SEQUENTIAL ESTIMATION IN SOME MARKOVIAN MODELS

Thesis Submitted to the

Cochin University of Science and Technology

for the Degree of

Doctor of Philosophy

UNDER THE FACULTY OF SCIENCE

By

T. M. JACOB

DEPARTMENT OF STATISTICS

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COCHIN - 682 022

AUGUST, 1998

CERTIFICATE

*Certified that the thesis entitled "**SEQUENTIAL ESTIMATION IN SOME MARKOVIAN MODELS**" is a bonafide record of work done by Sri.T.M Jacob 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 the award of any degree or title.*

Dr.N.Balakrishna,
Lecturer,
Department of Statistics,
Cochin University of
Science and Technology.

Cochin- 22
August 10, 1998.

Contents

		Page
CHAPTER 1	INTRODUCTION	
1.1	Introduction	1
1.2	Random Coefficient Autoregressive Models	2
1.3	Autoregressive Minification Processes	5
1.4	Statistical Inference for Markov Sequences	7
1.5	Sequential Estimation	9
1.6	Some Useful Definitions and Results	11
1.7	Summary of the Thesis	18
CHAPTER 2	PARAMETER ESTIMATION IN MINIFICATION PROCESSES	
2.1	Introduction	20
2.2	Some Probabilistic Properties of the Model.	24
2.3	Estimation of the Mean	27
2.4	Estimation of k	30
2.5	Estimation of Parameters in Some Special Cases	38
CHAPTER 3	SEQUENTIAL ESTIMATION FOR MINIFICATION PROCESSES	
3.1	Introduction	41
3.2.	Sequential point estimation of mean	42
3.3	Sequential Estimation of *k* in Exponential Minification processes	57
3.4	Sequential Interval Estimation	64

CHAPTER 4 ESTIMATION IN RANDOM COEFFICIENT AUTOREGRESSIVE MODEL

4.1	Introduction	71
4.2	The Model and its Properties	71
4.3	Properties of Least Square Estimators	76

CHAPTER 5 SEQUENTIAL ESTIMATION OF THE MEAN OF RCAR(1) PROCESS

5.1	Introduction	88
5.2	Sequential Point Estimation	89
5.3	Sequential Interval Estimation	102

CHAPTER 6 SEQUENTIAL ESTIMATION OF THE REGRESSION PARAMETER OF RCAR(1) MODEL

6.1	Introduction	106
6.2	Definition and Properties of the Stopping Rule	107
6.3	Properties of the Stopping Rule	112
6.4	Sequential Interval Estimation	123
6.5	Directions of future work	125

REFERENCES 126

CHAPTER 1

INTRODUCTION

1.1 Introduction

The outcome of any experiment or the result of any natural phenomenon depends on many unknown factors, which cannot be completely controlled or measured exactly. It is not possible to explain such situations by deterministic mathematical equations. A better way of studying the behaviour of such phenomena, when the outcomes are affected by many uncertain factors is by using stochastic models. These are the models defined in terms of random variables. For example, suppose that one wants to know the value of X, where X may be the price of certain commodity, or the contents of a reservoir, or the velocity of wind, or the amount of currency notes in the Reserve Bank of India, or the stock of radio active materiel etc.

Note the variable X in the above examples are random variable (r.v.) and they may vary at different time points. If we consider X_n as the value of X, at a time point n, then $\{X_n\}$ can be viewed as a realization of the stochastic process $\{X_n, n \in T\}$, where T is an index set. We take in our studies T as a set of integers.

In the classical setup, the statistical analysis of the data is performed by assuming that $\{X_n, n \in T\}$ is a sequence of independent and identically distributed (i.i.d) r.v.s having some common distribution function (d.f) F. However, even in the examples quoted above, the random variables X_n for different n need not be independent. The dependency among the r.v.s at different time points can be brought out by defining appropriate stochastic models. The purpose of defining such models is to identify the stochastic mechanism which generates the data and then use such models to predict its future behaviour. This involves estimation of the unknown parameters in the model and the study of other related statistical inference. Once we identify the stochastic model, the further analysis can be handled, by the help of the well developed theory of stochastic processes.

One of the important applications of stochastic processes is the analysis of time series. The models used in the classical analysis of time series are all linear in nature. Moreover, the time series $\{X_n\}$ is assumed to be a Gaussian sequence (See Box and Jenkins (1976)). One of the linear stochastic models used in the time series analysis is the p^{th} order autoregressive model defined by

$$X_n = a_1 X_{n-1} + a_2 X_{n-2} + \ldots + a_p X_{n-p} + \varepsilon_n \tag{1.1.1}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d r.v.'s assumed to follow normal distribution and $a_1, a_2, \ldots, a_p$ are constants referred to as autoregressive parameters. However, there are many practical situations where the models are non-linear and non-Gaussian, (see eg. Tong (1990)).

1.2. Random Coefficient Autoregressive Models

Nicholls and Quinn (1982) generalized the model (1.1.1) by allowing a_i's to be random variable to define a random coefficient autoregressive (RCAR) model. The sequence $\{X_n\}$ said to follow a p^{th} order RCAR (RCAR(p)) model if

$$X_n = \sum_{i=1}^{p} \{b_i + \beta_{in}\} X_{n-i} + \varepsilon_n , \; n = 1,2, \ldots \tag{1.2.1}$$

The following assumptions are made on this model:

A_1: $\{\varepsilon_n, n \in T\}$ is a sequence of i.i.d. r.v.'s with mean zero and variance σ^2.

A_2: $\{\beta_n = (\beta_{1n}, \beta_{2n}, \ldots, \beta_{pn}), n \in T\}$ is a sequence of i.i.d random vectors with mean **0** and dispersion matrix Γ.

A_3: The sequences $\{\varepsilon_n\}$ and $\{\beta_n\}$ are statistically independent.

A_4. $b = (b_1, b_2, \ldots, b_p)$ is a vector of real constants.

The model (1.2.1) is fitted to various data sets and shown to be performing well. For example, Nicholls and Quinn (1982) fitted an RCAR(2) model to the lynx data which consist of the annual records of the number of Canadian lynx trapped in the MacKenzie river district of North-Western Canada. Lawarance and Lewis (1985) used a more specific RCAR(2) model which generates a stationary sequence of exponential r.v.s to analyse the wind velocity data.

A score of literature is available on the study of first order Random Coefficient Autoregressive RCAR(1) models. For the sake of future reference we define the model as

$$X_n = \{b + \beta_n\} X_{n-1} + \varepsilon_n, n = 1, 2, \ldots \qquad (1.2.2)$$

where the r.v.'s satisfy the assumption A_1 - A_4 described above with $p = 1$. We say that the model (1.2.2) is stationary if it generates a stationary sequence $\{X_n, n \in T\}$.

The model (1.2.2) has many applications in the study of non-Gaussian time series modelling. The theory of non-Gaussian time series mainly concentrates on obtaining the stationary solution of the model (1.2.2) We say that the model (1.2.2) has a stationary solution if there exists a proper probability distribution for ε_n for a specified distribution of X_n for evey n. A standard technique adopted in obtaining the solution of (1.2.2) is using characteristic functions or Laplace transforms. Without much loss of generality one may take b=0 in (1.2.2). If $\phi(s)$ is the characteristic function of X_n for every n and $\psi(s)$ is that of ε_1, then

$$\psi(s) = \frac{\phi(s)}{\int \phi(\beta s) dG(\beta)}, \qquad (1.2.3)$$

where $G(.)$ is distribution function of β_n and the integral ranges over the support of $G(.)$.

A general discussion on conditions for existence of solutions to this model can be found in Paulson and Uppuluri (1972). The RCAR(1) model for defining a sequence of exponential r.v.'s are discussed by Gaver and Lewis (1980), Lawrance and Lewis (1981), Sim (1990), etc. The similar models are used to generate sequences of gamma r.v.'s by Gaver and Lewis (1980), Sim (1986), Lewis, McKenzie and Hugus (1989), Sim (1990), Adke and Balakrishna (1992a) etc. A discussion on Laplace RCAR(1) models may be found in Dewald and Lewis (1985).

The following are some of the more specific examples where RCAR(1) models are used to describe the practical situations. In these cases it is assumed that b=0. Hence the model is rewritten as

$$X_n = \beta_n X_{n-1} + \varepsilon_n, \qquad n = 1,2, \ldots \qquad (1.2.4)$$

Paulson and Uppuluri (1972) claim that the model (1.2.4) arises in the study of retention of a substance in a system when the substance is periodically introduced in random quantities and the system periodically eliminates a random proportion of this substance. Then one may be interested in the behaviour of the amount of a given substance present in the system at the end of epoch n-1, n=1,2,... with X_0 =0. Suppose an amount ε_n of this substance is introduced during the time interval (n-1, n] and during the same interval a modification of the amount X_{n-1} to $\beta_n X_{n-1}$ take place. Hence the total amount of the substance present at epoch n is X_n described in (1.2.4). More specific example where X_n

denotes the (i) balance of a saving account and (ii) the stock of radio active material at time n are described by Vervaat (1979). In example (i) ε_n denotes the deposit made just before time n and β_n the interest factor which may fluctuate stochastically with time. On the other hand, in (ii) ε_n is the quantity of the radio active material added or taken away just before time n and β_n is the natural decay of radioactivity.

Sim (1986) discussed the application of model (1.2.4) in Hydrolical modelling. In his example X_n and ε_n are content and random inputs of a dam respectively at random time T_n, and U_n is the random decay factor of the storage of the dam between time T_{n-1} and T_n. Sim further assumed that $\{T_n\}$ is a sequence of random times generated by a homogeneous Poisson process and $0 = T_0 < T_1 < \ldots$ Then by taking $U_n = \exp[-b(T_n - T_{n-1})]$, $b \geq 0$, it was proved that X_n has gamma distribution for each $n \geq 0$.

Similar examples may also be found in Andel (1976) and Hutton (1990). A vector valued version of the model and its properties are discussed by Glasserman and Yao (1995).

1.3. Autoregressive Minification Processes

In Section 1.2 we have seen that the solution for the model (1.2.4) exists if and only if $\psi(s)$ defines the characteristic function of a r.v. in the relation (1.2.3). Therefore, to check the existence of the solution, we should have a closed from expression for $\phi(s)$. But there are several standard distributions used in statistical studies, which donot have closed form expression for their characteristic functions. For example, the distributions, such as Pareto, Logistic, Weibull and extreme value type are useful to analyase the variety of real life data. In order to generate a sequence of dependent r.v.'s having any of these marginal distributions, the model of the type (1.2.4) are not of much use [see eg. Tavares (1977, 1980)]. As an alternative, a model of the following type is used for this purpose when the r.v.'s have closed form expressions for their survival functions.

Let $\{Z_n\}$ be a sequence of i.i.d r.v.'s with common d.f $G(.)$ and X_0 be a r.v. having d.f. F and is independent of Z_1. Now define X_n by

$$X_n = \begin{cases} X_0, & n = 0 \\ k \min(X_{n-1}, Z_n), & n = 1,2,\ldots \end{cases} \tag{1.3.1}$$

where $k>1$. The model (1.3.1) implies that

$$\overline{F}(x) = \overline{F}(\tfrac{x}{k})\overline{G}(\tfrac{x}{k}), \tag{1.3.2}$$

where $\overline{F}(x) = 1 - F(x)$ and $\overline{G}(x) = 1 - G(x)$. Arnold and Hallett (1989) proved that if the survival function of X_0 is chosen as $\overline{F}_0(x) = \prod_{j=1}^{\infty} \overline{G}\left(\frac{x}{k^j}\right)$ with $\overline{F}(0) = 1$, then $\{X_n, n\geq 0\}$ defined by (1.3.1) is a stationary sequence of r.v.'s having each X_n distributed as F. In this case the infinite product does not diverge to zero. Further the model has a solution if and only if

$$\overline{G}(x) = \frac{\overline{F}(kx)}{\overline{F}(x)} \tag{1.3.3}$$

defines a proper survival function. Lewis and Mckenzie (1991) discuss the existence of the solutions in different cases of this model in terms of survival functions and hazard rates.

The model (1.3.1) has most of the properties of a first order autoregressive model having "minimum" instead "addition" and hence the name autoregressisive minification model. This is useful in modelling a situation where the underlying distribution have closed survival function. For example, if one wants to model the stream flow of rivers, where during certain periods there will not be flow and when there is flow there will be lot of variations, which are very common in hydroloyical and geophysical sciences.

Even though weibull or extreme value r.v.'s are commonly used for modelling the data of the above type, the sequences of such r.v.s cannot be generated with linear random coefficient models of the type (1.2.2). Same is the problem with Pareto distribution, though it is very useful in modelling variety of socio-economic variables. In modelling Markov dependent time series with these marginal distributions it is found that the minification model is more appropriate. Studying the probabilistic properties of these models becomes easier here as the distributions have closed form expressions for their survival functions. Minification models for different special distributions are studied by various authors. For example Tavares (1980) defined the model for exponential variates, Sim (1986) defined for weibull, Yeh et al. (1988) discuss model for Pareto variables and so on.

1.4. Statistical Inference for Markov Sequences.

Statistical Inference is an integral part of Stochastic modelling. If we want to check the validity of any stochastic model, it is essential to have good statistical test procedures. This in turn demands estimation of the unknown parameters involved in the model. The classical theory of statistical inference is based on the assumption that the r.v.'s are i.i.d with common d.f. F. But in practice we come across many situations where the data is a realization of a sequence of dependent r.v.'s. To handle such situations the theory of statistical inference for stochastic processes is developed. We are interested in the inference for stochastic models which generate a sequence of r.v.'s having a special kind of dependence structure, defined below known as Markov dependence.

Definition 1.4.1 (Markov Sequence)

A sequence $\{X_n, n \in T\}$ of r.v.'s defined on a common probability space (Ω, F, P) is said to be Markovian if

$$\Pr[X_n \in B | X_{n-1} = x, X_{n-2} = x_{n-2}, \ldots] = \Pr[X_n \in B | X_{n-1} = x]$$

for any borel set B and $x \in \Omega$, where $\Pr[X_n \in B | X_{n-1} = x]$ is called the transition function of $\{X_n, n \in T\}$. In particular, $\Pr[X_n \leq y | X_{n-1} = x]$ is referred to as the transition distribution of X_n at y given $X_{n-1} = x$.

Definition 1.4.1 (Stationary Stochastic Sequence)

A sequence $\{X_n, n \in T\}$ is said to be stationary if for any positive integer k and $t_1, t_2, \ldots, t_n$ and h in T the joint distribution of $X(t_1), X(t_2), \ldots, X(t_k)$ is same as that of $X(t_1+h), X(t_2+h), \ldots, X(t_k+h)$.

Definition 1.4.3 (Marginal Stationary Sequence)

A sequence $\{X_n, n \in T\}$ of r.v.'s is said to marginally stationary if X_n's are identically distributed for every n.

The assumption on the models (1.2.2) and (1.3.1) immediately imply that the sequence $\{X_n\}$ generated by them are in fact Markovian. Moreover these sequences are stationary under some mild additional conditions. We will come back to these properties again in the forthcoming chapters.

The importance of Markov sequence in the analysis of practical situations, necessitated the development of related theory of statistical inference. One of the useful references for this subject is the book by Billingsley (1961). The statistical inference for stochastic processes in general and for many special models are discussed in Basawa and Prakasa Rao (1980).

1.5. Sequential Estimation

The statistical inference in classical setup is based on a random sample (X_1, X_2, ...,X_n) of size n where n is a fixed positive integer. In stochastic processes the inference is made by observing the realization for a duration of fixed length say 'l'. In both these cases, it is assumed that the sample size n and the duration l do not depend on the observations. That is, one does not take advantage of the information supplied by the observations for choosing the sample size. Choice of an optimum sample size is a crucial problem while planning any statistical experiment. In most of the experiment, sampling is very expensive and taking of each observation involves some cost. Since, cost of sampling is a concern, one has to find minimum size of the sample required to make an optimum decision. Now the problem becomes that of finding a value of n (sample size) which optimizes (minimizes or maximizes) an appropriate objective function. One of the procedures used in such cases is the sequential method.

Sequential procedure is a method of statistical inference whose characteristic feature is that the number of observations or the time required for observation of the process not determined in advance. The decision to terminate the observation on the process depends, at each stage, on the result of the observations previously made. A merit of this method is that test procedures and estimators can be derived in some smaller number of observations.

Another context, where sequential estimation becomes necessary is that when we want to determine the optimum sample size to estimate a parameter under certain optimality criteria in the presence of unknown nuisance parameters. This is illustrated in the example of Woodroof (1982) pp. 105.

Ghosh and Sen (1991) describe some situations where fixed sample size procedure is not suitable. The following are some of the specific examples discussed in Ghosh and Sen (1991).

a. Example where Sequential Analysis is intrinsic

Consider the situation when the blood pressure X_n of a patient under intensive care is monitored continuously in time n. The problem here may be how to analysis and interprect sudden fluctuations in pressure. To some extend the same is true in the classical secretary problem. Here one is dealing with k objects which are intrinsically ranked 1(best), ..., k (worst) according to some characteristics, but the observer can rank them only by visual comparison with each other. The observer assigns a rank X_n to the n^{th} arrival by comparing it with its $(n-1)^{th}$ predecessors who were all rejected. The problem is to design a stopping rule that maximise the probability of selecting the best one when the observations do not have access to the rejected ones. Clearly in both examples, a fixed sample analysis cannot be conceived.

b. Example where only Sequential Procedure yields solutions.

There are problems in point estimation, confidence intervals and hypothesis testing where fixed sample procedures can be conceptualized but cannot provide solutions. Suppose that the observations are i.i.d Bernoulli variables with $P(X_1 = 0) = p = 1 - P(X = 1)$ and one wants an unbiased estimate for p^{-1}. Such an estimate does not exist if one consideres the fixed data sample $(X_1, X_2, \ldots, X_n)$ for any $n \geq 1$. On the other hand the stopping rule with N = smallest n for which $X_n = 1$ yields N itself as an unbiased estimate of p^{-1}. This stopping rule is known as Haldane's inverse sampling procedure.

c. Example where Sequential Analysis is ethical.

Consider a clinical trial or reliability study designed to elucidate the differential if any between two competing treatments. The response data from the patients in the sample are recorded in the order in which they appear. As soon as one treatment could be judged superior to the other, ethical considerations demand curtailment of the study.

1.6 Some Useful Definitions and Results.

In this section we quote some useful definitions and results which are frequently used in our discussion. Proofs of these results may be found in the reference cited in the parenthesis.

Definition 1.6.1. (Stochastically bounded random variables).

A sequence of random variables is said to be stochastically bounded if for $\varepsilon > 0$ there is a $C > 0$ for which

$$P\{|Y_n| > C\} < \varepsilon \qquad \text{for all } n > 1.$$

In particular if $\{Y_n\}$ converges in distribution then $\{Y_n, n \geq 1\}$ is stochastically bounded.

Definition 1.6.2. (Ergodic Sequence).

A stationary process is said to be ergodic if $\Pr\{(X_0, X_1, \ldots) \in A\}$ is either zero or one whenever A is a shift invariant event.

Remark.1.6.1. (Karlin and Taylor (1974), pp. 488).

If sequence $\{X_n\}$ is stationary and ergodic then the sequence

$Y_n = \phi(X_n, X_{n-1},\ldots),\ n = 1,2,\ldots$

generates another ergodic stationary sequence.

Definition 1.6.3. (Uniformly Continuous in Probability (u.c.i.p)).

A sequence $\{Y_n\}$ of r.v's is said to be uniformly continuous in probability if for every ε>0 there is a δ>0 for which $P\left\{\underset{0\le k\le n\delta}{Max}\left|Y_{n+k}-Y_n\right|\ge\varepsilon\right\}<\varepsilon$ for all $n\geq 1$.

Remark.1.6.2. (Woodroofe (1982), pp.41).

If $\{Y_n,\ n\geq 1\}$ converges to a finite limit with probability one then $\{Y_n\}$ is u.c.i.p.

Remark.1.6.3. (Woodroofe (1982), pp.41).

If $\{Y_n\}$ and $\{Z_n\}$ are u.c.i.p then so is $\{Y_n+Z_n,\ n\geq 1\}$, if in addition $\{Y_n,\ n\geq 1\}$ and $\{Z_n,\ n\geq 1\}$ are stochastically bounded and if ϕ is any continuous function on R^2 then $\phi(Y_n, Z_n)$ is u.c.i.p.

Definition 1.6.4 (Uniform integrability (u.i)).

A sequence of r.v.'s $\{X_n\}$ is said to be uniformly integrable if

$$\lim_{\alpha\to\infty}\sup_{n\geq 1}\int_{\{|X_n|\geq\alpha\}}|X_n|dP = 0.$$

The uniform integrability gives a sufficient condition to interchange the limit and expectation of a sequence of r.v.s.

Definition 1.6.5 (Uniformly Mixing Sequences).

A Sequence $\{X_n\}$ of r.v.'s is said to be uniformly mixing if

$$|P(A\cap B)-P(A).P(B)| \leq P(A).\phi(h),$$

where $A \in \sigma\{X_0, X_1, \ldots, X_n\}$, $B \in \sigma\{X_{n+h}, X_{n+h+1}, \ldots\}$ and $\phi(h) \to 0$ ans $h \to \infty$.

Definition 1.6.6 (*m*-dependent r.v.s).

A Sequence $\{X_n\}$ of r.v.'s is said to be m-dependent if ($X_1, \ldots, X_k$) and ($X_{n+k}, X_{n+k+1}, \ldots$) are independent for any k when ever $n>m$.

Definition 1.6.7 (Martingale Sequences).

The $\{X_t,\ t\in T\}$ is said to constitute a martingale w.r.t a non-decreasing sequence of σ-field $\{\boldsymbol{D}_t\ t\in T\}$ if the following conditions hold

i For every $t\in T$, X_t is $\boldsymbol{D}_t$ measurable

ii $E[|X_t|]<\infty$ for every $t\in T$

iii For $s,t\in T$, $s<t$, the relation $E[X_t|\boldsymbol{D}_s] = X_s$ a.s.

Definition 1.6.8 (Submartingale Sequences).

The $\{X_t,\ t\in T\}$ is said to constitute a submartingale w.r.t a non-decreasing sequence of σ-field $\{\boldsymbol{D}_t\ t\in T\}$ if the following conditions hold

i. For every $t \in T$, X_t is $\boldsymbol{D}_t$ measurable

ii. $E[|X_t|] < \infty$ for every $t \in T$

iii. For $s, t \in T$, $s < t$, the relation $E[X_t|\boldsymbol{D}_s] \geq X_s$ a.s.

Remark 1.6.4 (c.f: Karlin and Taylor (1974), pp.250).

If $\{X_t,\ t \in T\}$ is a martingale and if g is a convex function on R then $\{g(X_t)\}$ is a submartingale provided $E[|g(X_t)|] < \infty$ for $t > 1$.

Definition 1.6.9 (Reverse (Backward) Martingales).

The $\{Z_t,\ t \in T\}$ is said to constitute a reverse martingale w.r.t a decreasing sequence of σ-field $\{\boldsymbol{G}_t\ t \in T\}$ if the following conditions hold

i. For every $t \in T$, Z_t is $\boldsymbol{G}_t$ measurable

ii. $E[|Z_t|] < \infty$ for every $t \in T$

iii. $E[Z_t|\boldsymbol{G}_{t-1}] = Z_{t-1}$ a.s.

Result 1.6.1 Minkowski Inequality [Chow and Teicher (1978), pp.108].

For $p \geq 1$, $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$

where $\|.\|_p$ denotes the pth norm defined as $\|X\|_p = E^{1/p}|X|^p$

Result 1.6.2 Holder's Inequality [Chow and Teicher (1978), pp.104].

$$E|XY| \leq E^{1/p}|X|^p\, E^{1/q}|Y|^q \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1.$$

Holder's inequality with $p = q = 2$ is called **Schwartz Inequality.**

Result 1.6.3 Markov Inequality [Chow and Teicher (1978), pp.85]

$$P(|X| > a) \le \frac{E|X|^r}{a^r}, \quad a>0,\ r\ 0.$$

Result 1.6.4 Liapounov Inequality [Chow and Teicher (1978), pp.104]

$$\|X\|_r \le \|X\|_p, \text{ for } r<p.$$

Result 1.6.5. Anscombe's Theorem [Woodroofe (1982), pp.11].

If $Y_1, Y_2, \ldots$ are u.c.i.p and t_a $a>0$ be an integer valued r.v. for which t_a/a converges to a finite positive constant c in probability and $N_a = [ac]$ where [x] denotes the greatest integer part of x. Then $Y_{t_a} - Y_{N_a} \to 0$ in probability as $a \to \infty$. If in addition Y_n converges in distribution to a r.v. Y then $Y_{t_a} \to Y$ as $a \to \infty$.

Result 1.6.6. Slutsky's Theorem (Chow and Teicher (1978), pp.249)

If $\{X_n\}$, $\{Y_n\}$ and $\{Z_n\}$ are three sequences of r.v.'s with $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} a$ and $Z_n \xrightarrow{p} b$, where a, b are finite constants, then

$$X_n Y_n + Z_n \xrightarrow{d} aX + b,$$

where $\xrightarrow{d}$ and $\xrightarrow{p}$ denote the convergence in distribution and probability respectively.

Result 1.6.7 Martingale Central limit Theorem (Nicholls and Quinn (1982), pp.14).

Let $\{\zeta_t\}$ be a sequence of r.v.'s with the property that ζ_t may be expressed as a function not depending on t, which is measurable w.r.t. σ-field $\boldsymbol{F}_t$ generated by a sequence $\{\alpha_t, \alpha_{t+1}, \ldots\}$ of strictly stationary ergodic r.v. Further more suppose that

$E(\zeta_t | F_t)=0$ and $E(\zeta_t^2) = c^2 < \infty$. Then $(c^2 N)^{-1/2} \sum_{t=1}^{N} \zeta_t$ converges in distribution to a standard normal variate.

Result 1.6.8 Maximal Inequality for submartingales (Karlin and Taylor (1974), pp. 251).

Let $\{X_n\}$ be a submartingale for which $X_n \geq 0$ for all n. Then for any positve λ

$$\lambda \Pr\left\{\underset{0 \leq k \leq n}{Max}\, X_k > \lambda\right\} \leq E[X_n].$$

Result 1.6.9 Maximal Inequality for reverse submartingales (Sen (1982), pp. 13)

Let $\{X_n\}$ be a reverse submartingale. Then

$$\left\| \underset{n_0 \leq n \leq n_1}{Max}\, X_n \right\|_p \leq \frac{p}{p-1} \left\| X_{n_0} \right\|_p, \text{ for } p>1.$$

Result 1.6.10 Marcinkiewicz-Zygmund Inequality (Chow and Teicher (1978), pp.356).

If $\{X_n, n \geq 1\}$ is an i.i.d. sequence with $EX_1 = 0$, $E|X_1|^p < \infty$, $p \geq 2$ and $S_n \sum_{i=1}^{n} X_i$ then $E|S_n|^p = O(n^{p/2})$, where the notation $f = O(g)$ means that $\frac{f}{g}$ is bounded.

Result 1.6.11 Burkholder Inequality [Chow and Teicher (1978), pp.384].

If $f = \{f_n, n \geq 1\}$ is a martingale and $p \in (1,\infty)$, then there exist constants $A_p = 18p^{3/2}(p-1)^{-1}$ and $B_p = 18p^{3/2}(p-1)^{-1/2}$

such that $A_p \|S_n(f)\|_p \leq \|f_n\|_p \leq B_p \|S_n(f)\|_p$ where $S_n(f) = \left(\sum_{j=1}^{n} d_j^2\right)^{1/2}$, $d_n = f_n - f_{n-1}$.

Result 1.6.12 [Sriram (1987)].

Let Y_n and Z_n be any sequence of r.v.s and $a, b \neq 0$ and $s>0$ be real numbers. If $P[\,|Y_n - a| > \varepsilon] = O(n^{-s}) = P[\,|Z_n - b| > \varepsilon]$ for every $\varepsilon > 0$, then

i. $P[\,|Y_n Z_n - ab| > \varepsilon] = O(n^{-s})$

ii. $P\left[\left|\frac{Y_n}{Z_n} - \frac{a}{b}\right| > \varepsilon\right] = O(n^{-s})$.

Proof:

Consider

$$P[\,|Y_n Z_n - ab| > \varepsilon] = P[\,|Y_n Z_n - Y_n b + Y_n b - ab| > \varepsilon]$$

$$\leq P[\,|Y_n (Z_n - b)| > \varepsilon/2\,] + P[\,|b(Y_n - a)| > \varepsilon/2]. \qquad (1.6.1)$$

But $[\,|Y_n (Z_n - b)| > \varepsilon/2\,] = [\,|Y_n (Z_n - b)| > \varepsilon/2,\ |Z_n - b| \geq \delta]$

$$\cup\, [\,|Y_n (Z_n - b)| > \varepsilon/2,\ |Z_n - b| < \delta]$$

$$\subseteq [\,|Z_n - b| > \delta] \cup [\,|Y_n| \geq \varepsilon/2\delta]$$

Hence $P[\,|Y_n (Z_n - b)| > \varepsilon/2\,] \leq P[\,|Z_n - b| > \delta] + P[\,|Y_n| \geq \varepsilon/2\delta]$

$$= O(n^{-s}) + O(n^{-s}).$$

$$= O(n^{-s}) \qquad (1.6.2)$$

Consider the second term in (1.6.1)

$$P[\,|b(Y_n - a)| > \varepsilon/2] = P[\,|Y_n - a| > \varepsilon/2b]$$

$$= O(n^{-s}). \qquad (1.6.3)$$

Application of (1.6.2) and (1.6.3) in (1.6.1) gives part (i). Proof of part (ii) is similar and hence omitted.

Result 1.6.13 Central limit Theorem for *m*-dependent random variables (Ibragimov and Linnik (1971), pp.370).

If $\{X_n\}$ be is a stationary and *m*-dependent sequence with $E[X_n] = 0$ and $E[X_n^2]<\infty$. Then $\sigma^2 = E[X_n^2] + 2\sum_{k=1}^{m} E[X_1 X_{1+k}]$ converges, and if $\sigma \neq 0$ $\sqrt{n}\sigma^{-1}\bar{X}_n \xrightarrow{d} N(0, \sigma^2)$.

Where $\xrightarrow{d}$ means convergence in distribution.

Result 1.6.14 Central limit Theorem for Uniform Mixing Sequence. (Billingsley, (1968), pp.174).

Suppose that $\{X_n\}$ is stationary ϕ-mixing sequence with $\sum_n \phi_n^{1/2} < \infty$ and that X_0 has mean zero and finite variance. Then the series $\sqrt{n}\,\bar{X}_n \xrightarrow{d} N(0, \sigma^2)$, where

$$\sigma^2 = E[X_0^2] + 2\sum_{k=1}^{\infty} E[X_0 X_k].$$

Result 1.6.15 Ergodic Theorem [Karlin and Taylor (1974), pp.487]

Let $\{X_n\}$ be an ergodic stationary process having a finite mean *m* . Then, with probability one

$$\lim_{n\to\infty} \frac{1}{n}(X_1 + X_2 + \dots + X_n) = m.$$

1.7 Summary of the Thesis

The rest of this thesis is divided into five chapters. The chapter 2 discusses the estimation problem for the minification model defined by (1.3.1.). In this chapter we study some of the probabilistic properties of the minification model such as ergodicity and uniform mixing. Based on these properties we study the performance of the estimators for the common mean and k. These are followed by estimation of the parameters in some special cases like, the minification models generating Exponential, Uniform and Pareto marginals.

The chapter 3 deals with the sequential point and interval estimation of the parameters of the above models. The optimal properties of sequential point and interval estimators are studied here.

The different properties of RCAR(1) model (1.2.1) and least squares estimators for the parameters of the this model form the subject matter of chapter 4. Least squares estimators suggested by Nicholls and Quinn (1982) are considered and their p^{th} moment convergence are studied in detail in this chapter.

Chapter 5 provides the sequential methods of estimation for RCAR(1) model. Results from chapter 4 are used to prove the optimal properties of sequential procedure for estimating location parameter of RCAR(1) model.

Chapter 6 is devoted to the sequential estimation of autoregressive parameter of RCAR(1) model. Appropriate stopping rule is developed and first order efficiencies of this stopping rule are established. It also includes a discussion on sequential interval estimation for autoregressive parameter.

The references used in the thesis at various stages are listed after the chapter 6.

Each chapter is divided into different sections. The equations are numbered as (a.b.c.). This means that equation number 'c' of section 'b' in chapter 'a'. Similarly the Theorems, Lemmas, Results and Definitions are also numbered. The references are arranged in the alphabetical order of authors.

CHAPTER 2

PARAMETER ESTIMATION IN MINIFICATION PROCESSES

2.1 Introduction

The problem of estimation is an important stage in stochastic modelling. In this chapter we estimate the parameters of minification model described in Section 1.3. This is one of the non-linear models used to generate the non-Gaussian time series.

In recent years it is found that the models of this type have many applications in analysing the real life situations. Further it is also found that these non-linear models are more suitable than the linear Gaussian models in certain situations, see for example, Tong (1990), Lawrance (1991) and references cited their in. One of the important nonlinear models used to generate a sequence $\{X_n\}$ of a non-negative random variable is defined by

$$X_n = \begin{cases} X_0 & n = 0 \\ kMin(X_{n-1}, Z_n) & k > 1,\ n = 1,2,\ldots \end{cases} \tag{2.1.1}$$

where $\{Z_n\}$ is a sequence of iid non-negative, nondegenerate r.v.s called innovations and X_0 is independent of Z_1. This model is referred to as a minification model. Various aspects of this model are discussed in Lewis and Mckenzie (1991).

Now we will consider some interesting properties of a minification model defined by (2.1.1).

The sequence has Markovian property.

Let $\{X_n\}$ be defined by (2.1.1) and consider the conditional probability distribution function

$$P[X_n \le x \mid X_{n-1} = y, X_{n-2} = x_{n-2}, X_{n-3} = x_{n-3}, \ldots, X_0 = x_0]$$

$$= 1 - P[X_n > x \mid X_{n-1} = y, X_{n-2} = x_{n-2}, X_{n-3} = x_{n-3}, \ldots, X_0 = x_0]$$

$$= 1 - P[k\,Min(X_{n-1}, Z_n) > x \mid X_{n-1} = y, X_{n-2} = x_{n-2}, \ldots, X_0 = x_0]$$

$$1 - P[X_{n-1} > x/k \mid X_{n-1} = y, X_{n-2} = x_{n-2}, \ldots, X_0 = x_0]\, P[Z_n > x/k]$$

$$= \begin{cases} 1 & \text{if } y \le x/k \\ 1 - P[Z_n > x/k] & \text{if } y > x/k. \end{cases}$$

This is same as the conditional probability distribution function $P[X_n \le x \mid X_{n-1} = y]$. Hence the process defined by (2.1.1) is Markovian.

The different aspects of the model (2.1.1) when X_n has a specified distribution are studied by various researchers. For example, Tavares (1980) discussed the minification process with exponential marginals, Sim (1986) defined this model for Weibull r.v.s., Yeh, Arnold and Robertson (1988) for Pareto r.v.s and Pillai (1991) studied a model with semi-Pareto marginals.

The distributional properties of minfication model in general set up are studied by Lewis and Mckenzie (1991) and Arnold and Hallett (1989). The applications of these models in various areas such as geophysical sciences, reliability etc. are discussed in the above mentioned references.

In model (2.1.1) let $F(x) = P(X_0 \leq x)$ and $G(y) = P(Z_1 \leq y)$. Lewis and Mckenzie (1991) have proved that the model (2.1.1) defines a stationary sequence $\{X_n\}$ if and only if

$$\overline{G}(x) = \frac{\overline{F}(kx)}{\overline{F}(x)}, \qquad x \geq 0, k > 1 \tag{2.1.2}$$

where $\overline{F}(x) = 1 - F(x)$ and $\overline{G}(x) = 1 - G(x)$. Arnold and Hallett (1989) showed that if the distribution of X_0 is chosen as

$$\overline{F}(x) = \prod_{i=1}^{\infty} \overline{G}\left(x / k^i\right) \tag{2.1.3}$$

then (2.1.1) defines a stationary sequence with X_n having the survival function (2.1.3) for every $n \geq 0$. In (2.1.3) it is assumed that the product does not diverges to zero.

For our convenience we present the condition (2.1.3) in terms of a sequence of iid non-negative r.v.s. Let $\{Z_{-n}, n=0,1,2,\ldots\}$be a sequence of iid non-negative r.v.s with common survival function $\overline{G}(.)$. Let us define

$$X_0 = \inf_{0 \leq j < \infty} k^j Z_{-j}. \tag{2.1.4}$$

Now it follows that the survival function X_0 is given by (2.1.3). We state some of the useful results related to the model (2.1.1) below. The proofs of theses results may be found in Lewis and Mckenzie (1991).

Result 2.1.1 Let $\{X_n, n \geq 0\}$ be a stationary Markov sequence defined by (2.1.1) with stationary density function $f(.)$. Then

$$\phi(j) = P[X_n = k^j X_{n-j}] = \int_0^\infty \frac{\overline{F}(k^j x)}{\overline{F}(x)} f(x) dx. \qquad (2.1.5)$$

Result 2.2.2 For the stationary sequence defined by (2.1.1) the covariance between X_n and X_{n-1} is given by

$$\mathrm{Cov}(X_n, X_{n-1}) = kE\left\{(X - m_x)\int_0^x \overline{G}(z)dz\right\} \qquad (2.1.6)$$

where $m_x = E(X) = E\left\{k\int_0^x \overline{G}(z)dz\right\}$.

Remark 2.1.1 Correlation between X_n and X_{n-1} denoted by $\rho_x(1)$ can be obtained from (2.1.6) by dividing by $\mathrm{Var}(X_n)$. If we denote $\rho_x(1) = c(k)$, then $\rho_x(j) = \mathrm{Corr}(X_n, X_{n-j})$ may be obtained by replacing k in $c(k)$ by k^j. That is $\rho_x(j) = c(k^j)$.

Remark 2.1.2 Autocorrelation function is said to be in geometric form if and only if $\rho_x(1) = \rho^\alpha$ for some $\alpha > 0$. Exponential, Uniform and Pareto minification process have this geometric auto correlation function. Also for these minification processes

$$\phi(j) = P[X_n = k^j X_{n-j}] = \mathrm{Corr}(X_n, X_{n-j}) = \rho_x(j). \qquad (2.1.7)$$

The relation $\phi(j) = \rho_x(j)$ doesnot hold in gneral. Lewis and Mckenzie (1991) showed that the relation $\phi(j) = \rho_x(j)$ is not true in case of Weibull minification processes. The quantity $\phi(j)$ defined in (2.1.5) is also a useful measure of dependence.

As far as statistical inference is concerned, little work is being done for these models. Adke and Balakrishna (1992) have estimated the parameters of exponential minification model. In this paper they proposed some sampling schemes to determine the exact value of k and then estimate the mean of $\{X_n\}$. Balakrishna (1998) discussed the estimation problem in semi-Pareto and Pareto processes. In this chapter, we estimate the common mean of $\{X_n\}$ and the parameter k of the general minification process defined by (2.1.1).

In Section 2.2, we prove that a stationary minification process is ergodic and uniformly mixing. These result are used to prove the optimal properties of estimators of common mean of X_n, in Section 2.3. Section 2.4 deals with estimation of k. Section 2.5 considers the estimation problems in some special cases.

2.2 Some Probabilistic Properties of the Model.

In this section we prove that the minification process is ergodic and uniformly mixing [See definitions 1.6.2 and 1.6.5].

Lemma 2.2.1 Let $\{X_n\}$ be a stationary Markov sequence defined by (2.1.1) with $k>1$ and the distribution of X_0 specified by (2.1.3). Then $\{X_n\}$ is ergodic.

Proof: Let $F_n = \sigma\{X_1, X_2, \ldots, X_n\}$, $G_n = \sigma\{X_0, Z_1, Z_2, \ldots, Z_n\}$ be the σ-field induced by $(X_1, X_2, \ldots, X_n\}$and $(X_0, Z_1, Z_2, \ldots, Z_n)$ respectively. Repeatedly using (2.1.1) we can write

$$X_n = Min\{k^n X_0, k^n Z_1, k^{n-1} Z_2, \ldots, kZ_n\}. \tag{2.2.1}$$

At this stage if we use the representation (2.1.4) of X_0 then

$$X_n = \inf_{j\leq 0}\{k^{-j+1} Z_{n-j}\} \tag{2.2.2}$$

The representations (2.2.1) and (2.2.2) imply that F_n is contained in G_n which is the minimal sigma field induced by a sequence of i.i.d r.v.s $\{Z_n\}$. Hence the tail sigma field τ of $\{X_n\}$is contained in the tail sigma field τ^* of the independent r.v.s. $\{Z_n\}$. It is well known by Kolmogorov zero-one law that each event of τ^* has probability zero or one. This implies that τ contains only events of probability zero or one, which is a sufficient condition for $\{X_n\}$ to be ergodic.(cf. Stout (1974), pp.182). Hence the lemma is proved.

Lemma 2.2.2 The minification sequence $\{X_n\}$ generated by (2.1.1) is uniformly mixing with mixing parameters

$$\phi(h) = P[X_n = k^h X_0], \quad h = 0,1,2. \tag{2.2.3}$$

Proof: Let A and B be two events such that

$$A \in \sigma\{X_0 X_1 \ldots X_n\} \text{ and } B \in \sigma\{X_{n+h}, X_{n+h+1} \ldots\}.$$

In order to prove that $\{X_n\}$ is uniformly mixing we have to show that (See Definition 1.6.5)

$$|P(A \cap B) - P(A)P(B)| \le \phi(h)P(A) \tag{2.2.4}$$

such that $\phi(h) \to 0$ as $h \to \infty$. We can prove this by closely inspecting the r.v.s, $X_{n+1}, X_{n+2}, \ldots, X_{n+h-1}$. By definition of the model (2.1.1)

$$X_{n+j} = k \operatorname{Min}(X_{n+j-1}, Z_{n+j}), \qquad j = 1,2,\ldots,h-1$$

$$= \begin{cases} kX_{n+j-1} & \text{if } X_{n+j-1} \le Z_{n+j} \\ \\ kZ_{n+j} & \text{if } X_{n+j-1} > Z_{n+j}. \end{cases} \tag{2.2.5}$$

Note that Z_{n+j} is independent of $X_{n+j-1}, X_{n+j-2}, \ldots$. If $X_{n+j} = k\, Z_{n+j}$ for some $j=1,2,\ldots,h-1$, then the events A and B will be independent and hence (2.2.4) will hold. Let N be the number of innovations occurring in the interval $(n+1, n+h-1]$. Then

$$P(A \cap B) = P(A)\, P(B) \qquad \text{if } N > 0$$

$$\ne P(A)\, P(B) \qquad \text{if } N = 0$$

That is

$$P(A \cap B | N > 0) = P(A \mid N>0)\, P(B \mid N>0). \tag{2.2.6}$$

From (2.2.5) it follows that $N=0$ if and only if $X_{n+h} = k^h X_n$ and in this case A and B are not independent.

Consider

$$P[A \cap (N>0)] = \int_0^\infty P\{[A \cap (X_{n+h} \neq k^h X_n)] | X_n = x\}\ dF(x)$$

$$= \int_0^\infty P\{A | X_n = x) P\{X_{n+h} \neq k^h X_n | X_n = x\}\ dF(x)$$

$$= \int_0^\infty P\{A | X_n = x) P\{N > 0)\ dF(x)$$

$$= P(N>0)P(A).$$

At the first stage of simplification we have used the Markov property of $\{X_n\}$ and then we made use of the fact that $[N=0]$ if and only if $[X_{n+h} = k^h X_n]$. Hence from (2.2.6)

$$P(A \cap B | N > 0) = P(A)\, P[B \mid N>0]. \tag{2.2.7}$$

Once again using Markov property of $\{X_n\}$ and (2.2.7) we can write

$$P(A \cap B) = P[A \cap B \cap (N=0)] + P[A \cap B \cap (N>0)]$$

$$= P[A \cap B \cap (N=0)] + P[A \cap B \mid (N>0)]P(N>0)$$

$$\leq P(N=0)P(A) + P(A)P(N>0)\, P[B \mid N>0]. \tag{2.2.8}$$

Hence using (2.2.8) and the fact that $P(B) = P[B \cap (N=0)] + P[B \cap (N>0)]$ we get

$$|P(A \cap B) - P(A)P(B)| \leq P(A)[P(N=0) + P(N>0)\, P(B \mid N>0)$$

$$- P(B \mid N>0)\, P(N>0) - P(B \mid N=0)\, P(N=0)]$$

$$\leq P(A)P(N=0)\, [1 - P(B \mid N=0)]$$

$$\leq P(A)\, \phi(h),$$

where $\phi(h) = P[N=0] = P[X_{n+h} = k^h X_n]$.

Since $\{X_n\}$ defined by (2.1.1) is stationary, from (2.1.5) we have

$$\phi(h) = \int_0^\infty \frac{\overline{F}(k^h x)}{\overline{F}(x)} dF(x).$$

Note that $\overline{F}(k^h x) / \overline{F}(x)$ is a decreasing function of h and hence we can write

$$\lim_{h\to\infty} \phi(h) = \int_0^\infty \lim_{h\to\infty} \frac{\overline{F}(k^h x)}{\overline{F}(x)} dF(x) = 0.$$

This completes the proof.

2.3 Estimation of the Mean

Let $\{X_n\}$ be a stationary sequence defined by (2.1.1) with common d.f. $F(.)$ and common mean $\mu = E(X_n)$. Assume further that $\text{Var}(X_n) = \sigma^2 < \infty$ for all n. The ergodicity of $\{X_n\}$ implies that the sample mean $\overline{X}_n = (X_1 + X_2 + \ldots + X_n)/n$ is a natural estimator of μ. The asymptotic properties of $\overline{X}_n$ are discussed in the following theorem.

Theorem 2.3.1 The time average $\overline{X}_n$ is strongly consistent and asymptotically normal (CAN) estimator of μ. The asymptotic variance (A.V) of $\overline{X}_n$ is given by

$$\text{A.V}\,(\overline{X}_n) = \frac{\sigma^2}{n} B(k) \tag{2.3.1}$$

where $B(.)$ is a continuous non-negative function.

Proof: By Lemma 2.2.1 and point wise ergodic theorem [See Result 1.6.16] it follows that $\overline{X}_n \to \mu$ almost surely (a.s) as $n \to \infty$. The uniform mixing property of $\{X_n\}$ implies that (cf. Result 1.6.14)

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} Z_1 \tag{2.3.2}$$

where $\xrightarrow{d}$ stands for convergence in distribution and Z_l is a normal r.v with mean zero and variance

$$\sigma_1^2 = \sigma^2 + 2\sigma^2 \sum_{j=1}^{\infty} \rho_x(j). \tag{2.3.3}$$

In this case $\rho(j)$ is the autocorrelation between X_l and X_{l+j}. Further $0< \sigma_1^2 <\infty$. Thus $\overline{X}_n$ is the CAN estimator of μ and

$$\text{A.V}(\overline{X}_n) = \frac{\sigma^2}{n}\{1 + 2\sum_{j=1}^{\infty} \rho_x(j)\}. \tag{2.3.4}$$

Let us denote by $\rho_x(1) = \text{Corr}(X_1, X_2)$ and assume that $\rho(1)$ is a continuous function of k say $c(k)$ [See Remark 2.1.1]. Hence we can write

$$\text{A.V}(\overline{X}_n) = \frac{\sigma^2}{n} B(k)$$

where

$$B(k) = 1 + 2\sum_{j=1}^{\infty} c(k^j), \tag{2.3.5}$$

which is continuous in k. Hence the proof is complete.

Remark 2.3.1 The uniform mixing property of $\{X_n\}$ implies that the summations in (2.3.4) and (2.3.5) are finite [cf. Ibragimov and Linnik (1978), pp.344].

Based on the above CAN property of $\overline{X}_n$ we can construct the asymptotic confidence interval of μ as follows.

Here we are interested in specifying an interval which covers the true parameter (Population mean μ) with an assigned probability say (1-α). This particular interval is known as confidence interval with confidence coefficient (1-α).

By the above Theorem 2.3.1, we have $\dfrac{(\overline{X}_n - \mu)}{\sqrt{\dfrac{\sigma^2}{n}B(k)}} \sim$ N(0,1) asymptoticaly.

Thus for large n, when σ^2 and k are known we can find out an $Z_{\alpha/2}$ from standard normal tables such that

$$P\left[-Z_{\alpha/2} \leq \frac{(\overline{X}_n - \mu)}{\sqrt{\frac{\sigma^2}{n}B(k)}} \leq Z_{\alpha/2}\right] \geq (1\text{-}\alpha)$$

$$P\left[\overline{X}_n - Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}B(k)} \leq \mu \leq \overline{X}_n + Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}B(k)}\right] \geq (1\text{-}\alpha).$$

Thus the 100(1-α)% confidence interval for μ is given by

$$\left[\overline{X}_n \pm Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}B(k)}\right]$$

We will use the above result to study the sequential interval estimation in Section 3.4.

2.4 Estimation of k

In this section we discuss the problem of estimating k. Let us define

$$W_n = \frac{X_n}{X_{n-1}} = \begin{cases} k & \text{if } X_{n-1} \le Z_n \\ \\ k(Z_n / X_{n-1}) & \text{if } X_{n-1} > Z_n . \end{cases} \tag{2.4.1}$$

So that $W_n \le k$ for all n. We propose $\tilde{k}_n$ as an estimator of k where,

$$\tilde{k}_n = \underset{1 \le i \le n}{Max} \; W_i \tag{2.4.2}$$

The properties of $\tilde{k}_n$ are studied below.

Theorem 2.4.1 The estimator $\tilde{k}_n$ is a strongly consistent estimator of k.

Proof: From (2.4.1) and (2.4.2) it is clear that $\tilde{k}_n = k$ if and only if $X_{i-1} < Z_i$ for at least one i, $i=1,2,\ldots$, n. Thus

$$P[\tilde{k}_n \ne k] = P[X_{i-1} > Z_i \text{ for all } i=1,2, \ldots, n].$$

But we have $X_{i-1} = k \operatorname{Min} (X_{i-2}, Z_{i-1})$ and hence

$$P[\tilde{k}_n \ne k] = P[kZ_{i-1} > Z_i,\ kX_{i-2} > Z_i, \text{ for } i=1,\ldots,n]$$

$$P[\tilde{k}_n \ne k] \le P[kZ_{i-1} > Z_i \text{ for all } i=1,2, \ldots, n]$$

$$= P[kZ_0 > Z_1,\ kZ_1 > Z_2,\ \ldots,\ kZ_{n-1} > Z_n]$$

$$\le P[kZ_1 > Z_2, kZ_3 > Z_4, \ldots, kZ_{2\left[\frac{n}{2}\right]-1} > Z_{2n}]$$

$$= \prod_{i=1}^{\left[\frac{n}{2}\right]-1} P[kZ_{2i-1} > Z_{2i}], \tag{2.4.3}$$

where [x] denotes the integer part of x.

We have used the independents of Z_i's to arive at (2.4.3).

This implies that $\sum_{n=1}^{\infty} P[\widetilde{k}_n \neq k] \leq \sum_{n=1}^{\infty} p^{\left[\frac{n}{2}\right]-1} < \infty$, where $p = P[kZ_{2i-1} > Z_{2i}] = P[kZ_1 > Z_2]$ and since Z_i's are iid non-degenerate r.v.s, we have $0 \leq p < 1$.

Now by Borel-Cantelli lemma, it follows that

$$P[\widetilde{k}_n \neq k \text{ infinitely often}] = 0.$$

Equivalently $\widetilde{k}_n = k$ infinitely often with probability one and hence $\widetilde{k}_n \to k$ a.s as $n \to \infty$.

Remark 2.4.1 Note that $0 < W_i = \dfrac{X_i}{X_{i-1}} \leq k$ for all i and $W_i = k$ if and only if $Z_i > X_{i-1}$. Thus the distribution function of W_i is concentrated on a finite interval with a positive jump at the end point. From the study of extreme value theory we know that if $\{V_i\}$ is a sequence of iid non-negative r.v.s concentrated in a finite interval, with a positive jump at the end point then $\underset{1 \leq i \leq n}{Max}\ V_i$ does not converge to a nondegenerate limit distribution for any norming sequence. This result also holds for our stationary sequence since it is uniformily mixing [cf. pp.13 and pp. 60 of Leadbetter, Lingren and Rootzen (1983)]. Hence our estimator $\widetilde{k}_n$ converges to a degenerate distribution for any norming sequence.

In the folowing , we propose an alternative estimators for k in some special cases which are CAN for k.

One of the wellknown minification models is that defined by Travers (1980) for exponential r.v.s. In this case X_0 has the distribution

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \qquad \lambda>0,\ x\geq 0$$

and the iid sequence $\{Z_n\}$ has the common distribution specified by

$$G(x) = P(Z_1 \leq x) = 1 - e^{-\lambda(k-1)x},\ x\geq 0.$$

Then X_n defined by

$$X_n = k\ \mathrm{Min}(X_{n-1}, Z_n),\ \ n = 1,2,\ldots$$

has exponential distribution $F(x)$ for all $n \geq 0$. For this process $\mathrm{E}(X_n) = 1/\lambda$, $\mathrm{Var}(X_n)=1/\lambda^2$, $\mathrm{Corr}(X_n, X_{n+h}) = k^{-h}$ $h=0,1,2\ldots$.

Now we suggest an estimator for k for the exponential minification process defined above.

Let

$$U_j = \begin{cases} 1 & \text{if } X_j \geq X_{j-1} \\ 0 & \text{if } X_j < X_{j-1} \end{cases} \qquad j = 1,2,\ldots . \tag{2.4.4}$$

Then

$$\mathrm{E}(U_j) = k/(2k-1)$$

and (2.4.5)

$$\mathrm{V}(U_j) = k\,(k-1)/(2k-1)^2.$$

Let $\overline{U}_n = n^{-1}\sum_{j=1}^{n} U_j$ be the arithmetic mean of $U_1, U_2, \ldots, U_n$. Now we study the properties of moment estimator of k based on $\overline{U}_n$.

Theorem 2.4.2: For the exponential minification process, the estimator $\hat{k}_n = \frac{\overline{U}_n}{2\overline{U}_n - 1}$ is strongly consistent and $\sqrt{n}(\hat{k}_n - k) \rightarrow Z_2$ as $n \rightarrow \infty$ where Z_2 is a normal r.v with mean zero and variance

$$\sigma_2^2 = k\,(k-1)\,(2k-1)^2 - 2\,(2k-1)^2\,(\,(k-1)^3 \sum_{h=1}^{\infty} \frac{1}{\{k - 1 + k^{h-1}(2k-1)\}}. \qquad (2.4.6)$$

Proof: By the ergodicity of $\{X_n\}$ we have as $n \rightarrow \infty$, $\overline{U}_n = E(U_i) = k/2k-1$ a.s and hence $\hat{k}_n \rightarrow k$ a.s. As U_n is a function of X_n and X_{n-1}, by lemma 2.2.2, it follows that $\{\overline{U}_n\}$ is also stationary and uniform mixing with coefficients

$$\phi^*(h) = \phi(h-1) = k^{-(h-1)}, \qquad h = 1,2, \ldots \qquad (2.4.7)$$

Now by applying Result 1.6.14 we get the result that

$$\sqrt{n}[\overline{U}_n - k/(2k-1)] \xrightarrow{d} Z \qquad (2.4.8)$$

where Z is a normal r.v with mean zero and variance

$$\mathrm{Var}(Z) = \mathrm{Var}(U_1) + 2\sum_{h=1}^{\infty} Cov(U_1, U_{1+h}). \qquad (2.4.9)$$

Now let us compute

$$\begin{aligned} r(h) &= Cov(U_1,\ U_{1+h}) \\ &= E(U_1,\ U_{1+h}) - E(U_1)\,E(U_{1+h}) \\ &= P(U_1 = 1,\ U_{1+h} = 1) - P(U_1 = 1)\,P(U_{1+h} = 1) \end{aligned} \qquad (2.4.10)$$

But

$$\begin{aligned} P(U_1 = 1,\ U_{1+h} = 1) &= P(X_1 > X_0,\ X_{1+h} > X_h) \\ &= P[k\,\mathrm{Min}(X_0, Z_1) > X_0,\ k\,\mathrm{Min}(X_h, Z_{h+1}) > X_h] \end{aligned}$$

$$= P[X_0 > X_0/k,\ Z_1 > X_0/k,\ X_h > X_h/k,\ Z_{h+1} > X_h/k]$$

$$= P[Z_1 > X_0/k,\ Z_{h+1} > X_h/k] = P_h, \text{ (say) since } k>1. \qquad (2.4.11)$$

Now consider P_h

$$P_h = 1 - P[Z_1 > X_0/k,\ Z_{h+1} > X_h/k]^c$$

$$= 1 - P[Z_1 \leq X_0/k] - P[Z_{h+1} \leq X_h/k] + P[Z_1 \leq X_0/k,\ Z_{h+1} \leq X_h/k] \qquad (2.4.12)$$

Denote the last term in (2.4.12) by I_h. Thus

$$I_h = P[Z_1 \leq X_0/k,\ Z_{h+1} \leq X_h/k]$$

$$= P[X_0 \geq kZ_1,\ X_h \geq kZ_{h+1}]$$

$$= P[X_0 \geq kZ_1,\ \mathrm{Min}(k^h X_0, k^h Z_1, k^{h-1} Z_2, \ldots, kZ_h) \geq kZ_{h+1}]$$

$$= P[X_0 \geq kZ_1,\ k^{h-1} X_0 \geq Z_{h+1},\ k^{h-2} Z_1 \geq Z_{h+1}, \ldots, Z_h \geq Z_{h+1}]$$

$$= \int_0^{\infty} P[X_0 \geq kZ_1, X_0 \geq \frac{z_{h+1}}{k^{h-1}}, Z_1 \geq \frac{z_{h+1}}{k^{h-1}}, Z_2 \geq \frac{z_{h+1}}{k^{h-2}}, \ldots, Z_h \geq \frac{z_{h+1}}{k^0}]\ ce^{-cz_{h+1}}\ dz_{h+1}$$

where, $c = (k-1)\lambda$.

Thus

$$I_h = \int_0^{\infty} P[X_0 \geq Max\left(kz_1, \frac{z_{h+1}}{k^{h-1}}\right),\ Z_1 \geq \frac{z_{h+1}}{k^{h-1}}] \exp\left\{-c\frac{z_{h+1}}{k^{h-2}}\right\} \exp\left\{-c\frac{z_{h+1}}{k^{h-3}}\right\} \ldots$$

$$\exp\{-cz_{h+1}\}\ c\exp\{-cz_{h+1}\}\, dz_{h+1}. \qquad (2.4.13)$$

Now consider

$$P[X_0 > Max\left(kz_1, \frac{z_{h+1}}{k^{h-1}}\right),\ Z_1 > \frac{z_{h+1}}{k^{h-1}}]$$

$$= \int_0^{\infty} P[X_0 > Max\left(kz_1, \frac{z_{h+1}}{k^{h-1}}\right),\ Z_1 > \frac{z_{h+1}}{k^{h-1}}]\ ce^{-cZ_1} dz_1$$

$$= \int_0^{\infty} \exp\left\{-\lambda \ \max\left(kz_1, \tfrac{z_{h+1}}{k^{h-1}}\right)\right\} I_{\left(Z_1 > \frac{z_{h+1}}{k^{h-1}}\right)}\ ce^{-cZ_1} dz_1$$

$$= \int_{\frac{z_{h+1}}{k^{h-1}}}^{\infty} e^{-\lambda k Z_1} ce^{-cZ_1} dZ_1 = \text{J (say)}.$$

Since $k>1$ and $Z_1 > \frac{z_{h+1}}{k^{h-1}}$ we have

$$\frac{z_{h+1}}{k^{h-1}} < Z_1 < kz_1.$$

Thus

$$\text{J} = \frac{c}{\lambda k + c} \exp\left\{-(\lambda k + c)\frac{z_{h+1}}{k^{h-1}}\right\}$$

$$= \frac{k-1}{2k-1} \exp\left\{-\lambda(2k-1)\frac{z_{h+1}}{k^{h-1}}\right\}. \qquad (2.4.14)$$

Using (2.4.14) and (2.4.13) I_h can be written as

$$I_h = \int_0^{\infty} \frac{c}{\lambda k + c} \exp\left\{-(\lambda k + c)\frac{z_{h+1}}{k^{h-1}}\right\} \exp\left\{-cz_{h+1}[1 + \tfrac{1}{k} + \tfrac{1}{k^2} + \ldots + \tfrac{1}{k^{h-2}}]\right\} ce^{-cz_{h+1}} dz_{h+1}$$

$$= \int_0^{\infty} \frac{c^2}{\lambda k + c} \exp\left\{-(\lambda k + c)\frac{z_{h+1}}{k^{h-1}}\right\} \exp\left\{-cz_{h+1}[2 + \tfrac{1}{k} + \tfrac{1}{k^2} + \ldots + \tfrac{1}{k^{h-2}}]\right\} dz_{h+1}$$

$$= \frac{\dfrac{c^2}{\lambda k + c}}{\dfrac{\lambda k + c}{k^{h-1}} + c[1 + 1 + \tfrac{1}{k} + \tfrac{1}{k^2} + \ldots + \tfrac{1}{k^{h-2}}]}.$$

Substituting for c we have

$$I_h = \frac{[(k-1)\lambda]^2 / (\lambda k + (k-1)\lambda)}{\dfrac{\lambda k + (k-1)\lambda}{k^{h-1}} + (k-1)\lambda\left\{\dfrac{k^{h-1}-1}{\frac{k^{h-1}(k-1)}{k}} + 1\right\}}$$

$$= \frac{(k-1)^2 / 2k-1}{\frac{2k-1+k^h-k+k^h-k^{h-1}}{k^{h-1}}} = \frac{\frac{(k-1)^2}{2k-1} k^{h-1}}{k-1+2k^h-k^{h-1}}. \qquad (2.4.15)$$

Now let us consider the other terms in (2.4.12)

$$P[Z_1 \le X_0/k] = \int_0^\infty P[x_0 > kz_1] ce^{-cZ_1} dZ_1 = \frac{c}{\lambda k + c} = \frac{k-1}{2k-1}. \qquad (2.4.16)$$

Similarly

$$P[Z_{h-1} \le X_h/k] = \frac{k-1}{2k-1}. \qquad (2.4.17)$$

Using (2.4.15), (2.4.16) and (2.4.17) in (2.4.12) we get

$$P_h = P[U_1 = 1, U_{1+h} = 1] = 1 - \frac{2(k-1)}{2k-1} + \frac{(k-1)^2 k^{h-1}}{(2k-1)(k-1+2k^h-k^{h-1})}$$

$$= \frac{k-1+k^{h+1}}{(2k-1)(k-1+2k^h-k^{h-1})}.$$

Using this value of P_h in (2.4.10),

$$r(h) = \mathrm{Cor}(U_1, U_{1+h}) = P_h - P(X_1 > X_0)\, P(X_{1+h} > X_h)$$

$$= P_h - \frac{k^2}{(2k-1)^2}. \qquad (2.4.18)$$

Where we used the fact

$$P(X_h > X_{h-1}) = P[\, k \,\mathrm{Min}(X_{h-1}, Z_h) > X_{h-1}]$$

$$= P[X_{h-1} > \tfrac{X_{h}}{k}, Z_h > \tfrac{X_{h-1}}{k}] = P[Z_h > \tfrac{X_{h-1}}{k}]$$

$$= \int_0^{\infty} P[Z_h > \tfrac{x}{k}]\lambda e^{-\lambda x} dx$$

$$= \int_0^{\infty} e^{-cx/k} \lambda e^{-\lambda x} dx$$

$$= \frac{\lambda}{\frac{c}{k} + \lambda} = \frac{k}{2k-1},$$

which is independent of n.

Note that

$$P_h = \frac{\frac{k-1}{k^{h-1}} + 1}{\frac{2k-1}{k}\left\{\frac{k-1}{k^h} + 2 - \frac{1}{k}\right\}} \to \frac{1}{\left(\frac{2k-1}{k}\right)\left(2 - \frac{1}{k}\right)}.$$

Thus as $h \to \infty$

$$P_h \to \frac{k^2}{(2k-1)^2}.$$

Now let us simplify $r(h)$ in (2.4.18)

$$r(h) = P_h - \frac{k^2}{(2k-1)^2}$$

$$= \frac{k-1+k^{h+1}}{(2k-1)\{k-1+k^{h-1}(2k-1)\}} - \frac{k^2}{(2k-1)^2}$$

$$= \frac{(1-k)^3}{(2k-1)^2\{k-1+k^{h-1}(2k-1)\}}. \tag{2.4.19}$$

An application of ratio test for convergence of series implies that

$$\sum_{h=1}^{\infty} |r(h)| < \infty$$

Now using (2.4.5) and (2.4.19) in (2.4.9) we get

$$\text{Var}(Z) = \frac{k(k-1)}{(2k-1)^2} - \frac{2(k-1)^3}{(2k-1)^2} \sum_{h=1}^{\infty} \frac{1}{k-1+k^{h-1}(2k-1)} \tag{2.4.20}$$

and 0< Var(Z) <∞.

Let us write

$$\sqrt{n}(\hat{k}_n - k) = -\frac{2k-1}{2\bar{U}_n - 1} \sqrt{n}(\bar{U}_n - \tfrac{k}{2k-1})$$

Since $\frac{2k-1}{2\bar{U}_n - 1} \to (2k-1)^2$ a.s as $n \to \infty$ by Slutsky's theorem we have

$$\sqrt{n}(\hat{k}_n - k) \xrightarrow{d} Z_2$$

where Z_2 is a normal r.v with mean zero and variance σ_2^2 given by (2.4.6). This completes the proof.

In the next section we consider the estimation of parameters in some other special cases.

2.5 Estimation of Parameters in Some Special Cases.

This section is devoted to estimation of parameters when marginal distribution of X_n is either one of Exponential or Uniform or Pareto. In Section 2.1 we have noted the conditions to be satisfied by the marginal distribution so that the non-negative X_n generated by the minification model (2.1.1) is stationary and Markov. The class of such distributions include Exponential, Weibull, Pareto and Uniform- U(0,1) distributions. We summarise the important features of these minification processes in the Table 2.5.1.

Table 2.5.1

Distribution of X_n	Density function f(x)	Mean	Variance	$\overline{F}_X(x)$	$\overline{G}_Z(x)$	$\rho_X(1)$	$\phi(h)$
Uniform	1, $0 \le x \le 1$ 0, otherwise	1/2	1/12	1, $x \le 0$ $1-x$, $0<x<1$ 0, $x \ge 1$	1, $x \le 0$ $(1-kx)/(1-x)$, $0<x<1$ 0, $x \ge 1$	k^{-1}	k^{-h}
Exponential	$\lambda e^{-\lambda x}$, $x \ge 0$, $\lambda > 0$	$1/\lambda$	$1/\lambda^2$	$e^{-\lambda x}$	$e^{-\lambda(k-1)x}$	k^{-1}	k^{-h}
Weibull	$c\,\theta\,x^{c-1}\exp(-\theta x^c)$ $c>0$, $x \ge 0$	$(\frac{1}{\theta})^c\,\Gamma(\frac{c+1}{c})$	$(\frac{1}{\theta})^{2c}$ $[\Gamma(\frac{c+2}{2}) - \{\Gamma(\frac{c+1}{c})\}^2$	$e^{-\theta x^c}$	$e^{-\theta x^c(k^c-1)}$	no general form	k^{-h}
Pareto	$\alpha(1+x)^{-\alpha-1}$, $\alpha>1$, $x \ge 0$	$\dfrac{\alpha}{\alpha-1}$	$\dfrac{\alpha}{(\alpha-1)^2(\alpha-2)}$	$(1+x)^{-\alpha}$	$\dfrac{\frac{k(1+x)}{(1+kx)^{\alpha}}-1}{k^{\alpha}-1}$	k^{-1}	k^{-h}

It can be noted from the Table 2.5.1 that Uniform, Exponential and Pareto minification processes have geometric autocorrelation function. But all the four minifcation process are uniformly mixing (See Lemma 2.2.2) with mixing parameter

$$\phi(h) = P[X_{n+h} = k^{h} X_n] = k^{-h},\ k > 1,\ h = 1, 2, \ldots \tag{2.5.1}$$

Since Weibull minification process doesnot have a closed form of this type for autocorrelation function, here we consider only Exponential, Uniform and Pareto minification processes.

In Theorem 2.3.1, we have noted that the sample mean $\bar{X}_n$ is CAN estimator for population mean. The asymptotic variance of $\bar{X}_n$ can be calculated using (2.1.7) and (2.5.1) as

$$\begin{aligned} \text{A.V}(\bar{X}_n) &= \frac{Var(X_n)}{n}\left\{1 + 2\sum_{j=1}^{\infty}\rho_X(j)\right\} \\ &= \frac{Var(X_n)}{n}\left\{1 + 2\sum_{j=1}^{\infty}k^{-j}\right\} \\ &= \frac{k+1}{k-1}\frac{Var(X_n)}{n} \end{aligned}$$

Remark 2.5.1 Confidence interval for population mean of Uniform, Exponential and Pareto minification process can be constructed as described in Section 2.3. The general form of such an interval is given by

$$\left[\bar{X}_n \pm Z_{\alpha/2}\frac{k+1}{k-1}\frac{Var(X_n)}{n}\right].$$

We will use the results of this chapter for the sequential estimation of parameters of the minification processes in the next chapter.

The material of this chapter are briefed in the paper Balakrishna and Jacob (1998a).

CHAPTER 3

SEQUENTIAL ESTIMATION FOR MINIFICATION PROCESSES

3.1 Introduction

In many statistical inference problems, some predetermined accuracy is required and usually the optimal fixed sample size to meet this accuracy depends on some nuisance parameters. For example, if we wish to construct a confidence interval for the unknown mean θ of a normal population, $N(\theta,\sigma^2)$ with preassigned accuracy width $2d$ and confidence level γ for given $d > 0$ and $\gamma \in (0,1)$, the optimal fixed sample size procedure requires a sample of size $n_0 = \left(\frac{z\sigma}{d}\right)^2$, where $z = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ and Φ is the cumulative distribution function of $N(0,1)$ r.v. Note however that, the sample size $n_0 = \left(\frac{z\sigma}{d}\right)^2$, depends on σ^2 which is often unknown. To solve such problems it is necessary to use a sequential scheme.

The most frequently used sequential sampling scheme is the fully sequential scheme due to Anscombe (1953), Robbins (1959) and Chow and Robbins (1965). In this scheme of sampling a sample of size m is drawn first and then observations are taken one by one. It renews the estimates of the unknown parameter and the total sample size after each new observation and checks weather enough observations have already been drawn. Not surprisingly this scheme is very efficient in terms of sample size.

The purpose of present chapter is to extend the sequential estimation techniques to minification processes.

The general minification processes and its probabilistic properties are studied in Chapter 2. Compared to i.i.d cases, the literature on sequential estimation in time series emerged some what recently. See Sriram (1987, 1988), Basawa, McCormick and Sriram (1990) for the history of sequential estimation in dependent cases.

The present chapter is organised into four sections. In Section 3.2 we propose sequential procedure to deal with point estimation of mean. Sequential estimation for k in exponential minification process is given in Section 3.3. Section 3.4 contains sequential interval estimation for mean and k.

3.2. Sequential point estimation of mean

Let $X_1, X_2, \ldots, X_n$ be the n observation from the model (1.3.1) and our aim is to estimate $\mu = E(X_i)$. As one can see in the literature on sequential estimation the loss function is often the sum of quadratic loss for the discrepancy between the target parameters and their estimates. Thus here the loss function is

$$L_n(\mu) = \frac{A}{\sigma^2}(\overline{X}_n - \mu)^2, \qquad A>0 \tag{3.2.1}$$

where $\sigma^2 = \mathrm{Var}(X_i)$. The loss function defined in (3.2.1) have the property that for a given μ, the loss increases as the difference between $\overline{X}_n$ and μ increases in either direction. Also this loss function is easy to handle mathematically compared to other loss functions. The expected value of a loss function is called risk function. The aim is to find an estimator for the unknown parameter which have minimum risk under any loss function. Such estimation procedures are known as minimum risk estimation method. Thus in minimum risk estimation problem, minimization of risk w.r.t the choice of sample size leads to the minimum risk estimator (MRE). Here we estimate the parameter μ such that the expected value of $L_n(\mu)$ in (3.2.1) is less than some prescribed value u_μ.

$$\mathrm{E}\, L_n(\mu) = R_n(\mu) = \frac{A}{\sigma^2}\mathrm{E}(\overline{X}_n - \mu)^2$$

Now using Theorem 2.3.1 we have $\mathrm{E}(\overline{X}_n - \mu)^2 \approx \dfrac{\sigma^2 B(k)}{n}$, where $\approx$ means asymptoticaly equal. Hence

$$R_n(\mu) = A\, n^{-1} B(k) + \mathrm{o}(n^{-1}), \qquad (3.2.2)$$

where $B(k)$ is a continuous function of k for $k>1$ and $f = \mathrm{o}(g)$ means that $\dfrac{f}{g} \to 0$.

Let $n_0(\mu)$ be the smallest integer n such that

$$R_n(\mu) \le u_\mu.$$

ie,

$$A\, n^{-1} B(k) \le u_\mu.$$

Thus

$$n_0(\mu) \cong A\, B(k)\, u_\mu. \qquad (3.2.3)$$

It is clear from the sample size defined in (3.2.3) that $n_0(\mu)$ depends on the parameter k. When the parameter k is unknown, nonsequential optimal solutions may not exist in general. As a remedy we go for sequential method of estimation by defining a stopping rule T_μ in analogy with $n_0(\mu)$ by

$$T_\mu = \inf\{n \ge m_u,\ n \ge u_\mu^{-1} A . B(\widetilde{k}_n)\} \qquad (3.2.4)$$

where m_u is an initial sample size imposed to avoid stopping too soon and that depends on the risk bound u_μ. $B(\widetilde{k}_n)$ is obtained by replacing k by $\widetilde{k}_n$ [See (2.4.1) and (2.4.2)] in

$B(k)$. The estimator $\widetilde{k}_n$ and its properties are discussed in Theorem 2.4.1. Now the sequential point estimator for μ is $\overline{X}_{T_\mu}$ with corresponding risk

$$R_{T_\mu} = \frac{A}{\sigma^2}\mathrm{E}(\overline{X}_{T_\mu} - \mu)^2$$

The efficiency of sequential procedure is measured in terms of the convergence properties of the following quantities, under some regularity conditions as cost per observations tends to zero. The quantities of interest are

$$\text{(i) } \frac{T}{n_0} \qquad \text{(ii) } \frac{E(T)}{n_0} \qquad \text{(iii) } \frac{R_T}{R_{n_0}}.$$

Here T denotes the stopping time, n_o the fixed sample size R_T denotes the risk under sequential setup and R_{n_0} the risk under fixed sample size procedure. If $\frac{T}{n_0}$ converges to 1, then we say that the sequential procedure is asymptotically consistent and if $\frac{E(T)}{n_0} \to 1$ we say that the sequential procedure is asymptotically efficient. As a measure of relative efficiency of sequential estimator w.r.t fixed sample size estimator we consider the ratio $\frac{R_T}{R_{n_0}}$. The sequential point estimator is risk efficient if $\frac{R_T}{R_{n_0}} = 1$. However, this is not true in general. But under some conditions if $\frac{R_T}{R_{n_0}}$ converges to 1, then we term the sequential procedure as asymptotically risk efficient.

The main results of this section are summarised in the following Theorem.

Theorem 3.2.1: If for $p > 2$, $\mathrm{E}|Z_1|^{2p} < \infty$ and m_u is such that $u_\mu^{-1/(h+1)} \leq m_u = \mathrm{O}(u_\mu^{-1})$ for $h \in (0, p\text{-}2)$ then as $u_\mu \to 0$

i. $\dfrac{T_\mu}{n_0(\mu)} \to 1$, a.s

ii. $E\left|\frac{T_\mu}{n_0(\mu)} - 1\right| \to 0$

iii. $\sqrt{T_\mu}\ (\overline{X}_{T_\mu} - \mu) \xrightarrow{d} N(0,\ \sigma_1^2)$ where $\sigma_1^2 = \sigma^2 B(k)$

iv. $\frac{R_{T_\mu}}{R_{n_0(\mu)}} \to 1.$

We need some lemmas to prove this theorem and we introduce the following notations for easy reference

$n_1 = n_0(\mu)\,(1-\varepsilon)$ $\qquad$ $n_2 = n_0(\mu)\,(1+\varepsilon)$ $\qquad$ $0<\varepsilon<1$

$C = [T_\mu \le n_1]$ $\qquad$ $D = [T_\mu \ge n_2]$ $\qquad$ $H = [n_1 < T_\mu < n_2].$

Lemma 3.2.1: If $E|Z_1|^{2p} < \infty, p\ge 1$ then

$$\left\|\overline{X}_n - \mu\right\|_{2p} = O(n^{-1/2}).$$

Proof: From the definition of the model (1.3.1) we can write

$$\overline{X}_n - \mu \le n^{-1}\sum_{i=1}^{n}(kZ_i - \mu)$$

$$= n^{-1} k \left[\sum_{i=1}^{n}(Z_i - \alpha) + (\alpha - \mu/k)\right] \tag{3.2.5}$$

where $\qquad \alpha = E(Z_i)$

Thus using Minkowski inequality (See Result 1.6.1)

$$\left\|\overline{X}_n - \mu\right\|_{2p} \le n^{-1} k \left\|\sum_{i=1}^{n}(Z_i - \alpha)\right\|_{2p} + k\left\|(\alpha - \mu/k)\right\|_{2p}$$

Now since Z_i's are iid with $E|Z_1|^{2p} < \infty$, we can use Marcinkiewicz -Zygmund inequality [See Result 1.6.10] to the first term to get

$$\left\| \sum_{i=1}^{n} (Z_i - \alpha) \right\|_{2p} = O(n^{1/2}).$$

Thus

$$\left\| \bar{X}_n - \mu \right\|_{2p} \leq n^{-1}\, O(n^{1/2}) = O(n^{-1/2})$$

The lemma is proved.

Lemma 3.2.2: If $\{Z_n\}$ is a sequence of nonnegative and non-degenerate r.v.s and m_u is such that $u_\mu^{1/(h-1)} \leq m_u = O(u_\mu^{-1})$, $h \in (0, p-2)$ for $p > 2$. Then as $u_\mu \to 0$,

1. $P[T_\mu \leq n_l] = O(u_\mu^{(p-1)(1+h)^{-1}})$

2. $\sum_{n > n_2} P[T_\mu > n] = O(u_\mu^{(p-2)/2})$.

Proof: From the definition of stopping time (3.2.4) if $T_\mu \leq n_l$ then

$$u_\mu^{-1}\, A.B(\tilde{k}_n) \leq n_l \quad \text{for some } m_u \leq n \leq n_l.$$

Thus

$$P[T_\mu \leq n_l] \leq P[u_\mu^{-1}\, A.\, B(\tilde{k}_n) \leq n_l \text{ for some } m_u \leq n \leq n_l]$$

$$\leq P[\, B(\tilde{k}_n) - B(k) \leq -\varepsilon B(k) \text{ for some } m_u \leq n \leq n_l]$$

$$\leq P\left[\underset{m_u \leq n \leq n_l}{Max} \left| B(\tilde{k}_n) - B(k) \right| > \varepsilon' \right], \text{ where } \varepsilon' = \varepsilon B(k)$$

$$\leq \sum_{n = m_u}^{\infty} P\left[\left| B(\tilde{k}_n) - B(k) \right| > \varepsilon' \right]$$

$$\le \sum_{n=m_u}^{\infty} P\left[\left|\tilde{k}_n - k\right| > \eta\right], \text{ for } \eta > 0. \qquad (3.2.6)$$

The last inequality is due to the fact that $B(.)$ is a continuous function of k.

Now consider

$$P\left[\left|\tilde{k}_n - k\right| > \eta\right] = P[(\tilde{k}_n\text{-}k)>\eta] + P[(\tilde{k}_n\text{-}k)<-\eta]$$

$$= 0 + P[(\tilde{k}_n\text{-}k)<-\eta]$$

$$\le P\left[\underset{1\le i\le n}{Max}\left(\frac{Z_i}{Z_{i-1}}\right) < k-\eta\right]$$

$$= P\left[\frac{Z_1}{Z_0} < k-\eta, \frac{Z_2}{Z_1} < k-\eta, \frac{Z_4}{Z_3} < k-\eta, \ldots, \frac{Z_n}{Z_{n-1}} < k-\eta\right]$$

$$\le P\left[\frac{Z_2}{Z_1} < k-\eta, \frac{Z_4}{Z_3} < k-\eta, \ldots, \frac{Z_n}{Z_{n-1}} < k-\eta\right]$$

$$\le P\left[\frac{Z_2}{Z_1} < k-\eta\right]^{\left[\frac{n}{2}\right]-1} = a^{\left[\frac{n}{2}\right]-1},$$

where $a = P\left[\frac{Z_2}{Z_1} < k-\eta\right]$.

Since $\{Z_n\}$ is a sequence of nondegenerate r.v.s it is true that $0 \le P\left[\frac{Z_2}{Z_1} < x\right] < 1$ for $x > 0$.

Consider

$$\frac{P\left[\left|\tilde{k}_n - k\right| > \eta\right]}{n^{-p}} \le \frac{a^{\left[\frac{n}{2}\right]-1}}{n^{-p}} \to \frac{0}{0} \text{ as } n \to \infty.$$

Therefore,

$$\lim_{n\to\infty}\frac{a^{\left[\frac{n}{2}\right]-1}}{n^{-p}} \le \lim_{n\to\infty}\frac{a^{\frac{n}{2}-2}}{n^{-p}} = \lim_{n\to\infty}\frac{n^p}{a^{-\frac{n}{2}+2}} = \frac{\infty}{\infty}$$

$$= \lim_{n\to\infty} \frac{pn^{p-1}}{a^{-\frac{n}{2}+2} . \log a}$$

$$=$$

$$= \lim_{n\to\infty} \frac{p(p-1)\ldots 1.n^0}{(\log a)^p a^{\frac{n}{2}+2}}$$

$$= \lim_{n\to\infty} \frac{(1.2\ldots p)}{(\log a)^p} a^{\frac{n}{2}-2}, \qquad 0 < a < 1$$

$$= 0.$$

Therefore,

$$P\left[\left|\widetilde{k}_n - k\right| > \eta\right] = o(n^{-p}).$$

This implies that

$$P\left[\left|\widetilde{k}_n - k\right| > \eta\right] = O(n^{-p}). \qquad (3.2.7)$$

Now using (3.2.7) in (3.2.6) we have

$$P[T_\mu \leq n_l] = \sum_{n=m_u}^{\infty} O(n^{-p})$$

$$= c\left[\frac{1}{m_u^p} + \frac{1}{(m_u+1)^p} + \ldots\right] \text{ for some } 0 < c < \infty.$$

Now we have the following relations

$$\frac{1}{m_u^p} + \frac{1}{(m_u+1)^p} + \ldots + \frac{1}{(2m_u-1)^p} < \frac{m_u}{m_u^p} = \frac{1}{m_u^{p-1}}$$

$$\frac{1}{(2m_u)^p} + \frac{1}{(2m_u+1)^p} + \ldots + \frac{1}{(4m_u-1)^p} < \frac{2m_u}{(2m_u)^p} = \frac{1}{(2m_u)^{p-1}}$$

$$\frac{1}{(4m_u)^p}+\frac{1}{(4m_u+1)^p}+\ldots+\frac{1}{(8m_u-1)^p}<\frac{4m_u}{(4m_u)^p}=\frac{1}{(4m_u)^{p-1}}$$

Hence $\frac{1}{m_u^p}+\frac{1}{(m_u+1)^p}+\ldots<\frac{1}{m_u^{p-1}}+\frac{1}{(2m_u)^{p-1}}+\frac{1}{(4m_u)^{p-1}}+$

$$=\frac{1}{m_u^{p-1}}\left[1+\frac{1}{2^{p-1}}+\frac{1}{2^{2(p-1)}}+\ldots\right]$$

Now the series $1+\frac{1}{2^{p-1}}+\frac{1}{2^{2(p-1)}}+\ldots$ is a geometric series whose common ratio is $\frac{1}{2^{p-1}}$ is less than unity since p>2. Hence the sum of this geometric series is finite. Thus we have

$$P[T_\mu \le n_1] = O(m_u^{p+1})$$

$$= O(u_\mu^{(p-1)(1+h)^{-1}}),$$

where we used the condition $u_\mu^{1/(h+1)} \le m_u$.

Hence first part of the Lemma is proved.

For the second part, from the definition of T_μ it follows that for $n \ge n_2$

$$P[T_\mu > n] \le P[u_\mu^{-1}\, A\, B(\widetilde{k}_n) > n]$$

$$= P[B(\widetilde{k}_n) > n_2 A^{-1} u_\mu]$$

$$= P[\, B(\widetilde{k}_n) - B(k) > \varepsilon B(k)\,]$$

$$= P[\,|\,\widetilde{k}_n - k\,| > \varepsilon'\,] = O(n^{-p})$$

As in the proof of part (1) here we can prove

$$\sum_{n \ge n_2} P[T_\mu > n] = O(u_\mu^{(p-2)/2}).$$

This completes the proof.

Lemma 3.2.3: Under the conditions of Lemma 3.2.2, $\left\{u_\mu^{-1}(\overline{X}_{T_\mu}-\mu)^2 I_{(n_1<T_\mu<n_2)}\right\}$, $0< u_\mu< u_0$, $u_0<1$ is uniformly integrable.

Proof: From the definition of uniform integrability [See Definition 1.6.4] it is enough to prove that

$$\underset{T_\mu}{Sup}\, E\left|u_\mu^{-1}(\overline{X}_{T_\mu}-\mu)^2 I_{(n_1<T_\mu<n_2)}\right|<\infty.$$

Using (3.2.5) we can write

$$E\left|u_\mu^{-1}(\overline{X}_{T_\mu}-\mu)^2 I_H\right| \le u_\mu^{-1} E\left|\underset{n_1<n<n_2}{Max}(\overline{X}_n-\mu)^2\right| I_H$$

$$\le u_\mu^{-1}k\, E\left|\underset{n_1<n<n_2}{Max}\, n^{-1}\sum_{i=1}^{n}(Z_i-\alpha)+(\alpha-\mu/k)\right|^2 I_H$$

$$\le u_\mu^{-1}kn_1^{-2}\, E(F_n^2\, I_H)+2\, n_1^{-1}\, u_\mu^{-1}k\,(\alpha-\mu/k)\, E(F_n\, I_H)$$

$$+u_\mu^{-1}k\, E(I_H)(\alpha-\mu/k)^2. \qquad (3.2.8)$$

where $F_n = \underset{n_1<n<n_2}{Max}\, n^{-1}\left|\sum_{i=1}^{n}(Z_i-\alpha)\right|$ is a submartingale w.r.t. $F_n=\sigma\{X_0, Z_1, \ldots, Z_n\}$. Now using Schwartz inequality, Maximal inequality for submartingales and M-Z inequality, the first term in (3.2.8) can be written as

$$u_\mu^{-1}kn_1^{-2}\, E(F_n^2\, I_H)\le u_\mu^{-1}kn_1^{-2}\, E^{1/2}(F_n^4)P^{1/2}(H)$$

$$\le u_\mu^{-1}kn_1^{-2}\, P^{1/2}(H)\, E^{1/2}\left\{\left|\underset{n_1<n<n_2}{Max}\sum_{i=1}^{n}(Z_i-\alpha)\right|^4\right\}$$

$$\leq u_{\mu}^{1} k n_{1}^{2} \; P^{1/2}(H) \; E^{1/2}\left\{\left|\sum_{i=1}^{n_2}(Z_i - \alpha)\right|^4\right\}$$

$$\leq u_{\mu}^{1} k n_{1}^{2} \; P^{1/2}(H) \, \mathrm{O}(n_2).$$

Now using Lemma 3.2.2, we have

$$u_{\mu}^{-1} k n_{1}^{2} \; E(F_n^2 \, I_H) < \infty.$$

Similarly the second term in (3.2.8) can be written as

$$2 n_1^{1} \, u_{\mu}^{1} k \, (\alpha - \mu k) \, E(F_n \, I_H) \leq n_1^{1} \, u_{\mu}^{1} k \, (\alpha - \mu k) \; n_1^{2} \; E^{1/2}(F_n^2) \, P^{1/2}(H).$$

Repeating the same arguments as above we have

$$n_1^{1} \, u_{\mu}^{-1} k \, E(F_n \, I_H)(\alpha - \mu k) < \infty.$$

Similarly the third term in (3.2.8) is finite. Thus we have proved the lemma.

Lemma 3.2.4: If $\mathrm{E}|Z_1|^{2p} < \infty$, $p \geq 1$ then $\{\sqrt{n}(\overline{X}_n - \mu), n > 1\}$ is uniformly continuous in probability.

Proof: We have $\sqrt{n}(\overline{X}_n - \mu) = \sum_{i=1}^{n}(X_i - \mu)/\sqrt{n}$

Letting $\sum_{i=1}^{n}(X_i - \mu) = Q_n$ and following Woodroofe (1982) (cf. pp.11)

we can write

$$\left|\frac{Q_{n+j}}{\sqrt{n+j}} - \frac{Q_n}{\sqrt{n}}\right| \leq \frac{1}{\sqrt{n}}\left|Q_{n+j} - Q_n\right| + \left[1 - \left(\frac{n}{n+j}\right)^{1/2}\right]\left|\frac{Q_n}{\sqrt{n}}\right|. \qquad (3.2.9)$$

If $j \le n\delta$ the second term on the right hand side (3.2.9) is stochastically bounded by

$$[1-(1+\delta)^{-1/2}]\ \|(\bar{X}_n - \mu)\|\sqrt{n} = [1-(1+\delta)^{-1/2}]\ O(1) \text{ by Lemma 3.2.1}$$

$$= O(\delta)$$

which tends to zero as $\delta \to 0$ uniformly in n.

Thus we have

$$P\left[\underset{0<j\le n\delta}{Max}\left|\left[1-\left(\tfrac{n}{n+j}\right)^{1/2}\right]\frac{Q_n}{\sqrt{n}}\right| > \frac{\varepsilon}{2}\right] \to 0, \text{ as } \delta \to 0$$

Now consider

$$P\left[\underset{0<j\le n\delta}{Max}\frac{1}{\sqrt{n}}\left|Q_{n+j} - Q_n\right| > \frac{\varepsilon}{2}\right] = P\left[\underset{0<j\le n\delta}{Max}\left|\sum_{i=n+1}^{n+j}(X_i - \mu)\right| > \frac{\varepsilon\sqrt{n}}{2}\right].$$

Using (3.2.5) we can write

$$P\left[\underset{0<j\le n\delta}{Max}\left|\sum_{i=n+1}^{n+j}(X_i - \mu)\right| > \frac{\varepsilon\sqrt{n}}{2}\right] \le P\left[\underset{0<j\le n\delta}{Max}\left|k\sum_{i=n+1}^{n+j}(Z_i - \alpha) + jk(\alpha - \tfrac{\mu}{k})\right| > \frac{\varepsilon\sqrt{n}}{2}\right].$$

Note that $k\left\{\sum_{i=n+1}^{n+j}(Z_i - \alpha) + jk(\alpha - \tfrac{\mu}{k})\right\}$ is a submartingale w.r.t $G_n = \sigma\{X_0\, Z_1 \ldots Z_n\}$.

Using maximal inequality for this submartingale we have

$$P\left[\underset{0<j\le n\delta}{Max}\frac{1}{\sqrt{n}}\left|Q_{n+j} - Q_n\right| > \frac{\varepsilon\sqrt{n}}{2}\right] \le \frac{4}{\varepsilon^2 n}\ \mathrm{E}\left|\sum_{i=n+1}^{n+n\delta}k(Z_i - \alpha) + n\delta k(\alpha - \tfrac{\mu}{k})\right|$$

$$\le \frac{4}{\varepsilon^2}\ k\delta\ \mathrm{E}[|(Z_i - \alpha)| + |(\alpha - \tfrac{\mu}{k})|]$$

$$= O(\delta) \to 0 \text{ uniformly in } n, \text{ since } E|Z_1|^{2p} < \infty.$$

Thus $\{\sqrt{n}(\overline{X}_n - \mu), n>1\}$ is u.c.i.p.

Now we are in position to prove the theorem.

Proof of Theorem 3.2.1

Since $\widetilde{k}_n$ is a strongly consistent estimator of k and $B(k)$ is a continuous function of k, it follows from (3.2.4) that $T_\mu < \infty$ and $T_\mu \to \infty$ as $u_\mu \to 0$. Also $B(\widetilde{k}_{T_\mu}) \xrightarrow{a.s} B(k)$.

From the definition of stopping rule (3.2.4) we have

$$T_\mu \geq m_u$$

$$T_\mu \geq u_\mu^{-1} A.\ B(\widetilde{k}_{T_\mu})$$

and

$$T_\mu < u_\mu^{-1} A.\ B(\widetilde{k}_{T_\mu}).$$

Thus

$$u_\mu^{-1} A.\ B(\widetilde{k}_{T_\mu}) < T_\mu \leq u_\mu^{-1} A.\ B(\widetilde{k}_{T_\mu}) + m_u. \qquad (3.2.10)$$

Dividing (3.2.10) by $n_0(\mu)$ and using the above arguments it follows that

$$\frac{T_\mu}{n_0(\mu)} \xrightarrow{a.s} 1$$

if m_u is such that, $\dfrac{m_u}{n_0(\mu)} \to 0$.

For **part (ii)** we can write

$$\mathrm{E}\left|\frac{T_\mu}{n_0(\mu)}-1\right| = \mathrm{E}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{+} + \mathrm{E}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{-} \qquad (3.2.11)$$

But $\left(\frac{T_\mu}{n_0(\mu)}-1\right) = \left(\frac{T_\mu}{n_0(\mu)}-1\right) I_C + \left(\frac{T_\mu}{n_0(\mu)}-1\right) I_D + \left(\frac{T_\mu}{n_0(\mu)}-1\right) I_H$

and $\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{+} I_c = 0.$

Now using Lemma 3.2.2,

$$\mathrm{E}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{+} = \mathrm{E}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{+} I_H + \mathrm{E}\left(\frac{T_\mu - n_2}{n_0(\mu)}\right)^{+} I_D + \mathrm{E}\left(\frac{n_2-1}{n_0(\mu)}\right)^{+} P(D)$$

$$= \varepsilon P(H) + n_0^{-1}(\mu)\,\mathrm{E}[(T_\mu - n_2)^{+} I_D] + n_0^{-1}(\mu)\,(n_2 - n_0^{-1}(\mu))^{+} P(D)$$

$$= \varepsilon + \varepsilon + \mathrm{o}(1),$$

since ε is arbitrary, as $u_\mu \to 0$, we have

$$\mathrm{E}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{+} \to 0. \qquad (3.2.12)$$

Now dominated convergence theorem can be applied to the second term in (3.2.11), since

$\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{-} \leq 1.$

Thus

$$\lim_{u_\mu \to 0} E\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{-} = E \lim_{u_\mu \to 0}\left(\frac{T_\mu}{n_0(\mu)}-1\right)^{-} = 0. \qquad (3.2.13)$$

Now part (ii)of the Theorem follows from (3.2.11), (3.2.12) and (3.2.13).

For **part (iii)** recall from Theorem 2.3.1 that

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} \mathrm{N}(0, \sigma_1^2).$$

We need to show that this result continues to hold when n is replaced by the stopping time T_μ. In Lemma 3.2.4 we have verified that $\{\sqrt{n}(\overline{X}_n - \mu),\ n \geq 1\}$ is u.c.i.p. Then one can conclude, using Anscombe's Theorem [See Result 1.6.5], (2.3.2) and from part (i) that

$$\sqrt{T_\mu}(\overline{X}_{T_\mu} - \mu) \xrightarrow{d} \mathrm{N}(0, \sigma_1^2), \text{ as } u_\mu \to 0.$$

In part (iv) we will show that the risk of the sequential procedure $\overline{X}_{T_\mu}$ defined by

$$R_{T_\mu} = A\,\mathrm{E}(\overline{X}_{T_\mu} - \mu)^2 \text{ is close to } R_{n_0(\mu)}, \text{ as } u_\mu \to 0.$$

Consider

$$\frac{R_{T_\mu}}{R_{n_0(\mu)}} = \frac{AE(\overline{X}_{T_\mu} - \mu)^2}{u_\mu \sigma^2} I_{[C \cup H \cup D]}, \tag{3.2.14}$$

where the events C,D,H are as defined earlier.

Using Schwartz inequality, Lemma 3.2.1 and Lemma 3.2.2, we write

$$\frac{AE(\overline{X}_{T_\mu} - \mu)^2}{u_\mu \sigma^2} I_C \leq \frac{A}{u_\mu \sigma^2} E^{1/p}\left[\underset{m_\mu < n < n_1}{Max}(\overline{X}_n - \mu)^{2p}\right] P^{1/q}(C),$$

where $\frac{1}{p} + \frac{1}{q} = 1$

$$= \frac{A}{u_\mu \sigma^2} \left[\sum_{n=m_\mu}^{\infty} n^{-p}\right]^{1/p} P^{1/q}(C)$$

$$= \frac{A}{\sigma^2} u_\mu^{\,1} \mathrm{O}(m_u^{(p\ 1)/p}) \mathrm{O}(u_\mu^{(p\ 1)/(h+1)q})$$

$$= \frac{A}{\sigma^2} u_\mu^{\,1}\ \mathrm{O}(u_\mu^{(p\ 1)\ (h+1)p}) \mathrm{O}(u_\mu^{(p-1)/(h+1)q})$$

$$= \mathrm{o}(1) \text{ as } u_\mu \to 0, \text{ since } h \in (0,p\text{-}2). \qquad (3.2.15)$$

Similarly using Schwatz inequality, Lemma 3.2.1 and part (2) of Lemma 3.2.2 we have

$$\frac{A}{u_\mu^{-1}\sigma^2} \mathrm{E}\ (\overline{X}_{T_\mu} - \mu)^2 I_D = \mathrm{o}(1). \qquad (3.2.16)$$

Now we will prove

$$\frac{A}{\sigma^2} u_\mu^{\,1}\ \mathrm{E}\ (\overline{X}_{T_\mu} - \mu)^2 I_H \to 1 \ \text{ as } u_\mu \to 0.$$

Using part (iii) and Lemma 3.2.3 it follows that

$$\mathrm{E}\ (\overline{X}_{T_\mu} - \mu)^2 I_H \to \frac{\sigma_1^2}{T_\mu} \ \text{ as } u_\mu \to 0.$$

But we have $\sigma_1^2 = B(k)\ \sigma^2$ and $n_0(\mu) \approx AB(k)\ u_\mu$ to write

$$\frac{A}{\sigma^2} u_\mu^{\,1}\ \mathrm{E}\ (\overline{X}_{T_\mu} - \mu)^2 I_H \to \frac{A}{\sigma^2} u_\mu^{\,1}\ \sigma_1^2\ n_0^{\,1}(\mu)$$

$$\approx \frac{A}{u_\mu} \frac{B(k)}{n_0(\mu)} \to 1. \qquad (3.2.17)$$

Now the asymptotic risk efficiency follows from (3.2.14), (3.2.15), (3.2.16) and (3.2.17). Hence the theorem is completely proved.

3.3 Sequential Estimation of k in Exponential Minification processes

Exponential minification process have many nice features compared to other minification processes. This section deals with sequential estimation of k of th exponential minification process.

Even though $\tilde{k}_n$ defined in (2.4.2) is consistent for k in a general minification process, it is not CAN. However, for the exponential case the estimator suggested is $\hat{k}_n$ given by

$$\hat{k}_n = \frac{\overline{U}_n}{2\overline{U}_n - 1} \tag{3.3.1}$$

is CAN as discussed in Theorem 2.4.1.

For the sake of algebraic simplicity we consider a loss function of the form

$$L_{nk} = C[(2k-1)\,\overline{U}_n - k]^2, \quad C>0. \tag{3.3.2}$$

for estimating k using (3.3.1). Using Theorem 2.4.1 (See (2.4.8) and (2.4.20)) the corresponding risk is given by

$$R_{nk} = \mathrm{E}[L_{nk}] = C(2k-1)^2\, \mathrm{E}[\overline{U}_n - (k/(2k-1))]^2$$

$$\cong C n^{-1} \left[k(k-1) - 2\sum_{h=1}^{\infty} \frac{(k-1)^3}{(k-1) + k^{h-1}(2k-1)} \right]$$

$$\cong C n^{-1} H(k) \text{ (say)},$$

where $$H(k) = k(k-1) - 2\sum_{h=1}^{\infty} \frac{(k-1)^3}{(k-1) + k^{h-1}(2k-1)} \tag{3.3.3}$$

Note that $H(k)$ defined by (3.3.3) is a continuous function of k for $k>1$. As in the case of population mean here also we calculate the sample size such that the risk is less than some prescribed limit say u_k. That is,

$$R_{nk} \leq u_k. \tag{3.3.4}$$

Let n_{0k} be the smallest integer n such that (3.3.4) holds. That is,

$$C\, n^{-1}\, H(k) \leq u_k.$$

Thus

$$n_{0k} \doteq C\, n^{-1}\, H(k). \tag{3.3.5}$$

Note that this fixed sample size procedure depends on the unknown parameter k.

Let us define a stopping time by

$$T_k = \inf\{\, n \geq m_k : n \geq C\, u_k^{-1}\ H(\hat{k}_n)\}, \tag{3.3.6}$$

where m_k is an initial sample size that may dependent on u_k, $H(\hat{k}_n)$ is obtained by replacing k by $\hat{k}_n$ in (3.3.3)

Based on this stopping rule the sequential point estimator of k is $\hat{k}_{T_k}$ with corresponding risk R_{T_k}. The optimal properties of this sequential procedure are summarised in the following Theorem.

Theorem 3.3.1: If E$\,|\,Z_I\,|^{2p} < \infty$ for $p>2$, and m_k is such that $u_k^{-1/(h+1)} \leq m_k = O(u^{-1})$ for $h \in (0, p-2)$, then as $u_k \to 0$

1. $\dfrac{T_k}{n_{0k}} \xrightarrow{a.s} 1$

2. $\mathrm{E}\left|\dfrac{T_k}{n_{0k}} - 1\right| \to 0$

3. $\sqrt{T_k}\,(\hat{k}_{T_k} - k) \xrightarrow{d} \mathrm{N}(0,\ \sigma_2^2)$, where σ_2^2 is defined by (2.4.6)

4. $\dfrac{R_{T_k}}{R_{n_0 k}} \to 1.$

The following lemma's are needed to prove this theorem and hence we prove them first.

Lemma 3.3.1: Let $\{Xn\}$ be an exponential minification sequence defined in section 2.4 and if $E|Z_1|^{2p} < \infty$ for $p \geq 1$ then

$\left\|(2k-1)\overline{U}_n - k\right\|_{2p} = O(n^{-1/2})$, where $\overline{U}_n = (U_1 + U_2 + \ldots + U_n)/n$ and U_j is defined by (2.4.4.)

Proof: Consider

$$\left\|(2k-1)\overline{U}_n - k\right\|_{2p} = (2k-1)\left\|\overline{U}_n - \frac{k}{(2k-1)}\right\|_{2p}$$

$$= (2k-1)n^{-1}\left\|\sum_{i=1}^{n}(U_i - \frac{k}{(2k-1)})\right\|_{2p}$$

Note that $\sum_{i=1}^{n}(U_i - \frac{k}{(2k-1)})$ is a zero mean martangale w.r.t $F_n = \sigma\{X_0, Z_1, \ldots Z_n\}$. Then by applying Burkholder inequality (See Result 1.6.1) and moment inequality we have

$$B_p^{-1} n^{-1/2}\left\|\sum_{i=1}^{n}(U_i - \frac{k}{(2k-1)})\right\|_{2p} \leq \left\|\left[n^{-1}\sum_{i=1}^{n}(U_i - \tfrac{k}{(2k-1)})^2\right]^{1/2}\right\|_{2p}$$

$$\leq \left\|\left[n^{-1}\sum_{i=1}^{n}(U_i - \tfrac{k}{(2k-1)})^p\right]^{1/p}\right\|_{2p}$$

$$= O(1),$$

where $B_p = 18p^{3/2}(p-1)^{-1/2}$ Hence

$$\left\|\sum_{i=1}^{n}(U_i - \frac{k}{(2k-1)})\right\|_{2p} = O(n^{1/2}).$$

Therefore

$$\left\|(2k-1)\bar{U}_n - k\right\|_{2p} = (2k\text{-}1)\, n^{-1}\, O(n^{1/2})$$

$$= O(n^{-1/2}).$$

The lemma is proved.

For the following lemma we need to introduce some notations

Let $n_1 = n_{0k}(1\text{-}\varepsilon)$, $n_1 = n_{0k}(1+\varepsilon)$, $A = [T_k \le n_1]$, $D = [T_k \ge n_2]$ and $E = [n_1 < T_k < n_2]$.

Lemma 3.3.2: If $E\,|\,Z_1\,|^{2p} < \infty$ for $p>2$ and m_k is such that $u_k^{-1/(h+1)} \le m_k = O(u_k^{-1})$, where u_k is as definded in Section 3.3. Then for $h \in (0,p\text{-}2)$,

1. $P[T_k \le n_1] = O(u_k^{(p-2)/(n+1)})$

2. $\sum_{n > n_2} [T_k \ge n_2] = O(u_k^{p-1})$

Proof: From the definition of stopping time T_k (3.3.6)

$T_k \le n_1$ implies

$C u_k^{-1} H(\hat{k}_n) \le n_1$ for some $m_k < n \le n_1$.

Thus using the definition of n_1

$$P[T_k \le n_1] \le P[C u_k^{-1} H(\hat{k}_n) \le n_1 \text{ for some } m_k < n \le n_1]$$

$$= P[H(\hat{k}_n) - H(k) \le -\varepsilon H(k) \text{ for some } m_k < n \le n_1]$$

$$\le P\left[\underset{m_k < n \le n_1}{Max}\left|H(\hat{k}_n) - H(k)\right| > \varepsilon'\right] \text{ where } \varepsilon' = \varepsilon H(k)$$

$$\leq \sum_{n \geq m_k}^{s} \mathrm{P}\left[\left|H(\hat{k}_n) - H(k)\right| > \varepsilon'\right]$$

$$\leq \sum_{n \geq m_k}^{s} \mathrm{P}\left[\left|\hat{k}_n - k\right| > \eta\right]. \tag{3.3.8}$$

Now we will prove

$$\mathrm{P}\left[\left|\hat{k}_n - k\right| > \eta\right] = \mathrm{O}(n^{-p/2}) \tag{3.3.9}$$

In view of relation (3.3.1) and Result 1.6.12 it is enough to prove

$$\mathrm{P}\left[\left|\overline{U}_n - \tfrac{k}{2k-1}\right| > \varepsilon\right] = \mathrm{O}(n^{-p}).$$

Consider

$$\left\|\overline{U}_n - \tfrac{k}{2k-1}\right\|_{2p}.$$

We have already proved in Lemma 3.3.1 that

$$\left\|\overline{U}_n - \tfrac{k}{2k-1}\right\|_{2p} = \mathrm{O}(n^{-1/2}).$$

Using Markov inequality for $\varepsilon > 0$, we get

$$\mathrm{P}\left[\left|\overline{U}_n - \tfrac{k}{2k-1}\right| > \varepsilon\right] \leq \frac{E\left|\overline{U}_n - \tfrac{k}{2k-1}\right|^p}{\varepsilon^p} = \mathrm{O}(n^{-p/2}).$$

Now using the Result 1.6.12,

$$\mathrm{P}\left[\left|\frac{\overline{U}_n}{2\overline{U}_n - 1} - \frac{\frac{k}{2k-1}}{\frac{2k}{2k-1} - 1}\right| > \varepsilon\right] = \mathrm{O}(n^{-p/2}).$$

Hence the required result (3.3.9) follows from the above equation.

Combining (3.3.8) and (3.3.9)

$$P[T_k \le n_l] \le \sum_{n > m_k} O(n^{-p/2})$$

$$= O(m_k^{-(p/2)+1})$$

$$= O(u_k^{(p/2)/2(h+1)}).$$

This proves the first part of the lemma.

On similar lines the second part can be proved. We have already provided a similar result in lemma 3.3.2. Hence we omit the details.

Lemma 3.3.3: Under the conditions of lemma 3.3.2,

$$\{ u_k^{-1} [(2k-1)\overline{U}_{T_k} -k]^2 \; I_{(n_1 < T_k < n_2)},\ 0< u_k < u_0\}$$

is uniformly integrable.

Proof: By the definition of uniform integrability it is enough to show that

$$\underset{T_k}{Sup}\ E \mid u_k^{-1} [(2k-1)\ \overline{U}_{T_k} -k]^2 I_E \mid < \infty.$$

Consider

$$E \mid u_k^{-1} [(2k-1)\ \overline{U}_{T_k} -k]^2 I_E \mid \le u_k^{-1} E \left| \underset{n_1 < n < n_2}{Max} [(2k-1)\overline{U}_n - k] \right|^2$$

$$\le u_k^{-1} (2k-1)^2 E \left| \underset{n_1<n<n_2}{Max} [\overline{U}_n - \tfrac{k}{2k-1}] \right|^2 I_E$$

$$\le u_k^{-1} \frac{(2k-1)^2}{n_1^2} E \left| \underset{n_1<n<n_2}{Max} \sum_{i=1}^{n} [U_i - \tfrac{k}{2k-1}] \right|^2 I_E$$

Note that $\sum_{i=1}^{n} [U_i - \frac{k}{2k-1}]$ is a martingle w.r.t $G_n = \sigma\{X_0\ Z_1\ ...Z_n\}$ and hence

$V_n = \left| \underset{n_1 \le n \le n_2}{Max} \sum_{i=1}^{n} [U_i - \frac{k}{2k-1}] \right|$ is a submartingale.

Now using Schwartz inequality and maximal inequality for submartingales,

$$\mathrm{E}\ | u_k^{-1} [(2k\text{-}1)\ \overline{U}_{T_k}\ \text{-}k]^2 I_E | \le u_k^{-1} \frac{(2k-1)^2}{n_1^2} \mathrm{E}^{1/2}\ V_{n_2}^4\ P^{1/2}(E). \tag{3.3.10}$$

Consider $\mathrm{E}(V_{n_2}^{\prime 2}) = \mathrm{E}\left|\sum_{i=1}^{n_2} [U_i - \frac{k}{2k-1}]\right|^4$

Using M-Z inequality we have

$$\mathrm{E}(V_{n_2}^4) = \mathrm{O}(n_2^2) \qquad \text{[See Lemma 3.3.1]}$$

Using (3.3.10) , lemma 3.3.2 and above arguments we have

$$\mathrm{E}\ | u_k^{-1} [(2k\text{-}1)\ \overline{U}_{T_k}\ \text{-}k]^2 I_E | \le u_k^{-1} \frac{(2k-1)^2}{n_1^2} \mathrm{O}(n_2)\, P^{1/2}(E) < \infty.$$

Hence the proof of the lemma is complete.

Lemma 3.3.4: $\{\sqrt{n}(\hat{k}_n - k),\ n \ge 1\}$ is stochastically bounded and uniformly continuous in probability.

Proof: Using (3.3.1), $\sqrt{n}(\hat{k}_n - k)$ can be written as

$$\sqrt{n}(\hat{k}_n - k) = \sqrt{n}\left(\frac{\overline{U}_n}{2\overline{U}_n - 1} - k\right)$$

$$= \frac{\sqrt{n}[(1-2k)\overline{U}_n + k]}{2\overline{U}_n - 1}$$

$$= \frac{[(1-2k)\overline{U}_n + k]/\sqrt{n}}{(2\overline{U}_n - 1)/n} \qquad (3.3.11)$$

We will prove that the terms in numerator and denominator of (3.3.11) are u.c.i.p and stochastically bounded and then use the Remark 1.6.3 to get the required result.

As $\overline{U}_n \leq 1$, it follows that $\{[(1-2k)\overline{U}_n + k]/\sqrt{n}\}$ and $[(2\overline{U}_n - 1)/n]$ converge to zero almost surely as $n \to \infty$. Thus by Remark 1.6.2 these terms are u.c.i.p and stochastically bounded. Since any continuous function of u.c.i.p and stochastically bounded sequences is again u.c.i.p (cf. Remark 1.6.3) lemma 3.3.4 now follows easily.

The proof of Theorem 3.3.1 is skipped as it is parallel to the proof of Theorem 3.2.1. In the next section we will consider sequential interval estimation for the mean and k.

3.4 Sequential Interval Estimation

This section is devoted to the study of sequential interval estimation for the mean of general minification processes defined by (2.1.1). In the iid setup Chow and Robbins (1965) proposed a sequential confidence interval for the mean θ of a population with finite variance as discribed below. They consider a situation where $\{X_n\}$ is a sequence of iid observations and $\hat{\theta}_{Ln}, \hat{\theta}_{Un}$ (both based on $X_1, X_2, \ldots, X_n$) such that $\hat{\theta}_{Ln} \leq \hat{\theta}_{Un}$ and $P[\hat{\theta}_{Ln} \leq \theta \leq \hat{\theta}_{Un}] \geq 1-\alpha$. In this case $1-\alpha$ is referred to as a confidence coefficient or the coverage probability and $\alpha \in (0,1)$. In the confidence interval, $\hat{\theta}_{Ln}$ and $\hat{\theta}_{Un}$ are the lower and upper confidence limits and the width of this interval is equal to $\hat{\theta}_{Un}$ - $\hat{\theta}_{Ln}$. In many problems of practical interest one wants to provide such a confidence interval for a parameter of interest satisfying the additional condition that for some preassigned $d(>0)$.

$$0 < \hat{\theta}_{Un} - \hat{\theta}_{Ln} \leq 2d.$$

Assume the estimator T_n for θ is strongly consistent and $\sqrt{n}\,(T_n - \theta)$ is asymptotically normally distributed as $n \to \infty$ say $N(0,\sigma^2)$.

Then

$$\lim_{n\to\infty} P_\theta(T_n - n^{-1/2}\sigma Z_{1-\alpha/2} \leq \theta \leq T_n + n^{-1/2}\sigma Z_{1-\alpha/2}) = 1\text{-}\alpha. \tag{3.4.1}$$

where $Z_{1-\alpha/2} = \Phi^{-1}(1- \frac{\alpha}{2})$, Φ being the standard normal distribution function. Consider the interval

$$I_{n_d} = [T_n - d,\ T_n + d] \tag{3.4.2}$$

as a possible confidence interval for θ. Its length is $2d$ and if σ^2 is known the best fixed sample size which minimizes the length can be obtained from (3.4.1) and (3.4.2) which is given by (cf. Chow and Robbins (1965)).

$$n_d = d^{-2}\ Z^2_{1-\alpha/2}\,\sigma^2. \tag{3.4.3}$$

and

$$\lim_{d\to 0} P_\theta(\theta \in I_{n_d}) = 1\text{-}\alpha.$$

For small d, I_{nd} provides a bounded length confidence interval for θ with asymptotic convergence probability $1\text{-}\alpha$.

However, when σ^2 is unknown, σ^2 in (3.4.3) it can be replaced by an estimator

$$S^2_n = S_n(X_1 X_2 \ldots X_n),$$

but then we cannot use the above fixed sample size procedure. So we replace n_d by a random sample or a stopping rule

$$N_d = \min\{n \geq n_0 : n \geq d^{-2}\ Z^2_{1-\alpha/2}\ S^2_n\},$$

where, n_0 is an initial sample size. Then we use the confidence interval

$$I_{N_d} = \{T_{Nd} - d, T_{Nd} + d\}$$

for estimating θ.

For the stopping rule N_d and the interval I_{N_d} Chow and Robbins (1965) have proved the following properties.

1. N_d is non decreasing in d
2. N_d is finite with probability one for every d>0
3. $N_d\, d \to 1$ as $d \to 0$
4. $\lim\limits_{n\to\infty} P_\theta(\theta \in I_{Nd}) = 1\text{-}\alpha.$

Our problem here in this section is to find a confidence interval for μ = E(X_i) for the model (2.1.1) having prescribed width $2d$ and a converge probability 1-α. That is to find I_{N_d} such that P[$\mu \in I_{N_d}$] = 1-α, $0 \le \alpha < 1$. We have proved in Theorem 2.3.1 [See (2.3.2)] that

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} \mathrm{N}(0, \sigma_1^2),$$

where σ_1^2 is as defined in (2.3.3) and

$$\sigma_1^2 = \frac{\sigma^2}{n} B(k).$$

$$B(k) = 1+2 \sum_{j=1}^{\infty} c(k^j).$$

Based on the above result an approximate confidence interval for μ, when σ^2 and k are known are constructed in Section 2.3.

Let I_{K_0} be the required confidence interval. Then

$$I_{K_0} = [\ \overline{X}_{K_0} - d,\ \overline{X}_{K_0} + d]$$

where

$$K_0 = [d^{-2}\ Z^2_{1-\alpha/2}\ \sigma^2 B(k)] \tag{3.4.4}$$

and $Z_{1-\alpha/2}$ is such that

$$\frac{1}{\sqrt{2\pi}} \int_{Z_{1-\alpha/2}}^{Z_{1-\alpha/2}} \exp\{-\tfrac{u^2}{2}\}\ du = 1\text{-}\alpha.$$

Note that from (3.4.4) that $K_0 \to \infty$ when $d \to 0$ and

$$\mathrm{P}[\mu \in I_{K_0}] = \mathrm{P}\left[\sqrt{K_0}\left|\overline{X}_{K_0} - \mu\right| \le \frac{d\sqrt{K_0}}{\Delta}\right] \to 1\text{-}\alpha$$

where

$$\Delta = [\sigma^2 B(k)]^{1/2} \tag{3.4.5}$$

When at least one of the parameters σ^2, k is unknown we proposes a sequential confidence interval. For that we define a stopping rule as in the case of point estimation,

$$N = \inf\{n \ge n_0 \colon n \ge d^{-2} Z^2_{1-\alpha/2}[S^2_n B(\widetilde{k}_n) + n^{-h}]\}, \tag{3.4.6}$$

where

n_0 is an initial sample size

$$S^2_n = n^{-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

$\widetilde{k}_n$ is as defined by (2.4.2) and h is a suitable constant to be defined later. Note that from the above definition of stopping rule $N \ge d^{-2}\ Z^2_{1-\alpha/2} . N^{-h}$

That is $N \ge \left(\dfrac{Z_{1-\alpha/2}}{d}\right)^{2/(1+h)}$

Thus when $d \to 0$, $N \to \infty$.

The performance of the above stopping time N and the corresponding confidence interval I_N are discussed in the following Theorem.

Theorem 3.4.1 If $\mathrm{E}|Z_l|^{2p} < \infty$ for $p>2$ and $h \in (0,p\text{-}2)$ then as $d \to 0$

i. $\dfrac{N}{K_0} \xrightarrow{a.s} 1$

ii. $\mathrm{E}\left(\dfrac{N}{K_0}\right) \to 1$

iii. $\mathrm{P}[\mu \in I_N] \to 1\text{-}\alpha$.

The following lemmas are needed to prove this theorem and hence we prove them first.

Lemma 3.4.1. If $\mathrm{E}|Z_l|^{2p} < \infty$ for $p>2$ then

$$\mathrm{P}[\,|S_n^2 A(\tilde{k}_n) - \sigma^2 A(k)| > \varepsilon] = o(n^{-p/2}).$$

Proof: We have proved in Section 3.2 that

$$\mathrm{P}[\,|A(\tilde{k}_n) - A(k)| > \varepsilon] = \mathrm{O}(n^{-p}),$$

which implies

$$\mathrm{P}[\,|A(\tilde{k}_n) - A(k)| > \varepsilon] = \mathrm{O}(n^{-p/2}).$$

In view of the Result 1.6.12 here it is enough to prove

$$\mathrm{P}[\,|S_n^2 - \sigma^2| > \varepsilon] = \mathrm{O}(n^{-p/2}). \tag{3.4.7}$$

As for (3.4.7) consider

$$\left\|S_n^2 - \sigma^2\right\|_p = \left\|n^{-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 - \sigma^2\right\|_p$$

$$= \left\|n^{-1}\sum_{i=1}^{n}(X_i - \mu)^2 + 2(\mu - \overline{X}_n)n^{-1}\sum_{i=1}^{n}(X_i - \mu) + (\mu - \overline{X}_n)^2 - \sigma^2\right\|_p$$

Now using (3.2.5), Minkoswski inequality and Schwarz inequality we can write

$$\left\|S_n^2 - \sigma^2\right\|_p \le \Big\|n^{-1}k^2\sum_{i=1}^{n}(Z_i-\alpha)^2 + 2kn^{-1}(k\alpha-\mu)\sum_{i=1}^{n}(Z_i-\alpha) + (k\alpha-\mu)^2$$

$$+\ 2kn^{-1}(\mu-\overline{X}_n)\sum_{i=1}^{n}(Z_i-\alpha) + 2(\mu-\overline{X}_n)\,(k\alpha-\mu) + (\mu-\overline{X}_n)^2 - \sigma^2\Big\|_p$$

$$\le \left\|n^{-1}k^2\sum_{i=1}^{n}[(Z_i-\alpha)^2-\theta]\right\|_p + 2n^{-1}(k\alpha-\mu)k\left\|\sum_{i=1}^{n}(Z_i-\alpha)\right\|_p$$

$$+\ 2n^{-1}k\left\|\sum_{i=1}^{n}(Z_i-\alpha)\right\|_{2p}\left\|(\mu-\overline{X}_n)\right\|_{2p} + 2(k\alpha-\mu)\left\|(\mu-\overline{X}_n)\right\|_p$$

$$+\left\|(\mu-\overline{X}_n)^2\right\|_p + \left\|k^2\theta-\sigma^2\right\|_p, \qquad (3.4.8)$$

where θ=E(Z_i - α)^2.

Note that $\sum_{i=1}^{n}[(Z_i-\alpha)^2-\theta]$ is a mean zero martinagle w.r.t $\boldsymbol{F}_n = \sigma\{X_0, Z_1, \ldots, Z_n\}$ and hence from M-Z inequality the first term in (3.4.8) can be calculated as

$$\left\|n^{-1}k^2\sum_{i=1}^{n}[(Z_i-\alpha)^2-\theta]\right\|_p = O(n^{-1/2})$$

Now by applyaing Schwartz inequality, Lemma 3.2.1 and M-Z inequality each term in (3.4.8) can shown to be of $O(n^{-1/2})$. Thus we have

$$\left\|S_n^2-\sigma^2\right\|_p = O(n^{-1/2}).$$

Now from Markov inequality

$$P[\,|\,S_n^2 - \sigma^2\,| > \varepsilon] \le \frac{E\left|S_n^2-\sigma^2\right|^p}{\varepsilon^p}$$

$$= O(n^{-p/2}).$$

Hence we have the required result (3.4.7). The lemma is proved.

Proof of the following lemma is omitted as it is similar to that of Lemma 3.2.2.

Lemma 3.4.2 If E$\,|\,Z_1\,|^{2p} < \infty$ for p>2 and $h\in(0,p-2)$ then

i. $P[N < K_d(1-\varepsilon)] = O\left(d^{-\frac{p-2}{2(1+h)}}\right)$

ii. $$\sum_{n \geq K_0(1+\varepsilon)} P[N > n] = O\left(d^{\frac{p-2}{2}}\right).$$

Proof of Theorem 3.4.1

We can prove part (i) and part (ii) using Lemmas 3.2.1 and 3.4.2. The proof is parallel to that of part (i) and part (ii) of Theorem 3.2.1. Hence we omitt the details.

For **part (iii)** consider,

$$\begin{aligned} \mathrm{P}[\mu \in I_N] &= \mathrm{P}\left[\left|\overline{X}_N - \mu\right| \leq d\right] \\ &= \mathrm{P}\left[\frac{\sqrt{N}\left|\overline{X}_N - \mu\right|}{\Delta} \leq \frac{d\sqrt{N}}{\Delta}\right] \\ &= \mathrm{P}\left[\frac{\sqrt{N}\left|\overline{X}_N - \mu\right|}{\Delta} \leq \frac{Z_{1-\frac{\alpha}{2}}\sqrt{N}}{\sqrt{K_0}}\right], \end{aligned} \tag{3.4.9}$$

where Δ is as defined in (3.4.5).

Recall from Section 3.2.1,

$$\sqrt{N}(\overline{X}_N - \mu) \xrightarrow{d} \mathrm{N}(0,\ \Delta^2).$$

That is,

$$\Delta^{-1}\sqrt{N}(\overline{X}_N - \mu) \xrightarrow{d} \mathrm{N}(0,1).$$

Also we have from part (i) of Theorem 3.4.1,

$$\frac{N}{K_0} \xrightarrow{a.s} 1$$

and hence $\sqrt{\dfrac{N}{K_0}} \xrightarrow{a.s} 1.$

Now part (iii) follows from (3.4.9) and the above arguement.

This completes the proof of the Theorem.

CHAPTER 4

ESTIMATION IN RANDOM COEFFICIENT AUTOREGRESSIVE MODEL

4.1 Introduction

The rest of this thesis is about ,sequential estimation of first order random coefficient autoregressive model RCAR(1). Linear time series models such as autoregressive models have been widely and successfully used in many fields. The reasons are that these models can be easily analysed and they provide fairly good approximations for the underlying chance mechanisms of numerous real life time-series. However, in some particular situations one may ask if there exist other models which can provide better fits. One is then led to consider nonstationary or nonlinear models. RCAR model is one such class of nonlinear models which have been found useful in many areas. Some of the specific applications of RCAR(1) models are discribed in Section 1.2. In the present chapter we consider the properties of RCAR(1) model and properties of least squares estimators of its parameters.

4.2 The Model and its Properties

Let $\{X_i\}$ be a sequence of r.v.s defined by an RCAR(1) model

$$X_i - \mu = (b + \beta_i)(X_{i-1} - \mu) + \varepsilon_i, \; i = 1,2, \ldots \qquad (4.2.1)$$

where $\mu = E(X_i)$ and the r.v.s satisfies the assumptions A_1 - A_4 in section 1.2 with p=1. They are

a$_1$) $\{\varepsilon_i, i = \pm 1, \pm 2, \ldots\}$ is a sequence of iid r.vs with mean zero and variance $\sigma^2 < \infty$

a₂) $\{\beta_i, i = \pm 1, \pm 2, \ldots\}$ is a sequence of iid r.vs with mean zero and variance $\gamma < \infty$

a₃) The sequence $\{\varepsilon_i\}$ and $\{\beta_i\}$ are statistically independent.

a₄) X_i is independent of ε_j and β_j for $j > i$.

Recursively using (4.2.1) we can express $X_i - \mu$ as

$$X_i - \mu = V_i + W_i, \quad i = 1, 2, \ldots \tag{4.2.2}$$

where

$$V_i = \varepsilon_i + \sum_{j=1}^{m} \left[\prod_{k=0}^{j} (b + \beta_{i-k}) \right] \varepsilon_{i-j}, \text{ for any } m \tag{4.2.3}$$

and

$$W_i = \left[\prod_{k=0}^{m} (b + \beta_{i-k}) \right] \left(X_{i-(m+1)} - \mu \right) \tag{4.2.4}$$

Here $\{V_i\}$ defined in (4.2.3) is an $(m+1)$ dependent stationary process. [See Definition 1.6.6].

In the following we discuss the conditions required for the stationarity of $\{X_i\}$. Using (4.2.2) we can write

$$(X_i - \mu) - V_i = W_i.$$

Thus

$$\mathrm{E}[(X_i - \mu) - V_i]^2 = \mathrm{E}[W_i^2]. \tag{4.2.5}$$

Now using the assumptions for the model (4.2.1) and (4.2.4)

$$\mathrm{E}[W_i^2] = \mathrm{E}\left[\prod_{k=0}^{m} (b + \beta_{i-k}) \right]^2 \mathrm{E}\left(X_{i-m-1} - \mu \right)^2.$$

$$= \prod_{k=0}^{m} (b^2 + \gamma) \ \mathrm{E}\left(X_{t-m-1} - \mu\right)^2$$

$$= (b^2 + \gamma)^{m+1} \ \mathrm{E}\left(X_{t-m-1} - \mu\right)^2$$

Now if $(b^2 + \gamma) < 1$ and $\mathrm{E}\left(X_{t-m-1} - \mu\right)^2 < \infty$, then as $m \to \infty$ $\mathrm{E}[W_t^2]$ converges to zero. Hence from (4.2.5) and from definition of convergence in mean square it follows that X_t - μ converge in mean square and, hence in probability to V_t. Thus we have

$$\left| (X_t - \mu) - V_t \right| \xrightarrow{p} 0 . \tag{4.2.6}$$

Thus there exist a solution for the model (4.2.1) if $(b^2 + \gamma) < 1$. The solution is given by

$$X_t - \mu = \varepsilon_t + \sum_{j=1}^{\infty} \left[\prod_{k=0}^{j} (b + \beta_{t-k}) \right] \varepsilon_{t-j} . \tag{4.2.7}$$

The solution for X_t - μ defined by (4.2.7) contains only iid r.v.s ε_t's and β_t's. Hence this solution is stationary also. Nicholls and Quinn (1982) proved that the solution to X_t - μ defined in (4.2.7) is ergodic. [See Theorem 2.7 of Nicholls and Quinn, (1982)].

We have noted in (4.2.6) that $\left| (X_t - \mu) - V_t \right| \xrightarrow{p} 0$. Now asymptotic properties of $(X_t - \mu)$ is same as that of V_t. [See Rao (1973), pp. 122]. Moreover, the asymptotic distribution of $\sqrt{n}(\overline{X}_n - \mu)$ is also same as that of $\sqrt{n}(\overline{V}_n - \mu_v)$ where μ_v is the mean of V_t and $\overline{V}_n = n^{-1} \sum_{t=1}^{n} V_t$. Since $\mathrm{E}(V_t^2)$ is finite, and V_t is $(m+1)$-dependent r.v.s, we have for fixed m [See Result 1.6.13],

$$\sqrt{n}(\overline{V}_n - \mu_v) \xrightarrow{d} N\left(0, \sum_{h=-m}^{m} Cov(V_t, V_{t+h})\right) . \tag{4.2.8}$$

Also as $m \to \infty$, V_i converges to $(X_i - \mu)$ in mean square. Hence $Cov(V_i, V_{i+h})$ also converges to $Cov(X_i - \mu, X_{i+h} - \mu)$ [cf. Rohatgi(1976), pp.248]. Thus as $m \to \infty$ the variance of the asymptotic distribution in (4.2.8) converges to

$$\sum_{h=-\infty}^{\infty} Cov(X_i - \mu, X_{i+h} - \mu).$$

That is,

$$\text{A.V.}[\ \sqrt{n}(\bar{V}_n - \mu_v)\ ] = \sum_{h=-\infty}^{\infty} Cov(X_i - \mu, X_{i+h} - \mu). \tag{4.2.9}$$

For the sequence defined by (4.2.1) using (4.2.7) and assumption on the model we have

$$\text{V}(X_i) = \text{E}\,(X_i - \mu)^2 = \text{E}\left[\varepsilon_i + \sum_{j=1}^{\infty}\left[\prod_{k=0}^{j}(b+\beta_{i-k})\right]\varepsilon_{i-j}\right]^2$$

$$= \text{E}[\varepsilon_i^2] + \text{E}[b+\beta_i]^2\,\text{E}[\varepsilon_{i-1}^2] + \text{E}[b+\beta_i]^2\,\text{E}[b+\beta_{i-1}]^2\,\text{E}[\varepsilon_{i-2}^2] + \ldots$$

$$+\ldots+ 2\,\text{E}[b+\beta_i]\,\text{E}(\varepsilon_i)\,\text{E}(\varepsilon_{i-1}) + \ldots$$

$$= \sigma^2 + (b^2+\gamma)\,\sigma^2 + (b^2+\gamma)^2\,\sigma^2 + \ldots$$

$$\frac{\sigma^2}{1-(b^2+\gamma)} = \text{V (say)}. \tag{4.2.10}$$

V = Var(X_i) defined in (4.2.10) is finite if E[ε_i^2] = $\sigma^2 < \infty$ and $b^2 + \gamma < 1$.

Now consider

$$r(h) = Cov(X_i, X_{i+h}).$$

Using the assumptions on the model we can write

$$X_{i+h} - \mu = \varepsilon_{i+h} + (b+\beta_{i+h})\,\varepsilon_{i+h-1} + \ldots + \prod_{k=1}^{h}(b+\beta_{i+h-k})(X_i - \mu).$$

Thus

$$r(h) = \mathrm{E}[\,(X_i - \mu)(X_{i+h} - \mu)\,]$$

$$= \mathrm{E}[\,\varepsilon_{i+h}(X_i - \mu)\,] + \mathrm{E}[\,(b + \beta_{i+h})\ \varepsilon_{i+h-1}\,(X_i - \mu)\,] + \ldots$$

$$+\ \mathrm{E}\left[\prod_{k=1}^{h}(b + \beta_{i+h-k})(X_i - \mu)^2\right]$$

$$= \mathrm{E}[\,\varepsilon_{i+h}\,]\mathrm{E}(X_i - \mu) + \ \mathrm{E}[\,(b + \beta_{i+h})\,]\mathrm{E}[\,\varepsilon_{i+h-1}\,]\ \mathrm{E}(X_i - \mu) + \ldots$$

$$+\ \mathrm{E}\left[\prod_{k=1}^{h}(b + \beta_{i+h-k})\right]\mathrm{E}\ (X_i - \mu)^2 \ = b^h\ \mathrm{Var}(X_i)$$

On similar lines

$$r(-h) = Cov(X_i, X_{i-h}) = b^h\ \mathrm{Var}(X_i) = r(h) \tag{4.2.11}$$

Now using (4.2.9), (4.2.10) and (4.2.11)

$$\mathrm{A.V}[\,\sqrt{n}(\overline{V}_n - \mu_v)\,] = \sum_{h=-\infty}^{\infty} r(h)$$

$$= \frac{\sigma^2}{1 - (b^2 + \gamma)}\ [\ 1 + 2\ (b + b^2 + \ldots)]$$

$$= \frac{\sigma^2}{1 - (b^2 + \gamma)}\left[1 + \frac{2b}{1 - b}\right]$$

$$= \frac{\sigma^2}{1 - (b^2 + \gamma)}\,\frac{1 + b}{1 - b}. \tag{4.2.12}$$

From (4.2.8) and from the above discussion we have

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{\ d\ } N\left(0,\ \frac{\sigma^2}{1 - (b^2 + \gamma)}\,\frac{1 + b}{1 - b}\right). \tag{4.2.13}$$

The next section deals with least squares estimation of parameters of RCAR(1) model defined in (4.2.1).

4.3 Properties of Least Square Estimators

The main objective of estimating the unknown parameters of a stationary time series $\{X_n\}$ is to provide predictors of X_n given the past values of the process. The least squares estimators are those estimators which minimize the sum of squares of errors Random coefficient autoregressive process are nonlinear in nature with two error components. Thus the least squares estimation procedure adopted here is a two step procedure. Many researchers have suggested estimators for regression parameter b in the model (4.2.1) that are efficient in the presence of nuisance parameters. For example see Koul and Schick (1996) and Schick (1996). The least squares estimators for b, σ^2 and γ obtained by Nicholls and Quinn (1982) are given below.

Assuming $\mu = E(X_t)$ is known,

$$\hat{b}_n = \sum_{t=1}^{n}(X_t - \mu)(X_{t-1} - \mu) / \sum_{t=1}^{n}(X_{t-1} - \mu)^2$$

$$\hat{\gamma}_n = \sum_{i=1}^{n}\hat{U}_i^2(Z_i - \bar{Z}) / \sum_{i=1}^{n}(Z_i - \bar{Z})^2$$

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{U}_t^2 - \hat{\gamma}_n\bar{Z}$$

where,

$$U_t = \beta_t(X_t - \mu) + \varepsilon_t = (X_t - \mu) - b(X_{t-1} - \mu)$$

$$\hat{U}_t = (X_t - \mu) - \hat{b}_n(X_{t-1} - \mu)$$

$$Z_t = (X_t - \mu)^2$$

$$\bar{Z} = n^{-1}\sum_{i=1}^{n} Z_i \text{ ,}$$

when μ is unknown the above estimators can be modified as

$$\hat{b}_n = \sum_{i=1}^{n}(X_i - \bar{X}_n)(X_{i-1} - \bar{X}^*_{n-1}) / \sum_{i=1}^{n}(X_{i-1} - \bar{X}^*_{n-1})^2 \qquad (4.3.1)$$

$$\hat{\gamma}_n = \sum_{i=1}^{n}\hat{U}_i^2(Z_i - \bar{Z}) / \sum_{i=1}^{n}(Z_i - \bar{Z})^2 \qquad (4.3.2)$$

$$\hat{\sigma}_n^2 = n^{-1}\sum_{i=1}^{n}\hat{U}_i^2 - \hat{\gamma}_n\bar{Z} \qquad (4.3.3)$$

where,

$$\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i \text{ ,} \qquad \bar{X}^*_{n-1} = n^{-1}\sum_{i=1}^{n} X_{i-1}$$

$$U_i = \beta_i\,(X_i - \bar{X}_n) + \varepsilon_i = (X_i - \bar{X}_n) - b(X_{i-1} - \bar{X}^*_{n-1})$$

$$\hat{U}_i = (X_i - \bar{X}_n) - \hat{b}_n(X_{i-1} - \bar{X}^*_{n-1}) \qquad (4.3.4)$$

$$Z_i = (X_i - \bar{X}^*_{n-1})^2$$

$$\bar{Z} = n^{-1}\sum_{i=1}^{n} Z_i \text{ .}$$

Some of the properties of these estimators useful in sequential analysis are discussed in the following lemmas.

Lemma 4.3.1: If $\mathrm{E}\left|\varepsilon_1^{4s}\right| < \infty$ and $\|b + \beta_1\|_{4s} < 1,\ s \geq 1,$ then $\|\bar{X}_n - \mu\|_{2p} = \mathrm{O}(n^{-1/2})$.

Proof: Assume without loss of generality $\mu = 0$. Then using (4.2.7)

$$\|\bar{X}_n\|_{4s} = \left\| n^{-1}\sum_{i=1}^{n}\left\{\sum_{j=1}^{\infty}\left[\prod_{k=0}^{j}(b + \beta_{i-k})\right]\varepsilon_{i-j} + \varepsilon_i\right\}\right\| .$$

Interchanging the order of summation and using Minkowski inequality we have

$$\left\|\bar{X}_n\right\|_{4s} \le n^{-1}\sum_{j=1}^{\infty}\left\|\sum_{i=1}^{n}\left[\prod_{k=0}^{j}(b+\beta_{i-k})\right]\varepsilon_{i-j}\right\|_{4s} + n^{-1}\left\|\sum_{i=1}^{n}\varepsilon_i\right\|_{4s} \tag{4.3.5}$$

By the Marcinkiewicz-Zygmund (M-Z) inequality (See Result 1.6.4),

$$\left\|\sum_{i=1}^{n}\left[\prod_{k=0}^{j}(b+\beta_{i-k})\right]\varepsilon_{i-j}\right\|_{4s} = O(n^{1/2})$$

and

$$\left\|\sum_{i=1}^{n}\varepsilon_i\right\|_{4s} = O(n^{1/2}).$$

Hence the required result.

The next lemma deals with p[th] moment convergence of $\hat{b}_n$.

Lemma 4.3.2: If $\mathrm{E}|\varepsilon_1|^{2p} < \infty$ and $\mathrm{E}|b+\beta_1|^{2p} < 1, p \ge 1$ then $P[|\hat{b}_n - b| > \varepsilon] = O(n^{-1/2})$.

Proof: When $\mu = \mathrm{E}(X_i)$ is known the estimator of b is given by

$$\hat{b}_n = \sum_{i=1}^{n}(X_i - \mu)(X_{i-1} - \mu)/\sum_{i=1}^{n}(X_{i-1} - \mu)^2$$

Using (4.2.1) $\hat{b}_n$ can be written as

$$\hat{b}_n = \sum_{i=1}^{n}[(b+\beta_i)(X_{i-1} - \mu) + \varepsilon_i]/\sum_{i=1}^{n}(X_{i-1} - \mu)^2$$

$$= \frac{b\sum_{i=1}^{n}(X_{i-1} - \mu)^2 + \sum_{i=1}^{n}\beta_i(X_{i-1} - \mu)^2 + \sum_{i=1}^{n}\varepsilon_i(X_{i-1} - \mu)^2}{\sum_{i=1}^{n}(X_{i-1} - \mu)^2}$$

$$= \frac{\sum_{i=1}^{n}\beta_i(X_{i-1}-\mu)^2 + \sum_{i=1}^{n}\varepsilon_i(X_{i-1}-\mu)^2}{\sum_{i=1}^{n}(X_{i-1}-\mu)^2} + b$$

Thus

$$\hat{b}_n\text{-}b = \frac{\sum_{i=1}^{n}\beta_i(X_{i-1}-\mu)^2 + \sum_{i=1}^{n}\varepsilon_i(X_{i-1}-\mu)^2}{\sum_{i=1}^{n}(X_{i-1}-\mu)^2}.$$

If μ is unknown the estimator for μ is $\overline{X}^*_{n-1} = n^{-1}\sum_{i=0}^{n-1} X_i$. Then

$$\hat{b}_n\text{-}b = \frac{n^{-1}J_n}{n^{-1}K_n}, \tag{4.3.6}$$

where,

$$J_n = \sum_{i=1}^{n}\beta_i(X_{i-1}-\overline{X}^*_{n-1})^2 + \sum_{i=1}^{n}\varepsilon_i(X_{i-1}-\overline{X}^*_{n-1})^2 \tag{4.3.7}$$

$$K_n = \sum_{i=1}^{n}(X_{i-1}-\overline{X}^*_{n-1})^2. \tag{4.3.8}$$

The ergodic theorem (See Result 1.6.15) for $\{X_i\}$ implies that as $n\to\infty$, $\overline{X}^*_{n-1}\to\mu$ a.s. The model (4.2.1) and the assumptions on that immediately imply that the sequences $\{\beta_i(X_{i-1}-\overline{X}^*_{n-1})^2\}$, $\{\varepsilon_i(X_{i-1}-\overline{X}^*_{n-1})^2\}$ and $(X_{i-1}-\overline{X}^*_{n-1})^2$ are stationary and ergodic [See Remark 1.6.1]. Now applying Ergodic Theorem for these sequences, it follows that $n^{-1} J_n\to 0$ and $n^{-1} K_n \to V$ a.s as well as in the p[th] moment, where V is defined in (4.2.10). As a consequence we have $(\hat{b}_n\text{-}b)\to 0$ a.s as $n\to\infty$.

Next we calculate $\|n^{-1}J_n\|_p$ and $\|n^{-1}K_n - V\|_p$.

Consider J_n defined in (4.3.7) and by some algebraic manipulations we can write

$$n^{-1} J_n = J_{n_1} + J_{n_2} + J_{n_3} + J_{n_4} + J_{n_5},$$

where

$$J_{n_1} = n^{-1} \sum_{i=1}^{n} \beta_i (X_{i-1} - \mu)^2$$

$$J_{n_2} = n^{-1} (\mu - \bar{X}^*_{n-1})^2 \sum_{i=1}^{n} \beta_i$$

$$J_{n_3} = n^{-1} \sum_{i=1}^{n} \varepsilon_i (X_{i-1} - \mu)^2 \tag{4.3.9}$$

$$J_{n_4} = 2(\mu - \bar{X}^*_{n-1}) \; n^{-1} \sum_{i=1}^{n} \beta_i (X_{i-1} - \mu)$$

$$J_{n_5} = n^{-1} (\mu - \bar{X}^*_{n-1}) \sum_{i=1}^{n} \varepsilon_i .$$

We write

$$P[|J_n|>\varepsilon] \le P[|J_{n_1}|>\varepsilon/5] + P[|J_{n_2}|>\varepsilon/5] + P[|J_{n_3}|>\varepsilon/5] + P[|J_{n_4}|>\varepsilon/5] + P[|J_{n_5}|>\varepsilon/5]. \tag{4.3.10}$$

Using Markov inequality we can write

$$P[|J_{n_1}|>\varepsilon/5] \le \frac{E|J_{n_1}|^p}{(\varepsilon/5)^p}. \tag{4.3.11}$$

If we define F_n as the σ-field induced by $\{(\beta_k, \varepsilon_k), k \le n\}$ then

$$E[\sum_{i=1}^{n} \beta_i (X_{i-1} - \mu)^2 \,|\, F_{n-1}] = \sum_{i=1}^{n-1} \beta_i (X_{i-1} - \mu)^2 + E[\beta_n (X_{n-1} - \mu)^2]$$

$$= \sum_{i=1}^{n-1} \beta_i (X_{i-1} - \mu)^2 .$$

Since β_n is independent of X_j for $j<n$ and $E(\beta_i) = 0$. Thus $\{\sum_{i=1}^{n} \beta_i (X_{i-1} - \mu)^2, \; n\ge1\}$ is a zero mean martingle w.r.t F_n. By using Burkholder inequality [See Result 1.6.11], moment inequalities, assumptions of the Lemma and independence of β_i and X_{i-1} we have

$$B_p^{-1} n^{-1/2} \left\| \sum_{i=1}^{n} \beta_i (X_i - \mu)^2 \right\|_p \le \left\| [n^{-1} \sum_{i=1}^{n} \beta_i^2 (X_i - \mu)^4]^{1/2} \right\|_p$$

$$\le \left\| n^{-1} \sum_{i=1}^{n} \left| \beta_i (X_i - \mu)^2 \right|^{1/p} \right\|_p$$

$$\le n^{-1} \sum_{i=1}^{n} \left\| \beta_i |(X_i - \mu)^2 \right\|_p$$

$$= O(1),$$

where $B_p = 18p^{3/2}(p-1)^{-1/2}$

Thus

$$\left\| \sum_{i=1}^{n} \beta_i (X_i - \mu)^2 \right\|_p = O(n^{-1/2}).$$

Hence from (4.3.11) it follows that

$$P[\,|\, J_{n_1} \,| > \varepsilon/5] = O(n^{-p/2}). \tag{4.3.12}$$

As for

$$P[\,|\, J_{n_2} \,| > \varepsilon/5] \le P\left(n^{-1/2} \left| \left(\sum_{i=1}^{n} \beta_i \right)^{1/2} \right| \left| \mu - \overline{X}^*_{n-1} \right| > \sqrt{\varepsilon/5} \right)$$

$$\le C.\, n^{-p/2}\ E\left\{ \left| \left(\sum_{i=1}^{n} \beta_i \right)^{p/2} \right| \left| \mu - \overline{X}^*_{n-1} \right|^p \right\}$$

$$= O(n^{-p/2}). \tag{4.3.13}$$

Where we used the moment inequality, Cauchy-Schwarz inequality, Lemma 4.3.1 with $s = p/2$ and the fact that $\left\|\sum_{i=1}^{n}\beta_i\right\|_p = O(n^{1/2})$. Note that $\sum_{i=1}^{n}\varepsilon_i(X_{i-1}-\mu)^2$ and $\sum_{i=1}^{n}\beta_i(X_{i-1}-\mu)$ are mean zero martingales w.r.t F_n. Now using similar arguments as in the case of J_{n_1} we can show that

$$P[\,|J_{n_3}| > \varepsilon/5] = O(n^{-p/2})$$

and

$$P[\,|J_{n_4}| > \varepsilon/5] = O(n^{-p/2}).$$

As for J_{n_5} use Schwartz inequality, Lemma 4.3.1 and M-Z inequality to get

$$P[\,|J_{n_5}| > \varepsilon/5] = O(n^{-p/2}).$$

Hence from the above arguments and (4.3.10) we have for $\varepsilon>0$

$$P[\,|n^{-1}J_n| > \varepsilon] = O(n^{-p/2}). \tag{4.3.14}$$

By writing

$$n^{-1}K_n - V = n^{-1}\sum_{i=1}^{n}[(X_{i-1}-\mu)^2 + (\mu-\overline{X}^*_{n-1})^2 + 2(X_{i-1}-\mu)(\mu-\overline{X}^*_{n-1})$$

$$- E(X_{i-1}-\mu)^2\,]$$

and repeating the similar arguments as in the case of J_n we get

$$\left\|n^{-1}K_n - V\right\|_p = O(n^{-1/2})$$

and hence

$$P[\,|n^{-1} K_n - V\,| > \varepsilon\,] = O(n^{-p/2}). \tag{4.3.15}$$

Now Lemma 4.3.2 follows from (4.3.14), (4.3.15), (4.3.7) and Result 1.6.12. This completes the proof.

In the expression (4.3.2) replacing $\hat{U}_i$ by U_i we write

$$\tilde{\gamma}_n = \frac{n^{-1}\sum_{i=1}^{n} U_i^2 (Z_i - \bar{Z})}{n^{-1}\sum_{i=1}^{n} (Z_i - \bar{Z})^2} \tag{4.3.16}$$

Now $\tilde{\gamma}_n - \gamma$ can be written as

$$\tilde{\gamma}_n - \gamma = \frac{n^{-1}\sum_{i=1}^{n} U_i^2 (Z_i - \bar{Z})}{n^{-1}\sum_{i=1}^{n} (Z_i - \bar{Z})^2} - \gamma$$

$$= \frac{n^{-1}\sum_{i=1}^{n} [U_i^2 (Z_i - \bar{Z}) - \gamma (Z_i - \bar{Z})^2]}{n^{-1}\sum_{i=1}^{n} (Z_i - \bar{Z})^2}$$

$$= \frac{n^{-1}\sum_{i=1}^{n} (U_i^2 - Z_i\gamma)(Z_i - \bar{Z})}{n^{-1}\sum_{i=1}^{n} (Z_i - \bar{Z})^2} \tag{4.3.17}$$

since $\sum_{i=1}^{n} (Z_i - \bar{Z})\bar{Z} = 0$.

Define $\qquad \xi_i = U_i^2 \quad \sigma^2 - \gamma Z_i \qquad$ (4.3.18)

and write

$$\tilde{\gamma}_n - \gamma = n^{-1} T_n \;\; n^{-1} R_n, \tag{4.3.19}$$

where

$$T_n = \sum_{i=1}^{n}(Z_i - \bar{Z})\xi_i \qquad (4.3.20)$$

$$R_n = \sum_{i=1}^{n}(Z_i - \bar{Z})^2 \qquad (4.3.21)$$

Now by repeating the arguments used to prove lemma 4.3.2 we can prove that as $n \to \infty$, $n^{-1}T_n \to 0$ and $n^{-1}R_n \to R$ a.s and in pth moment, where $R = \text{Var}(Z_i) < \infty$. Hence from (4.3.19) we have $\tilde{\gamma}_n - \gamma \to 0$.

Let us write

$$\hat{\gamma}_n - \tilde{\gamma}_n = \frac{n^{-1}\sum_{i=1}^{n}(\hat{U}_i^2 - U_i^2)(Z_i - \bar{Z})}{n^{-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2}$$

$$= \frac{n^{-1}\sum_{i=1}^{n}(Z_i - \bar{Z})(\hat{U}_i - U_i)(\hat{U}_i + U_i)}{n^{-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2}$$

$$= \frac{n^{-1}\sum_{i=1}^{n}(Z_i - \bar{Z})(b - b_n)(X_{i-1} - \bar{X}^*_{n-1})\{2U_i + (b - b_n)(X_{i-1} - \bar{X}^*_{n-1})\}}{n^{-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2}$$

$$= n^{-1}H_n / n^{-1}R_n\,, \qquad (4.3.22)$$

where,

$$H_n = \sum_{i=1}^{n}(Z_i - \bar{Z})(b - b_n)(X_{i-1} - \bar{X}^*_{n-1})\{2U_i + (b - b_n)(X_{i-1} - \bar{X}^*_{n-1})\} \qquad (4.3.23)$$

and R_n is as defined by (4.3.21). Using similar arguments as before we can prove $\hat{\gamma}_n - \tilde{\gamma}_n \to 0$ a.s.

Lemma 4.3.3: Under the conditions of Lemma 4.3.2

$$P[\,|\,\hat{\gamma}_n - \gamma\,| > \varepsilon] = O(n^{-p/2}).$$

Proof: We write

$$P[\,|\,\hat{\gamma}_n - \gamma\,| > \varepsilon] \le P[\,|\,\hat{\gamma}_n - \tilde{\gamma}_n\,| > \varepsilon/2\,] + P[\,|\,\tilde{\gamma}_n - \gamma\,| > \varepsilon/2\,]. \tag{4.3.24}$$

Now let us calculate $P[\,|\,\hat{\gamma}_n - \tilde{\gamma}_n\,| > \varepsilon/2\,]$ and $P[\,|\,\tilde{\gamma}_n - \gamma\,| > \varepsilon/2\,]$.

Using Minkowski inequality and Schwartz inequality in (4.3.20) we can write

$$n^{-1}\,\|T_n\|_p \le \left\|n^{-1}\sum_{i=1}^{n} Z_i\xi_i\right\|_p + \|\overline{Z}\|_{2p}\left\|n^{-1}\sum_{i=1}^{n}\xi_i\right\|_{2p} \tag{4.3.25}$$

Now using (4.3.18) and (4.3.4) each term in (4.3.25) can be written as

$$\left\|n^{-1}\sum_{i=1}^{n} Z_i\xi_i\right\|_p = \Big\|\; n^{-1}\sum_{i=1}^{n}[\beta_i^2(X_{i-1} - \overline{X}^*_{n-1})^4 + \varepsilon_i(X_{i-1} - \overline{X}^*_{n-1})$$

$$+ 2\beta_i\varepsilon_i(X_{i-1} - \overline{X}^*_{n-1})^3 - (X_{i-1} - \overline{X}^*_{n-1})^2\sigma^2 - (X_{i-1} - \overline{X}^*_{n-1})^4\gamma \;\Big\|_p$$

$$\|\overline{Z}\|_{2p} = \Big\|\; n^{-1}\sum_{i=1}^{n}[(X_{i-1} - \mu)^2 + (\mu - \overline{X}^*_{n-1})^2 + 2(X_{i-1} - \mu)(\mu - \overline{X}^*_{n-1}) \;\Big\|_{2p}$$

$$\left\|n^{-1}\sum_{i=1}^{n}\xi_i\right\|_{2p} = \Big\|\; n^{-1}\sum_{i=1}^{n}[\beta_i^2(X_{i-1} - \overline{X}^*_{n-1})^2 + \varepsilon_i + 2\beta_i\varepsilon_i(X_{i-1} - \overline{X}^*_{n-1}) - \sigma^2$$

$$- (X_{i-1} - \overline{X}^*_{n-1})^2\gamma \;\Big\|_{2p}$$

From the proofs of earlier lemmas it follows that

$$n^{-1}\,\|T_n\|_p = O(n^{-1/2}).$$

Using Markov inequality for $\varepsilon>0$

$$P[\,|\,n^{-1}\,T_n - 0\,| > \varepsilon] = O(n^{-p/2}). \tag{4.3.26}$$

On similar lines it can be shown that

$$P[\,|\,n^{-1} R_n - R\,| > \varepsilon] = O(n^{-p/2}). \tag{4.3.27}$$

Hence from (4.3.19), (4.3.26), (4.3.27) and Result 1.6.12 we get

$$P[\,|\,\hat{\gamma}_n - \tilde{\gamma}_n\,| > \varepsilon/2\,] = O(n^{-p/2}). \tag{4.3.28}$$

H_n defined in (4.3.23) can be written as

$$H_n = \sum_{i=1}^{n} [2\beta_i (X_{i-1} - \overline{X}^{*}_{n-1})^4 (b - \hat{b}_n) + 2\varepsilon_i (X_{i-1} - \overline{X}^{*}_{n-1})^3 (b - \hat{b}_n)$$

$$+ (X_{i-1} - \overline{X}^{*}_{n-1})^3 (b - \hat{b}_n)^2 - 2\overline{Z}\beta_i (X_{i-1} - \overline{X}^{*}_{n-1})^2 (b - \hat{b}_n)$$

$$- 2\overline{Z}\varepsilon_i (X_{i-1} - \overline{X}^{*}_{n-1})(b - \hat{b}_n) - \overline{Z}(X_{i-1} - \overline{X}^{*}_{n-1})^2 (b - \hat{b}_n)^2].$$

Now repeating arguments in Lemma 4.3.2 here also we can show that

$$P[\,|\,n^{-1} H_n - 0\,| > \varepsilon] = O(n^{-p/2}). \tag{4.3.29}$$

Now (4.3.27), (4.3.29), (4.3.22) and Result 1.6.12 leads to

$$P[\,|\,\tilde{\gamma}_n - \gamma\,| > \varepsilon/2\,] = O(n^{-p/2}). \tag{4.3.30}$$

Application of (4.3.28) and (4.3.30) in (4.3.24) gives

$$P[\,|\,\hat{\gamma}_n - \gamma\,| > \varepsilon/2\,] = O(n^{-p/2}).$$

This completes the proof.

Lemma 4.3.4: Suppose that the conditions of Lemma 4.3.2 hold, then for $\varepsilon > 0$

$$P[\,|\,\hat{\sigma}^2_n - \sigma^2\,| > \varepsilon] = O(n^{-p/2}).$$

Proof: Let $\tilde{\sigma}_n^2$ be the expression for $\hat{\sigma}_n^2$ obtained by replacing U_i in the place of $\hat{U}_i$ in equation (4.3.3).

Consider

$$\tilde{\sigma}_n^2 - \sigma^2 = n^{-1}\sum_{i=1}^{n} U_i^2 - \tilde{\gamma}_n \overline{Z} - \sigma^2$$

and

$$\hat{\sigma}_n^2 - \tilde{\sigma}_n^2 = n^{-1}\sum_{i=1}^{n}(\hat{U}_i^2 - \tilde{U}_i^2) - (\hat{\gamma}_n - \tilde{\gamma}_n)\overline{Z}\,.$$

Now using lemma 4.3.3 and using similar arguments as before we can show that

$$P[\,|\,\tilde{\sigma}_n^2 - \sigma^2\,| > \varepsilon/2] = O(n^{-p/2}) \qquad (4.3.31)$$

and

$$P[\,|\,\hat{\sigma}_n^2 - \tilde{\sigma}_n^2\,| > \varepsilon/2] = O(n^{-p/2}). \qquad (4.3.32)$$

But

$$P[\,|\,\hat{\sigma}_n^2 - \sigma^2\,| > \varepsilon] \le P[\,|\,\tilde{\sigma}_n^2 - \sigma^2\,| > \varepsilon/2] + P[\,|\,\hat{\sigma}_n^2 - \tilde{\sigma}_n^2\,| > \varepsilon/2]. \qquad (4.3.33)$$

Now Lemma 4.3.4 follows from (4.3.31), (4.3.32) and (4.3.33).

This completes the proof of Lemma 4.3.4.

The results of this chapter will be used in the following chapters for studying the sequential estimation.

CHAPTER 5

SEQUENTIAL ESTIMATION OF THE MEAN OF RCAR(1) PROCESS

5.1 Introduction

There are two basic reasons why sequential methods are used in Statistics. Firstly, it is possible to reduce the sample size on an average as compared to corresponding fixed sample size procedure. Secondly to solve certain problems which cannot be solved by any procedure based on a predetermined sample size. Some of the examples to this effect are discussed in Section 1.5. The discussion in the present chapter focuses on the first aspects of the subject and deals in particular with Random Coefficient Autoregressive Processes of order one RCAR(1). The main problems discussed in this chapter are the sequential point estimation, and interval estimation. We have already discussed in detail the properties of this model in Chapter 4.

The problem of sequential estimation of the parameters of AR(1) model are studied by Sriram (1987, 1988). Recently Sriram's results have been extended to AR(p) model and linear processes by Fakhre-Zakeri and Lee (1992) and Lee (1992).

2 Sequential Point Estimation

We study the problem of sequential point estimation of mean of RCAR(1) ocess in this section

Since $\{X_i\ ,\ i \geq 0\}$ defined in (4.2.1) is a stationary and ergodic sequence, a tural estimator for $\mu = E(X_i)$ is the sample mean

$$\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i .$$

ippose that we want to estimate μ by $\bar{X}_n$ using the loss function

$$L_{n,\lambda} = A(\bar{X}_n - \mu)^2 + \lambda n, \qquad A, \lambda > 0 \tag{5.2.1}$$

here A is a known constant and λ is the cost per observation. The loss function defined by .2.1) is the weighted error plus cost of inspection. An approximate expression of the risk in be calculated using (4.2.13) and is given by

$$R_{n,\lambda} = E(L_{n,\lambda}) = AE(\bar{X}_n - \mu)^2 + \lambda n$$

$$\approx An^{-1} \frac{\sigma^2}{1-(b^2+\gamma)} \frac{1+b}{1-b} + \lambda n. \tag{5.2.2}$$

et n_0 be the value of n for which $R_{n,\lambda}$ is minimum. Treating n as a continuous variable, we fferentiate $R_{n,\lambda}$ w.r.t. n and obtaine n_0. Thus

$$\frac{\partial R_{n,\lambda}}{\partial n} = 0 \text{ implies}$$

$$-An^{-2} \frac{\sigma^2}{1-(b^2+\gamma)} \frac{1+b}{1-b} + \lambda = 0.$$

This gives

$$n_0 \approx A^{1/2}\,\lambda^{-1/2}\,\sigma[1-(b^2+\gamma)]^{-1/2}\left(\frac{1+b}{1-b}\right)^{1/2} \tag{5.2.3}$$

Clearely $\dfrac{\partial^2 R_{n,\lambda}}{\partial n^2}$ is positive at n_0.

We refer n_0 as the best fixed sample size procedure. The corresponding minimum value of risk $R_{n_0,\lambda}$ can be obtained from (5.2.2) and is given by

$$R_{n_0,\lambda} = 2\,\lambda\, n_0. \tag{5.2.4}$$

If at least one of the parameters b, σ^2, γ in (5.2.3) is unknown, there does not exist any best fixed sample size procedure that will achieve the minimum risk $R_{n_0,\lambda}$. As a remedy we go for sequential procedure to estimate μ by choosing a sample size such that the associated risk will be close to $R_{n_0,\lambda}$, as cost per observation becomes small. Towards this end we use the least squares estimators of b, σ^2 and γ. Properties of these estimators are already discussed in Section 4.3.

Let us define a stopping time T by

$$T = \inf\left\{\, n\geq m\colon\ n\geq A^{1/2}\,\lambda^{-1/2}\left[\sigma[1-(b^2+\gamma)]^{-1/2}\left(\frac{1+b}{1-b}\right)^{1/2} + n^{-h}\right]\right\}, \tag{5.2.5}$$

where m is an initial sample size, $h>0$ is a suitable constant to be defined later. Based on this stopping rule the sequential point estimator of μ is $\overline{X}_T$ and the associated risk is

$$R_{T,\lambda} = A\mathrm{E}(\overline{X}_T - \mu)^2 + \lambda\,\mathrm{E}(T). \tag{5.2.6}$$

The main theorem in this chapter is stated below. This theorem establishes the optimal properties of the sequential procedure for estimating μ using the stopping rule (5.2.5).

Theorem 5.2.1: For $p>2$, if $E|\varepsilon_1|^{2p} < \infty$, $E|b+\beta_1|^{2p}<1$ and $h \in (0, (p-2)/4)$

then as $\lambda \to 0$

i. $\dfrac{T}{n_0} \to 1$, a.s

ii. $E\left|\dfrac{T}{n_0} - 1\right| \to 0$

iii. $\dfrac{R_{T,\lambda}}{R_{n_0,\lambda}} \to 1$

iv. $\sqrt{T}(\bar{X}_T - \mu) \xrightarrow{d} N\left(0, \dfrac{\sigma^2}{1-(b^2+\gamma)}\dfrac{1+b}{1-b}\right)$.

The proof of this Theorem depends on some lemmas, which are proved below. The following notations are introduced for easy reference.

$$n_1 = \left(\tfrac{A}{\lambda}\right)^{1/2(1+h)} \quad n_2 = n_0(1-\varepsilon) \quad n_3 = n_0(1+\varepsilon),\ 0<\varepsilon<1,\ K = \left(\tfrac{A}{\lambda}\right)^{1/2},$$

$$E = [n_2 < T < n_3] \qquad B = [T \le n_2] \qquad C = [T \ge n_3]$$

I_F and F^C denote the indicator and complement of a set F respectively.

Lemma 5.2.1: Suppose that $E|\varepsilon_1^{2p}| < \infty$ and $E|b+\beta_1|^{2p}<1$ for $p>2$, then for every $\varepsilon > 0$

$$P\left[\left|\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}-\frac{\sigma^2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right|>\varepsilon\right]=O(n^{-p/2}).$$

Proof: Using Lemma 4.3.2, Lemma 4.3.3, Lemma 4.3.4 and Result 1.6.12 we can write

$$P\left[\left|[1-(\hat{b}_n+\hat{\gamma}_n)]-[1-(b+\gamma)]\right|>\varepsilon\right]=O(n^{-p/2})$$

$$P\left[\left|\frac{(1+\hat{b}_n)}{(1-\hat{b}_n)}-\frac{1+b}{1-b}\right|>\varepsilon\right]=O(n^{-p/2})$$

and

$$P[|\hat{\sigma}_n^2-\sigma^2|>\varepsilon]=O(n^{-p/2})$$

Once again use Result 1.6.12 to obtain

$$P\left[\left|\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}-\frac{\sigma^2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right|>\varepsilon\right]=O(n^{-p/2}).$$

This completes the proof of Lemma 5.2.1.

Lemma 5.2.2: Suppose that $E|\varepsilon_1^{2p}|<\infty$ and $E|b+\beta_1|^{2p}<1$ for $p>2$, then for every $\varepsilon>0$

i. $P[T\leq n_2]=O\left(\lambda^{\frac{p-2}{4(h+1)}}\right)$

and

ii. $\sum_{n\geq n_3} P[T>n]=O\left(\lambda^{\frac{p-2}{4}}\right).$

Proof: From the definition of stopping rule (5.2.5), we have

$$T\geq\left(\tfrac{A}{\lambda}\right)^{1/2}T^{-h}$$

That is

$$T \geq \left(\tfrac{A}{\lambda}\right)^{1/2(1+h)} = n_1. \tag{5.2.7}$$

Now from (5.2.5) and (5.2.7),

$$\mathrm{P}[T \leq n_2] \leq \mathrm{P}\Big[\left(\tfrac{A}{\lambda}\right)^{1/2}\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}\right]^{1/2} \leq n \text{ for some } n_1 < n \leq n_2\Big]$$

$$\leq \mathrm{P}\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n} \leq K^{-2}n_2^2 \ \textit{for some}\ n_1 < n \leq n_2\right]$$

$$\leq \mathrm{P}\left[\underset{n_1 \leq n \leq n_2}{Max}\left|\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\left(\frac{1+b}{1-b}\right) - \frac{\sigma^2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right| > \frac{\sigma^2(2-\varepsilon)\varepsilon}{1-(b^2+\gamma)}\left(\frac{1+b}{1-b}\right)\right]$$

$$\leq \sum_{n=n_1}^{\infty} \mathrm{P}\left[\left|\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\left(\frac{1+\hat{b}_n}{1-\hat{b}_n}\right) - \frac{\sigma^2}{1-(b^2+\gamma)}\left(\frac{1+b}{1-b}\right)\right| > \frac{\sigma^2(2-\varepsilon)\varepsilon}{1-(b^2+\gamma)}\left(\frac{1+b}{1-b}\right)\right].$$

Now from Lemma 5.2.1 we have

$$\mathrm{P}[T \leq n_2] = \mathrm{O}(n_1^{-(\frac{p}{2}-1)}) = \mathrm{O}\left(\lambda^{\frac{p-2}{4(h+1)}}\right).$$

This proves the first part of the Lemma.

For the second part, from the definition of T it follows that for $n \geq n_3$,

$$\mathrm{P}[T > n] = \mathrm{P}\Big[\left(\tfrac{A}{\lambda}\right)^{1/2}\left\{\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}\right]^{1/2} + n^{-h}\right\} > n\Big]$$

$$= \mathrm{P}\left\{\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}\right]^{1/2} > K^{-1}n - n^{-h}\right\}$$

$$= \mathrm{P}\left\{\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_{n})}\frac{1+\hat{b}_n}{1-\hat{b}_n}\right]^{1/2} - \left[\frac{\sigma 2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right] > K^{-1}(n_3-n_0)-n_3^{-h}\right\}.$$

But

$$K^{-1}(n_3-n_0)-n_3^{-h} = \frac{\varepsilon\sigma^2(1+b)}{[1-(\hat{b}_n^2+\hat{\gamma}_n)](1-b)}$$

$$- \left\{\left[\frac{[1-(\hat{b}_n^2+\hat{\gamma}_n)](1-b)}{\sigma^2(1+b)}\right]^{1/2}\left(\frac{\lambda}{a}\right)^{1/2}\frac{1}{1+\varepsilon}\right\}^h.$$

Choose λ small enough so that the above expression for $K^{-1}(n_3-n_0)-n_3^{-h}$ is greater than

$$\left[\frac{\sigma^2(1+b)}{4[1-(\hat{b}_n^2+\hat{\gamma}_n)](1-b)}\right]^{1/2}$$

Thus we can write

P[$T>n$]

$$\le \mathrm{P}\left\{\left[\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n}\right]^{\frac{1}{2}} - \left[\frac{\sigma^2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right]^{\frac{1}{2}} > \left[\frac{\sigma^2(1+b)}{4[1-(b^2+\gamma)(1-b)]}\right]^{\frac{1}{2}}\right\}$$

$$\le \mathrm{P}\left\{\left|\frac{\hat{\sigma}_n^2}{1-(\hat{b}_n^2+\hat{\gamma}_n)}\frac{1+\hat{b}_n}{1-\hat{b}_n} - \frac{\sigma^2}{1-(b^2+\gamma)}\frac{1+b}{1-b}\right| > \frac{\varepsilon^2\sigma^2(1+b)}{4[1-(b^2+\gamma)(1-b)]}\right\}.$$

Now using Lemma 5.2.1 and repeating the same argument as in the first part we get the result.

Proof of Theorem 5.2.1

In section 4.3 we have proved that as $n \to\infty$, $\hat{b}_n \xrightarrow{a.s} b$, $\hat{\sigma}_n^2 \xrightarrow{a.s} \sigma^2, \hat{\gamma}_n \xrightarrow{a.s} \gamma$. Also we have noted in (5.2.7) that

$$T \geq \left(\tfrac{A}{\lambda}\right)^{1/2(1+h)}$$

Thus $T\to\infty$ as $\lambda\to 0$.

Hence it follows that as $\lambda\to 0$

$$\begin{aligned} &\hat{b}_T \xrightarrow{a.s} b \\ &\hat{\sigma}_T^2 \xrightarrow{a.s} \sigma^2 \\ \text{and} \qquad & \\ &\hat{\gamma}_T \xrightarrow{a.s} \gamma. \end{aligned} \tag{5.2.8}$$

From the definition of stopping rule T we can write

$$\left(\tfrac{A}{\lambda}\right)^{1/2}\left[\frac{\hat{\sigma}_T^2}{1-(\hat{b}_T^2+\hat{\gamma}_T)}\frac{1+\hat{b}_T}{1-\hat{b}_T}\right]^{1/2} \leq T$$

$$\leq \left(\tfrac{A}{\lambda}\right)^{1/2}\left\{\left[\frac{\hat{\sigma}_{T-1}^2}{1-(\hat{b}_{T-1}^2+\hat{\gamma}_{T-1})}\frac{1+\hat{b}_{T-1}}{1-\hat{b}_{T-1}}\right]^{1/2}+(T-1)^{-h}\right\}+m. \tag{5.2.9}$$

Hence dividing (5.2.9) by n_0 and using (5.2.8) and then letting $\lambda \to 0$, we obtain

$$\frac{T}{n_0} \xrightarrow{a.s} 1.$$

As for **part (ii)** we have the result

$$\mathrm{E}|X| = \mathrm{E}X^+ + \mathrm{E}X^-$$

where

$$X^+ = \mathrm{Max}\,(X,0)$$

and

$$X^{-} = \text{Max}\,(-X, 0).$$

Here observe that

$$\left(\frac{T}{n_0} - 1\right)^{-} \leq 1.$$

Therefore, by dominated convergence theorem and part (i) of the Theorem 5.2.1 we have

$$\text{E}\left(\frac{T}{n_0} - 1\right)^{-} \to 0 \text{ as } \lambda \to 0.$$

Now we write

$$\left(\frac{T}{n_0} - 1\right) = \left(\frac{T}{n_0} - 1\right) I_B + \left(\frac{T}{n_0} - 1\right) I_E + \left(\frac{T}{n_0} - 1\right) I_C \tag{5.2.10}$$

and hence

$$\text{E}\left(\frac{T}{n_0} - 1\right)^{+} \leq (1-\varepsilon)\,\text{P}(B) + \varepsilon + n_0^{-1} \sum_{n \geq n_1} P[T > n] + \text{P}(C). \tag{5.2.11}$$

Since $0 < \varepsilon < 1$ is arbitrary, from Lemma 5.2.2 we have

$$\text{E}\left(\frac{T}{n_0} - 1\right)^{+} \to 0 \text{ as } \lambda \to 0.$$

So part (ii) of the thorem is also proved.

In order to prove the part (iii), (that is T is asymptotically risk efficient) assume without loss of generality that $\mu = 0$.

Now using (5.2.4) and (5.2.6)

$$\frac{R_{T,\lambda}}{R_{n_0,\lambda}} = \frac{AE\overline{X}_T^2}{2\lambda n_0} + \frac{\lambda ET}{2\lambda n_0}.$$

Since we have already proved (ii) it is enough to show

$$\frac{AE\overline{X}_T^2}{\lambda n_0} \to 1, \text{ as } \lambda \to 0. \tag{5.2.12}$$

Instead of proving (5.2.12) we will prove

$$\frac{AE\overline{X}_T^2 I_E^c}{\lambda n_0} \to 0 \tag{5.2.13}$$

and

$$\frac{AE(\overline{X}_T - \overline{X}_{n_0})^2 I_E}{\lambda n_0} \to 0 \qquad \text{as } \lambda \to 0. \tag{5.2.14}$$

Towards that end consider (5.2.13) and write

$$\frac{AE\overline{X}_T^2 I_E^c}{\lambda n_0} = \frac{AE\overline{X}_T^2 I_B}{\lambda n_0} + \frac{AE\overline{X}_T^2 I_C}{\lambda n_0} \tag{5.2.15}$$

using (4.2.7) we can write

$$\begin{aligned}
\mathrm{E}\,\overline{X}_T^2\, I_B &\le \mathrm{E} \underset{n_1 \le n \le n_2}{Max} \overline{X}_n^2\, I_B \\
&\le \mathrm{E} \underset{n_1 \le n \le n_2}{Max} \left[\sum_{j=0}^{\infty} Y_{nj}\right]^2 I_B \\
&\le \mathrm{E} \left|\sum_{j=0}^{\infty} M_{nj}\right|^2 I_B \\
&\le \sum_{j=0}^{\infty} EM_{nj}^2\, I_B + 2 \sum_{j<j'} EM_{nj} M_{nj'}\, I_B,
\end{aligned} \tag{5.2.16}$$

where,

$$Y_{nj} = n^{-1} \sum_{i=1}^{n} \left[\sum_{k=0}^{j-1} (b + \beta_{i-k})\right] \varepsilon_{i-j} \tag{5.2.17}$$

and

$$M_{nj} = \underset{n_1 \le n \le n_2}{Max}\, Y_{nj}\,.$$

Observe that for $j \ge 0$, the sequence $\{Y_{nj}\}$ is a reverse martingale w.r.t. $\{G_n\}$, where

$G_n = \sigma\{(\beta_k, \varepsilon_k),\ k \ge n\}$.

Since

$$E[Y_{n-1j} \mid G_n] = E\left\{(n-1)^{-1}\sum_{i=1}^{n-1}\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}$$

$$= (n-1)^{-1}\sum_{i=1}^{n} E\left\{\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}$$

$$= (n-1)^{-1}(n-1)E\left\{\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}$$

$$= E\left\{\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}. \tag{5.2.18}$$

Leting $Z_n = nY_{nj}$ and using (5.2.17)

we have

$$Z_n = E[Z_n \mid Z_n, Z_{n+1}, \ldots]$$

$$= E[Z_n \mid G_n]$$

$$= \sum_{i=1}^{n} E\left\{\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}$$

$$= n\ E\left\{\left[\sum_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}\middle|G_n\right\}. \tag{5.2.19}$$

Using (5.2.19) in (5.2.18) we get

$$E[Y_{n-1,j} \mid G_n] = Y_{nj}.$$

Hence $\{M_{nj}\}$ and $\{M_{nj}^2\}$ are reverse submartingales. By Schwarz inequality and Maximal inequality for reverse submartingale we have

$$E\, M_{nj}^2\, I_B \le E^{1/2}(M_{nj}^4)\, P^{1/2}(B)$$

$$\le \tfrac{16}{9} E^{1/2} \left| n_1^{-1} \sum_{i=1}^{n_1} \left[\prod_{k=0}^{j-1} (b + \beta_{i-k}) \right] \varepsilon_{i-j} \right|^4 P^{1/2}(B).$$

An application of M-Z inequality gives

$$E^{1/2} \left| n_1^{-1} \sum_{i=1}^{n_1} \left[\prod_{k=0}^{j-1} (b + \beta_{i-k}) \right] \varepsilon_{i-j} \right|^4 = O(n_1^{-1}). \tag{5.2.20}$$

Application of lemma 5.2.2 and (5.2.20) leads to

$$E\, M_{nj}^2\, I_B = O(\lambda^{1/2(1+h)})\, O(\lambda^{(p-2)/8(1+h)}).$$

Since h < $\frac{p-2}{4}$ we have

$$A \frac{EM_{nj}^2 I_B}{\lambda n_0} \to 0 \text{ as } \lambda \to 0. \tag{5.2.21}$$

Using Schwarz inequality for the second term in (5.2.16)

$$E[M_{nj}\, M_{nj'}\, I_B] \le E^{1/4}(M_{nj}^4) E^{1/4}(M_{nj'}) P^{1/2}(B).$$

So that

$$\frac{A}{\lambda n_0} \mathrm{E}[M_{nj} M_{nj'} I_B] \to 0 \text{ as } \lambda \to 0. \tag{5.2.22}$$

Using (5.2.21) and (5.2.22) in (5.2.16)

$$\frac{AE\overline{X}_T^2 I_B}{\lambda n_0} \to 0 \text{ as } \lambda \to 0.$$

Repeating the same arguments as above and using part (ii) of Lemma 5.2.2

$$\frac{AE\overline{X}_T^2 I_C}{\lambda n_0} \to 0 \text{ as } \lambda \to 0.$$

Thus we have proved (5.2.13)

Next consider

$$\mathrm{E}(\overline{X}_T - \overline{X}_{n_0})^2 \, I_E \le \mathrm{E}\left[\underset{n_2 \le n \le n_3}{Max} \left|\overline{X}_T - \overline{X}_{n_0}\right|^2\right] I_E$$

$$\le \mathrm{E}\left[\sum_{j=0}^{\infty} \underset{n_2 \le n \le n_3}{Max} \left|W_{n,n_0j}\right|\right]^2$$

$$\le \sum_{j=0}^{\infty} E \underset{n_2 \le n \le n_3}{Max} W_{n,n_0j}^2 + 2\sum_{j<j'} E \underset{n_2 \le n \le n_3}{Max} W_{n,n_0j} \underset{n_2 \le n \le n_3}{Max} W_{n,n_0j'}, \tag{5.2.23}$$

where $W_{n,n_0j} = n^{-1}\sum_{i=1}^{n}\left[\prod_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j} - n_0^{-1}\sum_{i=1}^{n_0}\left[\prod_{k=0}^{j-1}(b+\beta_{i-k})\right]\varepsilon_{i-j}$.

Note that for each fixed $j \ge 0$ $\{W_{n,n_0j},\ n_0 \le n \le n_3\}$ is a reverse martingale w.r.t. $\{G_n\}$.

Consider

$$\mathrm{E}\left\{\underset{n_2 \le n \le n_3}{Max} W_{n,n_0j}^2\right\} = \mathrm{E}\left\{\underset{(n_2 \le n \le n_0)\cup(n_0 \le n \le n_3)}{Max} W_{n,n_0,j}^2\right\} \tag{5.2.24}$$

Now applying Schwarz and maximal inequalities for reverse submartingale $\{ W_{n,n_0 J} \}$,

$$\mathrm{E}\left\{ \max_{n_0 \le n \le n_3} W^2_{n,n_0 J} \Big/ \lambda n_0 \right\} \to 0$$

and

$$\mathrm{E}\left\{ \max_{n_2 \le n \le n_0} W^2_{n,n_0 J} \right\} / \lambda n_0 \to 0.$$

Thus from (5.2.24) we get

$$\mathrm{E}\left\{ \max_{n_2 \le n \le n_3} W^2_{n,n_0 J} \right\} / \lambda n_0 \to 0.$$

The second tern in (5.2.23) can be handled similarly. This completes the proof of part (iii).

For **part (iv)** we have noted in Chapter 4 [See (4.2.13)] that as $n \to \infty$

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N\left(0,\ \frac{\sigma^2}{1-(b^2+\gamma)} \frac{1+b}{1-b}\right).$$

Now write

$$\sqrt{n}(\overline{X}_n - \mu) = \sqrt{T/n_0}\sqrt{n_0}(\overline{X}_T - \overline{X}_{n_0}) + \sqrt{T/n_0}\sqrt{n_0}(\overline{X}_{n_0} - \mu). \tag{5.2.25}$$

From (5.2.14) we have

$$\frac{AE(\overline{X}_T - \overline{X}_{n_0})^2 I_E}{\lambda n_0} \to 0 \text{ as } \lambda \to 0.$$

Using (5.2.3)

$$\frac{\lambda n_0[1-(b^2+\gamma)](1-b)}{\sigma^2(1+b)} \frac{E(\overline{X}_T - \overline{X}_{n_0})^2}{\lambda n_0} \to 0 \text{ as } \lambda \to 0.$$

That is, as $\lambda \to 0$,

$$n_0 \, E(\overline{X}_T - \overline{X}_{n_0})^2 \to 0.$$

which implies that

$$\sqrt{n_0}(\overline{X}_T - \overline{X}_{n_0}) \xrightarrow{p} 0. \tag{5.2.26}$$

From part(i) we have

$$\frac{T}{n_0} \xrightarrow{a.s} 1 \text{ as } \lambda \to 0.$$

Application of (5.2.26), part (i) of Theorem 5.2.1, (4.2.13) and Slutsky's Theorem in (5.2.25) we get (iv). This completes the proof the theorem.

5.3 Sequential Interval Estimation

In Section 3.4 we have discussed the general frame work of sequential interval estimation. Our problem in this section is to find an interval I_n for the population mean of RCAR(1) process having prescribed width 2d and a coverage probibility $1-\alpha$.

That is to find an interval I_n such that

$$P\left[\mu \in I_n\right] = 1 - \alpha.$$

Recall from section 4.2 (see (4.2.13)) that

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} \mathrm{N}(0, \frac{\sigma^2}{1-(b^2+\gamma)} \cdot \frac{1+b}{1-b}).$$

Based on this result an appropriate confidence interval for μ when b, σ^2 and γ are known is given by

$$I_{n_0} = [\ \overline{X}_{n_0} - d,\ \overline{X}_{n_0} + d].$$

where

$$n_0 = [d^{-2}\ Z^2_{1-\alpha/2}\ \frac{\sigma^2}{1-(b^2+\gamma)} \cdot \frac{1+b}{1-b}]. \tag{5.3.1}$$

and $Z_{1-\alpha/2}$ is such that

$$\frac{1}{\sqrt{2\pi}} \int_{-Z_{1-\alpha/2}}^{Z_{1-\alpha/2}} \exp\{-\tfrac{u^2}{2}\}\ du\ = 1\text{-}\alpha.$$

Note from (5.3.1) that $n_0 \to \infty$ when $d \to 0$ and

$$\mathrm{P}[\mu \in I_{n_0}] = \mathrm{P}\left[\sqrt{n_0}\left|\overline{X}_{n_0} - \mu\right| \le \frac{d\sqrt{n_0}}{\xi}\right] \to 1\text{-}\alpha,$$

where

$$\xi\ =\ \left(\frac{\sigma^2}{1-(b^2+\gamma)} \cdot \frac{1+b}{1-b}\right)^{1/2} \tag{5.3.2}$$

When at least one of the parameters b, σ^2 and γ is unknown we proposes a sequential confidence interval. For that we define a stopping rule as in the case of point estimation,

$$N = \inf\left\{n \ge m\colon n \ge d^{-2} Z^2_{1-\alpha/2}\left[\left(\frac{\hat{\sigma}^2_n}{1-(\hat{b}^2_n+\hat{\gamma})} \cdot \frac{1+\hat{b}_n}{1-\hat{b}_n}\right) + n^{-h}\right]\right\}, \tag{5.3.3}$$

where m is an initial sample size and h is a suitable constant to be defined later. Note that from the above definition of stopping rule $N \ge d^{-2}\ Z^2_{1-\alpha/2} . N^h$.

That is $N \ge \left(\frac{Z_{1-\alpha/2}}{d}\right)^{2/(1+h)}$

Thus when $d \to 0$, $N \to \infty$.

The performance of the above stopping time N and the corresponding confidence interval I_N are discussed in the following Theorem.

Theorem 5.3.1 For $p>2$, if $E|\varepsilon_1|^{4p} < \infty$, $E|b+\beta_1|^{4p} < 1$ and $h\varepsilon(0, \frac{p-2}{4})$

then as $d \to 0$

(i) $\quad \frac{N}{n_0} \xrightarrow{a.s} 1$

(ii) $\quad E\left(\frac{N}{n_0}\right) \to 1$

(iii) $\quad P[\mu \in I_N] \to 1\text{-}\alpha$.

Proof: Proof of part (i) and part (ii) are very much similar to the proof of part (i) and (ii) of Theorem 5.2.1 and hence we omit the details.

For **part (iii)**

$$P[\mu \in I_N] = P\left[\left|\overline{X}_N - \mu\right| \le d\right]$$

$$= P\left[\frac{\sqrt{N}}{\xi}\left|\overline{X}_N - \mu\right| \le \frac{d\sqrt{N}}{\xi}\right], \tag{5.3.4}$$

where ξ is as defined in as (5.3.2).

Recall from Section (5.2) that

$$\sqrt{N}(\overline{X}_N - \mu) \xrightarrow{d} N(0, \xi^2).$$

Now using the definition of n_0 (5.3.4) becomes

$$P\left[\frac{\sqrt{N}}{\xi}\left|\overline{X}_N - \mu\right| \le Z_{1-\frac{\alpha}{2}} \cdot \frac{\sqrt{N}}{\sqrt{n_o}}\right].$$

Also we have noted in Part(i) that $\dfrac{N}{n_o} \to 1$ a.s.

That is $\sqrt{\dfrac{N}{n_o}} \to 1$ a.s.

Now Part (iii) follows from (5.3.4) and the above arguments. The proof of the Theorem is complete.

The work of this chapter is summarised in Balakrishna and Jacob (1998). In the next chapter we discuss the sequential estimation of b.

CHAPTER 6

SEQUENTIAL ESTIMATION OF THE REGRESSION PARAMETER OF RCAR(1) MODEL

6.1 Introduction

The main problems of this chapter are to obtain a sequential point and interval estimation of the autoregressive parameter of an RCAR(1) model.

For algebraic simplicity we assume $\mu = 0$ and hence our model in this chapter is

$$X_i = (b + \beta_i) X_{i-1} + \varepsilon_i, \ i=1,2, \ \ldots\ldots\ldots \tag{6.1.1}$$

In addition to the assumptions a_1-a_4 made in Section 4.2, here we further assume that

$$\gamma_k = \mathrm{E}\left(\beta_i^k\right) < \infty$$

$$\sigma_k = \mathrm{E}\left(\varepsilon_i^k\right) < \infty, \qquad \text{for} \qquad k = 1,2,3,4. \tag{6.1.2}$$

Sriram (1988) considered the sequential point estimation of autoregressive parameter in a AR(1) model. Basu and Das (1995) obtained sequential least square estimator for the autoregressive parameters in a AR(p) model.

In this chapter, Section 6.2 deals with the definition of stopping rule and its properties for the point estimation of regression parameter. The main result of this chapter is included in Section 6.3. Section 6.4 deals with the sequential interval estimation.

6.2 Definition and Properties of the Stopping Rule

Given a sample of size n one wishes to estimate the autoregressive parameter b by the least squares estimator given by

$$\hat{b}_n = \frac{\sum_{i=1}^{n} X_i X_{i-1}}{\sum_{i=1}^{n} X_{i-1}^2}. \tag{6.2.1}$$

Using the equation (6.1.1) and making some algebraic manipulations as in the proof of lemma 4.3.2 we can write

$$\hat{b}_n - b = \frac{J_n}{K_n}, \tag{6.2.2}$$

where

$$\begin{aligned} J_n &= \sum_{i=1}^{n} \beta_i X_{i-1}^2 + \sum_{i=1}^{n} \varepsilon_i X_{i-1} \\ K_n &= \sum_{i=1}^{n} X_{i-1}^2 . \end{aligned} \tag{6.2.3}$$

The purpose of estimating the autoregressive parameter b is to use the model in predicting the future values of the process given the past observations. The minimum mean square error predictor of X_{n+1} given X_0, X_1 ..... Xn based on the above RCAR(1) model is $\hat{X}_{n+1} = \hat{b}_n X_n$. From that point of view it may be legitimate to employ the mean squared error loss function with a sampling cost given by

$$L_n = An^{-1} \sum [\hat{X}_i - E(X_i | X_{i-1})]^2 + \lambda n,$$

where λ is the sampling cost per unit.

Nicholls and Quinn (1982) pp.44 showed that if $E(X_{i-1}^4) < \infty$ then as $n \to \infty$,

$$\sqrt{n}\,(\hat{b}_n - b) \xrightarrow{d} N(0, \sigma_2 V^{-1} + V^{-2} E(X_{i-1}^4)\gamma_2),$$

where V is as defined in (4.2.10). Also we have noted in Section 4.3 that

$$n^{-1}\sum_{i=1}^{n} X_{i-1}^2 \xrightarrow{a.s} V \qquad as\ n \to \infty.$$

Hence

$$(\sum_{i=1}^{n} X_{i-1}^2)^{1/2}(\hat{b}_n - b) \xrightarrow{d} N(0, \sigma_2 + V^{1} E(X_{i-1}^4)\gamma_2).$$

Since the process is stationary we have $E(X_{i-1}^4) = E(X_i^4)$ for every i. Hence the asymptotic distribution can be written as

$$(\sum_{i=1}^{n} X_{i-1}^2)^{1/2}(\hat{b}_n - b) \xrightarrow{d} N(0, H^2) \tag{6.2.4}$$

where

$$H^2 = \sigma_2 + \left\{ \frac{[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]} \right\} \frac{\gamma_2}{\sigma_2}. \tag{6.2.5}$$

Now the risk function $R_n = E(L_n)$ can be evaluated using (6.2.4). Thus

$$R_n = E(L_n) = An^{-1} H^2 + \lambda n + o(n^{-1}). \qquad (6.2.6)$$

The value of n which minimizes (6.2.6) can be obtained from

$$\frac{\partial R_n}{\partial n} = 0 = -An^{-2} H^2 + \lambda.$$

The best fixed sample size n_0 is given by

$$n_0 \cong (A/\lambda)^{1/2} H \qquad (6.2.7)$$

with corresponding minimum risk

$$R_{n_0} = 2\lambda n_0. \qquad (6.2.8)$$

Note that the fixed sample size procedure n_0 defined by (6.2.7) depends on the nuisance parameters and hence the sample size cannot be specified in advance. Under these circumstances a sequential method of estimation is preferable. For this purpose, we define a stopping rule T by

$$T = \inf\{n \geq m: n \geq (A/\lambda)^{1/2}[\hat{H}_n + n^{-h}]\}, \qquad (6.2.9)$$

where m is an initial sample size, $\hat{H}_n$ is obtained from H by replacing b, γ_2, γ_3, γ_4, σ_2 and σ_4 by their respective estimators. These estimators are obtained by using the method suggested by Beran and Hall (1992).

For the model (6.1.1), we define

$$U_i = \beta_i X_{i-1} + \varepsilon_i,$$

and

$$\hat{U}_i = X_i - \hat{b}_n X_{i-1}.$$

Constrain both $\hat{\sigma}_{1n}$ and $\hat{\gamma}_{1n}$ to be zero and assume that the estimates $\hat{\sigma}_{2n}$, $\hat{\gamma}_{2n}$ ...., $\hat{\sigma}_{k-1,n}$, $\hat{\gamma}_{k-1,n}$ have already been computed.

Put

$$\overline{X}_k = n^{-1} \sum_{i=1}^{k-1} X_{i-1}^k$$

$$W_{ik} = \hat{Z}_j^k - \sum_{l=1}^{k-1} {}^kC_l X_{i-1}^l \hat{\gamma}_{l,n} \hat{\sigma}_{k-l,n}$$

$$\overline{W}_{.k} = n^{-1} \sum_{i=1}^{n} W_{ik} .$$

Then

$$\hat{\gamma}_{k,n} = \frac{\sum_{i=1}^{n} W_{ik}(X_{i-1}^k - \overline{X}_k)}{\sum_{i=1}^{n} (X_{i-1} - \overline{X}_k)^2}$$

and

$$\hat{\sigma}_{k,n} = \overline{W}_{.k} - \hat{\gamma}_{k,n} \overline{X}_k .$$

For our model

$$\hat{U}_i^2 = (X_i - \hat{b}_n X_{i-1})^2$$

$$\hat{\sigma}_{1n} = 0 = \hat{\gamma}_{1n}$$

$$W_{i2} = \hat{U}_i^2$$

$$\hat{\gamma}_{2n} = \left\{ \sum_{i=1}^{n} \hat{U}_i^2 (X_{i-1} - \overline{X}_2) \right\} \left\{ \sum_{i=1}^{n} (X_{i-1}^2 - \overline{X}_2)^2 \right\}^{-1}$$

$$\hat{\sigma}_{2n} = n^{-1} \sum_{i=1}^{n} \hat{U}_i^2 - \hat{\gamma}_{2n} \overline{X}_2 .$$

Which are equilant to estimators for $\gamma_2 = E(\beta_i^{\,2})$ and $\sigma_2 = E(\varepsilon_i^{\,2})$ suggested by Nichollas and Quinn (1982) pp.46 & 47.

Further,

$$W_{i3} = (X_i - \hat{b}_n X_{i-1})^3 - {}^3C_1 \, X_{i-1} \, \hat{\gamma}_{1n} \, \hat{\sigma}_{2n} - {}^3C_2 \, X_{i-1}^2 \, \hat{\gamma}_{2n} \, \hat{\sigma}_{1n}$$

$$= (X_i - \hat{b}_n X_{i-1})^3 .$$

$$W_{i4} = (X_i - \hat{b}_n X_{i-1})^4 - {}^4C_2 \, X_{i-1}^2 \, \hat{\gamma}_{2n} \, \hat{\sigma}_{2n}$$

$$= (X_i - \hat{b}_n X_{i-1})^4 - 6 \, X_{i-1}^2 \, \hat{\gamma}_{2n} \, \hat{\sigma}_{2n} .$$

$$\hat{\gamma}_{3n} = \left\{ \sum_{i=1}^{n} (X_i - \hat{b}_n X_{i-1})^3 (X_{i-1}^3 - \overline{X}_3) \right\} \left\{ \sum_{i=1}^{n} (X_{i-1}^3 - \overline{X}_3)^2 \right\}^{-1}$$

$$\hat{\sigma}_{3n} = n^{-1} \sum_{i=1}^{n} (X_i - \hat{b}_n X_{i-1}^3) - \hat{\gamma}_{3n} \, \overline{X}_3$$

$$\hat{\gamma}_{4n} = \left\{ \sum_{i=1}^{n} [(X_i - \hat{b}_n X_{i-1})^4 - 6X_{i-1}^2 \hat{\gamma}_{2n} \hat{\sigma}_{2n}](X_{i-1}^4 - \overline{X}_4) \right\} \left\{ \sum_{i=1}^{n} (X_{i-1}^4 - \overline{X}_4)^4 \right\}^{-1}$$

$$\hat{\sigma}_{4n} = n^{-1} \sum_{i=1}^{n} [(X_i - \hat{b}_n X_{i-1}^4) - 6X_{i-1}^2 \hat{\gamma}_{2n} \hat{\sigma}_{2n}] - \hat{\gamma}_{4n} \, \overline{X}_4 .$$

Based on the stopping time define by (6.2.9) the sequential point estimator of b is $\hat{b}_T$ and the associated risk is

$$R_T = AT^{-1}\sum_{i=1}^{T}\left[\hat{X}_i - E(X_i|X_{i-1})\right]^2 + \lambda T. \qquad (6.2.10)$$

In the next section we will prove some results which establish the optimal properties of the above sequential procedure.

6.3 Properties of the Stopping Rule

The important properties of the sequential procedure are stated in the following theorem.

Theorem 6.3.1: Let T be the stopping rule defined by (6.2.9) and $p>2$. Suppose that $E|\varepsilon_I|^{8p}<\infty$, $E|b+\beta_1|^{8p}<1$ and $h \in (0, (p-2)/4)$ then as $\lambda \to 0$

i. $\dfrac{T}{n_0} \xrightarrow{a.s} 1$

ii. $E\left|\dfrac{T}{n_0} - 1\right| \to 0$

iii. $\sqrt{T}\,(\hat{b}_T - b) \xrightarrow{d} N\left(0, \sigma_2 V^{-1} + \dfrac{V^{-2}[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^2+\gamma_2)][1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]}\right)$

and

$$\left(\sum_{i=1}^{T} X_i\right)^{1/2}(\hat{b}_T - b) \xrightarrow{d} N(0, H^2)$$

iv. $\dfrac{R_T}{R_{n_0}} \to 1.$

Proof of this theorem depends on a number of lemmas. The following notations will be useful in the further discussion.

Let

$$K = \left(\tfrac{4}{\lambda}\right)^{1/2}, \ n_1 = K^{1/(1+h)}, \quad n_2 = n_0(1-\varepsilon),\ n_3 = n_0(1+\varepsilon),\ 0<\varepsilon<1,$$

$$E = [n_2 < T < n_3] \qquad B = [T \le n_2] \qquad C = [T \ge n_3]$$

I_F and F^C denote the indicator and complement of a set F respectively.

The asymptotic properties of the estimators are discussed in the following lemma. We have proved that $\hat{\gamma}_{2n} \to \gamma$ and $\hat{\sigma}_{2n} \to \sigma_2$ almost surely as $n \to \infty$, in chapter 4. Exactly in the same manner we can prove the a.s convergence of $\hat{\gamma}_{3n}$, $\hat{\gamma}_{4n}$ and $\hat{\sigma}_{4n}$.

Lemma 6.3.1 For $p \ge 1$, if $E(\varepsilon_1^{2kp}) < \infty$, $E|b+\beta_1|^{2kp} < 1$ and $h \in (0, (p-2)/4)$ then as $n \to \infty$, $\hat{\gamma}_{kn} \xrightarrow{a.s} \gamma_k$, $\hat{\sigma}_{kn} \xrightarrow{a.s} \sigma_k$ for $k = 1,2,3,4$ and $P\left[\left|\hat{H}_n - H\right| > \varepsilon\right] = O(n^{-p/2})$ for $\varepsilon > 0$.

Proof: Once we prove the asymptotic convergence of for k=1 3 and 4 proof of the last part of this lemma is a direct consequence of Result 1.6.12. We have already proved a similar result in Lemma 5.2.1 and hence omit the details.

We also skip the proof of the next Lemma as it is similar to the proof of lemma 5.2.2.

Lemma 6.3.2 Under the conditions of lemma 6.2.1 for $\varepsilon > 0$

$$P[T \le n_2] = O\left(\lambda^{\frac{(p-2)}{4(h+1)}} \right)$$

and

$$\sum_{n \ge n_3} \mathrm{P}[T>n] = \mathrm{O}\left(\lambda^{\frac{(p-2)}{4}} \right).$$

Lemma 6.3.3 If J_n is as defined by (6.2.3), then the sequence $\left\{ \frac{J_n}{\sqrt{n}},\ n \ge 1 \right\}$ is stochasticaly bounded and uniformly continuous in probability.

Proof: We have already proved in Lemma 4.3.2 that

$$\|J_n\| = \mathrm{O}(n^{1/2}), \text{ and hence } \left\{ \frac{J_n}{\sqrt{n}},\ n \ge 1 \right\} \text{ is stochasticaly bounded.}$$

Now to prove u.c.i.p we write

$$\left| \frac{J_{n+k}}{\sqrt{n+k}} - \frac{J_n}{\sqrt{n}} \right| \le \frac{1}{\sqrt{n}} \left| J_{n+k} - J_n \right| + \left| 1 - \frac{\sqrt{n}}{\sqrt{n+k}} \right| \left| \frac{J_n}{\sqrt{n}} \right|, \text{ for } k \ge 0. \qquad (6.3.1)$$

For $k \le n\delta$ the second term in (6.3.1) is bounded by $[1-(1+\delta)^{-1/2}] \left\| \frac{J_n}{\sqrt{n}} \right\|$ which tends to zero as $\delta \to 0$ uniformly in $n \ge 1$. For the first term applying the maximal inequality for the martingale $\{J_n\}$ and then using the independence of X_j and (ε_k, β_k) for $j<k$ we get

$$\mathrm{P}\left\{ \underset{0 \le k \le n\delta}{Max} \left| J_{n+k} - J_n \right| \le \frac{\varepsilon\sqrt{n}}{2} \right\} \le \frac{4}{\varepsilon^2 n} \mathrm{E}\left| J_{n+[n\delta]} - J_n \right|^2$$

$$= \frac{4}{\varepsilon^2 n} \mathrm{E} \left| \sum_{i=n+1}^{n+n\delta} \beta_i X_{i-1}^2 + \sum_{i=n+1}^{n+n\delta} \varepsilon_i X_{i-1} \right|^2$$

$$= \frac{4}{\varepsilon^2 n} \mathrm{E}\left\{\gamma \sum_{i=n+1}^{n+n\delta} E(X_{i-1}^4) + \sigma_2 \sum_{i=n+1}^{n+n\delta} E(X_{i-1}^2)\right\}$$

$$= \mathrm{O}(\delta) \text{ as } \delta \to 0 \text{ uniform in } n \geq 1.$$

From the above argument, it follows that $\left\{\frac{J_n}{\sqrt{n}},\ n \geq 1\right\}$ is u.c.i.p by Definition 1.6.3. This completes the proof of the lemma.

Lemma 6.3.4: The sequence $\left\{\frac{K_n}{n},\ n \geq 1\right\}$ is u.c.i.p and stochastically bounded, where K_n is defined by (6.2.3).

Proof: We have the result (See Lemma 4.3.2) that

$$n^{-1} K_n \xrightarrow{a.s} \frac{\sigma_2}{1-(b^2+\gamma_2)} < \infty.$$

But stochastic boundedness and u.c.i.p holds if a sequence converges to a finite limit a.s (cf. Remark 1.6.2. and Definition 1.6.1.). Hence the lemma follows.

Lemma 6.3.5 The sequences $\{\sqrt{n}\,(\hat{b}_n - b)\}$ and $\left\{\frac{J_n}{\sqrt{K_n}}\right\}$ are u.c.i.p and stochastically bounded.

Proof: We have noted in Remark 1.6.3 that any continuous function of u.c.i.p and stochasticaly bounded sequence is also u.c.i.p and stochasticaly bounded..

Here we can write

$$\sqrt{n}\,(\hat{b}_n - b) = \frac{J_n}{\sqrt{n}} \Big/ n^{-1} K_n .$$

Hence $\{\sqrt{n}\,(\hat{b}_n\text{-}b)\}$ is u.c.i.p follows from Remark 1.6.3, Lemma 6.3.3 and Lemma 6.3.4. Writing $\frac{J_n}{\sqrt{K_n}} = \frac{J_n/\sqrt{n}}{\sqrt{K_n/n}}$ and repeating the above arguments we can prove that $\left\{\frac{J_n}{\sqrt{K_n}}\right\}$ is u.c.i.p. and stochasticaly bounded.

Lemma 6.3.6 If for $p\geq 1$, $E|\varepsilon_l|^{4p}<\infty$, $E|b+\beta_1|^{4p}<1$ and

$$E\left|\left[\sum_{i=1}^{M}(\beta_{i+1}-\varepsilon_i+\beta_{i+1}X_i-\beta_i X_{i-1})^2\right]^{-1}\right|^{2p} <\infty,$$

for some positive integer M, then $\{|nW_n|^q\}$ is uniformly integreable (u.i) for all $q<p$, where

$$W_n = n^{-1} K_n\,(\hat{b}_n\text{-}b)^2 \tag{6.3.2}$$

Proof: We have

$$n\,W_n = K_n\,(\hat{b}_n\text{-}b)^2 = \frac{J_n^2}{K_n}.$$

An application of Schwartz inequality gives

$$\left\|\frac{J_n^2}{K_n}\right\|_p \leq \left\|\frac{J_n^2}{n}\right\|_{2p}\left\|\frac{n}{K_n}\right\|_{2p}$$

But note that $\|J_n\| = O(n^{1/2})$. [See Lemma 4.3.2]

and hence it follows that

$$\left\| \frac{J_n^2}{n} \right\| = O(1).$$

Hence from the definition of uniform integerability it is enough to prove that

$$\left\| \sup_{n \geq M+1} \frac{n}{K_n} \right\|_{2p} < \infty, \text{ for some } M.$$

To this end consider

$$2\sum_{i=1}^{n} X_{i-1}^2 = 2\sum_{i=0}^{n-1} X_i^2 = \sum_{i=0}^{n-2} X_i^2 + \sum_{i=0}^{n-2} X_{i+1}^2 + X_0^2 + X_{n-1}^2$$

$$\geq \sum_{i=0}^{n-2} [X_i^2 + X_{i+1}^2]$$

$$\geq \tfrac{1}{2} \sum_{i=0}^{n-2} [X_i - X_{i+1}]^2.$$

Letting

$$d_i = X_i - X_{i+1} \tag{6.3.3}$$

we have

$$\sum_{i=1}^{n-2} (\varepsilon_{i+1} - \varepsilon_i + \beta_{i+1} X_i - \beta_i X_{i-1})^2 = \sum_{i=0}^{n-2} [d_i - bd_{i-1}]^2$$

$$= (1 + |b|)^2 \sum_{i=0}^{n-2} d_i^2 \tag{6.3.4}$$

where, $d_{-1} = -X_0$.

Now using (6.3.3) and (6.3.4) we can write

$$\sum_{i=0}^{n} X_{i-1}^2 \geq \tfrac{1}{4} (1 + |b|)^{-2} = \sum_{i=1}^{n-2} (\varepsilon_{i+1} - \varepsilon_i + \beta_{i+1} X_i - \beta_i X_{i-1})^2.$$

For $n \geq M+1$, then there exist a q such that

$$qM + 1 \leq n \leq (q+1)M.$$

Since $\sum_{i=0}^{n} X_{i-1}^2$ is increasing in n we have

$$\frac{n}{\sum_{i=0}^{n} X_{i-1}^2} \le \frac{(q+1)M}{\sum_{i=0}^{qM+1} X_{i-1}^2}$$

$$\le 4M\,(1+|b|)^2 \frac{q}{\sum_{i=1}^{q} B_{iM}},$$

$$\le 4M\,(1+|b|)^2 \underset{q\ge 1}{Sup}\; q^{-1}\sum_{i=1}^{q} 1/B_{iM},$$

where $B_{iM} = \sum_{j=(i-1)M+1}^{iM} (\varepsilon_{j+1} - \varepsilon_j + \beta_{j+1}X_j - \beta_j X_{j-1})^2$.

But $\left\{q^{-1}\sum_{i=1}^{q}(1/B_{iM}), G_{qM},\ q \ge 1\right\}$ is a reverse submartingale,

where $G_{qM} = \sigma\left\{l^{-1}\sum_{i=1}^{l} 1/B_{iM},\ l \ge q\right\}$.

An application of the maximal inequality for reverse submartingale yields

$$\left\| \underset{q\ge 1}{Sup}\, q^{-1}\sum_{i=1}^{q}(1/B_{iM}) \right\|_{2p} \le \frac{2p}{2p-1}\|1/B_{1M}\|_{2p},$$

which is finite by assumption. Hence the proof of the Lemma is complete.

Now we proceed to prove the main theorem.

Proof of Theorem 6.3.1

We have noted that $\hat{b}_n \xrightarrow{a.s} b$, $\hat{\gamma}_{kn} \xrightarrow{a.s} \gamma_k$, $\hat{\sigma}_{kn} \xrightarrow{a.s} \sigma_k$ $k = 1,2,3,4$ as $n \to \infty$. Thus $\hat{H}_n \xrightarrow{a.s} H$, as $n \to \infty$. Also from the definition of T we have

$$T \geq (A\ \lambda)^{1/2(h+1)} \text{ and hence}$$

$T \to \infty$, as $\lambda \to 0$.

Hence

$$\hat{H}_T \xrightarrow{a.s} H. \tag{6.3.5}$$

Using (6.3.5) and the definition of the stopping rule (6.2.9) we can write

$$(A\ \lambda)^{1/2}\ \hat{H}_T \leq T \leq (A/\lambda)^{1/2}\ [\hat{H}_{T-1} + (T-1)^{-h}] + m. \tag{6.3.6}$$

Now by taking $\lambda \to 0$ and dividing (6.3.6) by n_0 and using the expression (6.2.7) for n_0, we have

$$\frac{T}{n_0} \xrightarrow{a.s} 1.$$

Thus we have proved the part (i) of the theorem.

Part (ii) can be proved using Lemma 6.3.2 and Lemma 6.3.3. However we skip the details as they follow similar to the proof of part (ii) of Theorem 5.2.1.

We have noted in Section 6.2 that $n \to \infty$,

$$\sqrt{n}\,(\hat{b}_n - b) \xrightarrow{d} \mathrm{N}\left(0, \sigma_2 V^{-1} + \frac{V^{-2}[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^2+\gamma_2)][1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]}\right)$$

and

$$\left(\sum_{i=1}^{n} X_i\right)^{1/2} (\hat{b}_n - b) \xrightarrow{d} \mathrm{N}(0, H^2).$$

Now Anscombe's Theorem [See Result 1.6.5] can be applied to get part (iii) of the theorem. The conditions for Anscombe's Theorem are verified in Lemma 6.3.5 and in part (i) of the theorem.

Now let us prove **part (iv).**

We have

$$R_{n_0} = 2(A\lambda)^{1/2} H$$

and

$$R_T = A\mathrm{E}W_T + \lambda\, \mathrm{E}T,$$

where W_n is as defined by (6.3.2). Thus we have

$$\frac{R_T}{R_{n_0}} = \tfrac{1}{2} H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T + \tfrac{1}{2}\, \mathrm{E}(T/n_0). \qquad (6.3.6)$$

By part (ii) of the theorem we have $\mathrm{E}\left(\dfrac{T}{n_0}\right) \to 1$ as $\lambda \to 0$.

Hence in order to prove $\dfrac{R_T}{R_{n_0}} \to 1$ it is enough to show that as $\lambda \to 0$,

$$H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T\, I_{E^c} \to 0 \qquad (6.3.7)$$

and

$$H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T\, I_E \to 1. \qquad (6.3.8)$$

Consider (6.3.7)

$$H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T\, I_{E^c} = H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T\, I_B + H^{-1} (A/\lambda)^{1/2}\, \mathrm{E}W_T\, I_D. \qquad (6.3.9)$$

Using Schwartz inequality we can write the first term in (6.3.9) as

$$H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}W_T I_B \le H^{-1}(A/\lambda)^{1/2}\left\|W_T I_{\{T\le n_2\}}\right\|_2 \mathrm{P}^{1/2}\ [T\le n_2]$$

$$\le H^{-1}(A/\lambda)^{1/2}\left(\sum_{n=n_1}^{n_2} EW_n^2\right)^{1/2}\mathrm{P}^{1/2}\ [T\le n_2]$$

$$\le H^{-1}(A/\lambda)^{1/2}\ \underset{n>1}{Sup}\|nW_n\|.\left(\sum_{n=n_1}^{\infty} n^{-2}\right)^{1/2}\mathrm{O}\left(\lambda^{\frac{(p-2)}{8(h+1)}}\right)$$

$$=(A/\lambda)^{1/2}H^{-1}\,\mathrm{O}(n_1^{-1/2})\mathrm{O}\left(\lambda^{\frac{(p-2)}{8(h+1)}}\right)$$

$$\to 0,\ \text{since } h\in(0,(p\text{-}2)/4) \text{ as } \lambda\to 0 \qquad (6.3.10)$$

The last inequality is due to Lemma 6.3.2, Lemma 6.3.5 and Lemma 6.3.6.

Using similar arguments we can show that the second term in (6.3.9) as

$$H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}W_T I_D \to 0 \text{ as } \lambda\to 0. \qquad (6.3.11)$$

Now combining (6.3.10) and (6.3.11) we get the required result (6.3.7).

In order to prove (6.3.8) it is enough to show that as $\lambda\to 0$,

$$H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}(W_T)I_E \xrightarrow{d} \chi_1^2. \qquad (6.3.12)$$

and

$$\{H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}(W_T)\,I_E\} \text{ is uniformly integrable}. \qquad (6.3.13)$$

Here χ_1^2 denote a Chi-square r.v with one degree of freedom.

From part (iii) of the theorem using (6.3.2) we have

$$H^{-1}\sqrt{TW_T} \xrightarrow{d} \mathrm{N}(0,1) \text{ as } \lambda \to 0$$

and hence we have

$$H^{-2}TW_T \xrightarrow{d} \chi_1^2 \text{ as } \lambda \to 0$$

Also using part (i), definition of n_o, (6.2.7), Slutsky Theorem and the fact $I_E \xrightarrow{a.s.} 1$ as $\lambda \to 0$, we have

$$H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}(W_T)I_E \xrightarrow{d} \chi_1^2, \text{ as } \lambda \to 0.$$

Hence we the equation (6.3.12).

To prove (6.3.13) we have consider for $t>1$,

$$H^{-t}(A/\lambda)^{t/2}\,\mathrm{E}\,W_T^t\, I_B \le H^{-t}(A/\lambda)^{t/2}\,\mathrm{E}\,\underset{n_2<n<n_3}{Max}\, W_n^t$$

$$\le H^{-t}(A/\lambda)^{t/2}\, n_2^{-2t} \left\| \underset{n\ge M+1}{Sup} \frac{n}{K_n} \right\|_{2t}^{t} \left\| \underset{n_2<n<n_3}{Max} J_n \right\|_{2t}^{2t}.$$

$$\le H^{-t}(A/\lambda)^{t/2}\, n_2^{-2t}\, \mathrm{O}(n_3^t) = \mathrm{o}(1).$$

The last inequality follows from Lemma 6.3.5 and maximal inequality for the martiangale $\{J_n\}$.

Hence $H^{-1}(A/\lambda)^{1/2}\,\mathrm{E}\,W_T\, I_B$ is u.i. Thus

$$\frac{R_T}{R_{n_o}} \to 1 \text{ as } \lambda \to 0.$$

This complete the proof of the Theorem.

6.4 Sequential Interval Estimation

In Section 5.3 we have discussed the sequential interval estimation problem for the mean of RCAR(1) model. Our problém in this section is to find an interval I_n for the regression parameter b of RCAR(1) process having prescribed width 2d and a coverage probibility $1-\alpha$.

That is to find an interval I_n such that

$$P\left[b \in I_n\right] = 1-\alpha .$$

Recall from Section 6.2 that

$$\sqrt{n}(\hat{b}_n - b) \xrightarrow{d} N\left(0, \sigma_2 V^{-1} + \frac{V^{-2}[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^2+\gamma_2)][1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]}\right)$$

Based on this result an appropriate confidence interval for b when all the parameters are known is given by

$$I_n = [\ \hat{b}_n - d,\ \hat{b}_n + d].$$

where

$$n = d^{-2}\, Z_{1-\alpha/2}^2 \left(\sigma_2 V^{-1} + \frac{V^{-2}[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^2+\gamma_2)][1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]}\right). \quad (6.4.1)$$

and $Z_{1-\alpha/2}$ is such that

$$\frac{1}{\sqrt{2\pi}} \int_{-Z_{1-\alpha/2}}^{Z_{1-\alpha/2}} \exp\{-\tfrac{u^2}{2}\}' du = 1-\alpha.$$

Note from (6.4.1) that $n \to \infty$ when $d \to 0$

and

$$P[b \in I_n] = P\left[\frac{\sqrt{n}}{\kappa}\left|\hat{b}_n - b\right| \le \frac{d\sqrt{n}}{\kappa}\right] \to 1\text{-}\alpha,$$

where

$$\kappa = \left(\sigma_2 V^{-1} + \frac{V^{-2}[1-(b^2+\gamma_2)]\sigma_4 + 6\sigma_2^4(b^2+\gamma_2)}{[1-(b^2+\gamma_2)][1-(b^4+\gamma_4+6b^2\gamma_2+4b\gamma_3)]}\right)^{1/2} \qquad (6.4.2)$$

When at least one of the parameters b, σ_2, σ_3, σ_4, γ_2, γ_3, γ_4 is unknown we proposes a sequential confidence interval. For that we define a stopping rule as in the case of point estimation,

$$N_b = \inf\left\{n \ge m: n \ge d^{-2} Z_{1-\alpha/2}^2\left[\hat{\kappa}_n^2 + n^{-h}\right]\right\},$$

where $\hat{\kappa}_n$ is obtained from (6.4.2) by replacing the parameters by corresponding estimators defined in Section 6.2.

Here m is an initial sample size and h is a suitable constant to be defined later. Note that from the above definition of stopping rule $N_b \ge d^{2}\ Z_{1-\alpha/2}^2 . N_b^{-h}$.

That is $N_b \ge \left(\frac{Z_{1-\alpha/2}}{d}\right)^{2/(1+h)}$

Thus when $d \to 0$, $N_b \to \infty$.

The performance of the above stopping time N_b and the corresponding confidence interval I_{N_b} are discussed in the following Theorem. Since the proof is similar to that of Theorem 5.3.1 we omit the details.

Theorem 6.4.1 For $p>2$, if $E\left|\varepsilon_1\right|^{8p} < \infty$, $E\left|b+\beta_1\right|^{8p} < 1$ and $h \varepsilon \left(0, \frac{p-2}{4}\right)$ then as $d \to 0$

(i) $\quad \frac{N_b}{n} \xrightarrow{a.s} 1$

(ii) $\quad \mathrm{E}\left(\frac{N_b}{n}\right) \to 1$

(iii) $\quad \mathrm{P}[b \in I_{N_b}] \to 1-\alpha.$

6.5 Directions of future work

In this thesis we have considered the sequential estimation for some Markovian models, like autoregressive minification models and Random Coefficient Autoregressive model. We have also discussed the optimality properties of the sequential procedures.

We have the plan to extend this work to p[th] order Random Coefficient Autoregressive model (RCAR (p)), vector valued RCAR (p) model. Also we would like to work on the second order approximations for the expected values of the stopping rules. The investigations towards these directions are in progress.

The work of this chapter is summarised in Balakrishna and Jacob (1997).

REFERENCES

Adke,S.R and Balakrishna,N (1992). Estimation of the mean of some stationary Markov sequences. *Commun. Statist.-Theory & Meth.*, 21(1), 137-159.

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

Andel , J (1976). Autoregressive series with random parameters. *Math . Operations forech.U. Statist.* , 7, 735-741.

Anscombe,F.J. (1953). Sequential estimation, *J. Roy. Statist. Soc. B,* 15, 1-29.

Arnold,B.C and Hallet , J.T (1989). A characterization of the Pareto process among stationary stochastic processes of the form $X_n = c \min (X_{n-1}, Y_n)$. *Statist. Prob . Letters,*

Balakrishna,N (1998). Estimation for semi Pareto processes. To appear in *Commun. Statist. Theory & Meth.*, 27(9).

Balakrishna,N and Jacob,T.M. (1997). Sequential estimation of Autoregressive parameter of RCAR(1) model, *Far East J. Theo. Stasit.,* 1(1), 1-14.

Balakrishna,N and Jacob,T.M. (1998). Sequential estimation of the mean of a first order random coefficient autoregressive process. (Submitted the revised version).

Balakrishna,N and Jacob,T.M. (1998a). Parameter estimation in minification processes. (presented at the International Conference on Stochastic Processes and Their Applications held at Anna University, Madras, January 1998

Basu,A.K. and Das,J.K. (1995), Sequential least squares and shrinkage estimators of the autoregressive parameters in AR(p) model, *Proc. XVth ISPS* - Thirunelveli.

Baswa,I.V, McCormick,W.P and Srirarm,T.N. (1990). Sequential estimation for dependent observations with an application to non standard autoregressive process. *Stoch. Proc. and Their. Applic.*, 35, 149-168.

Baswa,I.V and Prakasa Rao,B.L.S. (1980). **Statistical Inference for Stochastic Processes,** Academic Press, London.

Beran,R. and Hall,P. (1992). Estimating coefficent distribution in random coefficient regressions, *Ann. Statist.*, 20, 1970-1984.

Billingsley,P. (1961). Statistical Inferences for Markov Process, University of Chicago Press.

Billingsley,P. (1968). **Convergence of Probability Measures**, Jhon Wiley and Sons, New York.

Box,G.E.P. and Jinkins,G.M (1970). **Timeseries analysis, forecast and control,** Holden Day, San Francisco.

Chow,Y.S. and Robbins,H. (1965). On asymptotic theory of fixed width confidence interval for the mean., *Ann. Math. Statist.*, 36, 457-462.

Chow,Y.S.and Teicher,H. (1978). **Probability Theory, Independence, Interchangebility, Martingales,** Springer-verlag, New York.

Dewald and Lewis (1985). A new Laplace second order auto regressive time series model - NLAR(2), *IEEE Trans. Inf. Thoery*, 31, 645-652.

Fakhre-Zakeri,I. and Lee,S. (1992). Sequential estimation of the mean of a linear process. *Sequential Anal.*,11, 181-197.

Glasserman,P and Yao,D.D. (1995). Stochastic vector difference equations with stationary coefficients, *Adv. Appl. Prob.*, 32, 851-866.

Gaver,D.P and Lewis,P.A.W. (1980). First order autoregressive gamma sequence and point processes. *Adv. Appl. Prob.*, 12, 729-745.

Ghosh,B.K. and Sen,P.K. (1991). **Handbook of Sequential Analysis,** Marcel Dekker Inc, New York.

Hutton,J.L (1990) Non-negative time series models for dry river flow, *J. Appl. Prob.*, 27, 171-182.

Ibragimov,I.A and Linnik,Y.U (1971) **Independent and stationary sequence of random variables**, Wolters-Noordhott Publishing Groninyon , The Netherlands.

Karlin,S and Taylor,H.M (1974). **A First Course in Stochastic Processes**, Second Edition, Academic Press, New York.

Koul,H.L. and Schick.A (1996) Adaptive estimation in a random coefficent autoregressive model, *Ann. Statist.*, 24(2), 1025-1052.

Lawrance,A.J. (1991). Directionality and Reversibility in time series, *Int. Statis. Rev.*, 59(1), 67-79.

Lawrence,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.

Lawrence,A.J and Lewis,P.A.W. (1985). Modelling and residual analysis of non linear autoregressive time series in exponential variables, *J. Roy, Statist. Soc. B,* 47(2), 165-202.

Leadbetter,M.R., Lingren,G. and Rootzer, H. (1983), **Extremes and Related Properties of Random Sequences and Processes**, Springer Verlag,New York.

Lee,S (1994). Sequential estimation for the parameters of a stationary autoregressive model. *Sequential Anal.*, 13, 301-317.

Lewis,P.A.W. and Mckenzie,E.D.(1991). Minification processes and their transformations *J. Appl. Prob.*, 28, 45-57.

Lewis,P.A.W, Mckenzie,E.D. and Hugus,D.K.(1989). Gamma processes, *Commun. Statist.- Stichastic models,* 5(1), 1-30.

Nicholls, D.F and Quinn,B.C (1982). **Random coefficient autoregressive models, An Introduction**: Lecture notes in Statistics 11, Springer Verlag, New York.

Paulson,A.S. and Uppuluri,V.R.R (1972) Limit laws of a sequence determined by a random difference equation governing a one compartment system. *Math. Biosci.*, 13, 325-333.

Pillai,R.N (1991). Semi-Pareto processes, *J. Appl. Prob.*, 28, 461-465.

Rao,C.R. (1973). **Linear Statistical Inference and its Applications**, Second Edition, Wiley Eastern Limited, New Delhi.

Robbins,H (1959). Sequential estimation of the mean of a normal population, *Prob. Statist.*, 234-245.

Rohatgi,V.K (1976). **An Introduction to Probability and Mathematical Statistics**, Wiley Eastern Limited, New Delhi.

Schick, A (1996). $\sqrt{n}$-consistent estimation in a random coefficent autoregressive model., *Austral. J. Statist.*, 38(2), 155-160.

Sen,P.K. (1982). **Sequential Non Parametrics**, Jhon Wiley, New York.

Sim,C.H. (1986), Simulation of Weibull and gamma autoregressive stationarity processes, *Commun. Statist.- Simula. Camputa*, 15(4), 1141-1146.

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

Srirarm, T.N. (1987) Sequential estimation of the mean of a first order stationary autoregressive process, *Ann. Statist.*, 15, 1079-1090

Srirarm,T.N. (1988) Sequential estimation of the autoregressive parameter in a first order autoregressive processes, *Sequential. Anal.*, 7, 55-74.

Stout,W.F. (1974). **Almost Sure Convergence,** Academic Press New York

Travers,L.V. (1977) The exact distribution of extremes of a non-Gaussian process. *Stoch. Proc. Appl.*, 5, 151-156.

Travers,L.V. (1980). An exponential markovian stationary process, *J. Appl. Prob.*, 17, 1117-1120.

Tong,H. (1990). **Nonlinear Time Series: A Dynamical Approach** , Univ. Press, San Fransico

Vervaat,W. (1979). On a stochastic difference equation and a representation of non-negative infinitly divisible random variable, *Adv. Appl. Prob.*, 11, 750-783.

Woodroof,M. (1982). **Nonlinear Renewal Theory in Sequential Analysis,** SIAM Philadelphia.

Yeh,H.C., Arnold,B.C. and Robertson, C.A. (1988). Pareto processes, *J. Appl. Prob.*, 25, 291-301.