems of Mathematical Statistics*. John Wiley & Sons, New York.

Sim, C.H. (1990). First order autoregressive models for gamma and exponential processes. *J. Appl. Prob.*, **27**, 325-332.

Tavares, L.V. (1980). An exponential Amrkovian Stationary Process. *Jr. Appl. Prob.*, **17**, 1117-1120.

Taylor, S (1988). *Modelling Financilal Time Series* , John Wiley & Sons, New York.

Tong , H (1977 b). Discussion of a paper by A.J.Lawrance and N.T.Kottegoda. *J.R.stastist. Soc., A*, **140**, 34-35.

Tong H and Lim, K.S. (1980) Threshold autoregression, limit cycles and cyclical data (with discussion). *J.R. Statist. Soc.B*, **142,** 245-292.

Tong, H (1978). On Thresholld model. In pattern recognition and signal processing. (C.H.Chen, ed.). The Netherlands : Sijthoff and Noordhoff.

Tong, H (1980 a). A view on non-linear time series model building. Time Series (O.D.Anderson, eds.). The Netherlands: Sijthoff and Noordhoff.

Tong,H (1990) Non-linear Time Series *a Dynamic System Approach*. New York: Oxford University Press.

Tsay, R.S. (1986) Non-linearity tests for time series. *Biometrica*, **73**, 461-466

Tsay, R.S. (1989) Testing and modelling threshold autoregressive processes. *J. Am. Statist. Ass.* , **84**, 231-240.

Watier, L. and Richardson, S. (1995) Modelling of an epidimiological time series by a threshold autoregressive model. *The Statistician* , **44**, No.3, 353-364.

Zacks. S (1971). *The Theory of statistical Inferences.* JohnWiley & Sons , New York.