

LIBRARY
Michigan State
University

This is to certify that the

dissertation entitled

SEQUENTIAL ESTIMATION OF PARAMETERS
IN A FIRST ORDER AUTOREGRESSIVE MODEL

presented by

THARUVAI NARAYANASWAMI SRIRAM

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Statistics



Major professor

(H. L. Koul)

Date 8-7-86.



RETURNING MATERIALS:

Place in book drop to
remove this checkout from
your record. FINES will
be charged if book is
returned after the date
stamped below.

--	--	--

**SEQUENTIAL ESTIMATION OF PARAMETERS
IN A FIRST ORDER AUTOREGRESSIVE MODEL**

by

THARUVAI NARAYANASWAMI SRIRAM

A DISSERTATION

**Submitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of**

DOCTOR OF PHILOSOPHY

**Department of Statistics and Probability
1986**

ABSTRACT

SEQUENTIAL ESTIMATION OF PARAMETERS IN A FIRST ORDER AUTOREGRESSIVE MODEL

by

THARUVAI NARAYANASWAMI SRIRAM

In this thesis two sequential estimation problems concerning the parameters in a first order, non-explosive, autoregressive model are considered. The first is the sequential point and fixed width interval estimation of the location parameter. The second is the sequential point estimation of the autoregressive parameter when the location is assumed to be known.

For the location problem, sequential estimators based on sample mean are proposed using certain stopping times. The sequential point estimator is shown to be asymptotically risk efficient as the cost per observation tends to zero. The sequential interval estimator is shown to be asymptotically consistent and the corresponding stopping rule is shown to be asymptotically efficient as the width of the interval tends to zero.

For the estimation of the autoregressive parameter, sequential estimator based on least squares estimator is proposed using certain stopping time. The sequential point estimator is shown to be asymptotically risk efficient as the estimation error tends to infinity. Furthermore, a second order expansion to the expected stopping time is also

obtained. Finally, a thorough analysis of the 'regret' is also carried out as the estimation error tends to infinity.

To my parents and my wife

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to Professor Hira L. Koul for his continuous encouragement, guidance and especially his patience during the preparation of this dissertation.

I would also like to thank Professors James Hannan, Joseph Gardiner and Habib Salehi for serving on my committee. My special thanks go to Professor James Hannan for his careful reading of this dissertation and for giving useful suggestions, which helped greatly in the final presentation of this thesis.

I would also like to thank Professor Michael Woodroffe for many helpful discussions.

I would like to thank my wife Vidya for her moral support and encouragement during the preparation of this thesis.

Finally, special thanks go to Miss Sharon Carson for her excellent typing of the manuscript.

TABLE OF CONTENTS

Chapter		Page
0	Introduction and Summary	.1
1	Sequential Estimation of the Mean of a First Order Autoregressive Process	.9
	1.1. Introduction	.9
	1.2. Proofs	13
2	Sequential Estimation of the Auto- regressive parameter in a First Order Autoregressive Process	28
	2.1. Introduction	28
	2.2. Risk Efficiency	30
	2.3. A Second Order Expansion for the Expected Stopping Time. Asymptotic Normality and Uniform Integrability of Standardized Stopping Time	40
	2.4. Regret Analysis	55
	Bibliography	72

CHAPTER 0

INTRODUCTION AND SUMMARY

The sequential point and fixed width interval estimation has seen a proliferation of literature ever since the fundamental papers of Robbins (1959) and Chow and Robbins (1965). See, for example, the book by Sen (1982), the monograph by Woodroffe (1982) and detailed references given in Chow and Martinsek (1982) and Martinsek (1983, 84, 85).

Two basic problems of estimation are the following. The first problem is to determine the sample size that minimizes the risk R_n (expected loss) for a suitably defined loss function, L_n . The second problem is to construct a confidence interval for a parameter θ of prescribed width $2d$ and coverage probability $1 - \alpha$, for $0 < \alpha < 1$. In either case, the best fixed sample size procedure (BFSP), say n_0 , possessing the desired property generally depends on the underlying nuisance parameter(s). Therefore, the sample size cannot be specified in advance to solve these problems. In order to overcome this dependence on nuisance parameter(s) it is customary to define a sequential sampling rule (often called stopping time) similar to the rule considered by Robbins (1959). Essentially, this rule, say t , samples observations sequentially, updates a suitable estimator of nuisance parameter(s) in BFSP and stops sampling as soon as the number of observations exceed that of the estimated BFSP.

Thus one is led to solve these problems using sequential procedures.

In the point estimation problem, the performance of a stopping rule t is usually measured by

- (i) the risk efficiency: R_t/R_{n_0} , and
- (ii) the regret: $R_t - R_{n_0}$.

where R_t is the risk due to using the stopping time t . In the interval estimation problem, the performance of t is usually measured by

- (i) the consistency: $P[\theta \in I_t]$, and
- (ii) the efficiency: $E(t/n_0)$.

where I_t is the fixed width interval estimator using t .

The problem of sequential point estimation of the mean of independent, identically distributed (i.i.d) observations, with unknown variance and loss equal to linear combination of squared error and sample size, has been considered by Robbins (1959), Starr (1966), Starr and Woodroffe (1969) and Woodroffe (1977) for the normal case; by Starr and Woodroffe (1972) and Woodroffe (1977) for the gamma case; by Vardi (1979) for the poisson case; by Ghosh and Mukhopadhyay (1979), Chow and Yu (1981), Chow and Martinsek (1982) and Martinsek (1983) for the unknown distribution case. In all of these papers the sample mean is used to estimate the population mean. In some of the papers the penalty for estimation error is A units whenever the cost c for each observation is taken to be one

unit and vice versa. Accordingly, the asymptotic results are obtained as $A \rightarrow \infty$ or $c \rightarrow 0$.

For the Normal case, Robbins (1959) introduced a stopping time based on updated versions of sample variance and obtained some numerical and Monte Carlo results which suggested the boundedness of the regret. Using the same stopping rule, Starr (1966) established that $R_t/R_{n_0} \rightarrow 1$ as $A \rightarrow \infty$ (asymptotically risk efficient), and a much stronger result $R_t - R_{n_0} = o(1)$ as $A \rightarrow \infty$ (bounded regret) has been established by Starr and Woodroffe (1969). Furthermore, Woodroffe (1977) has given second order expansions for the expected stopping time and has also shown that $R_t - R_{n_0} = 1/2 + o(1)$ as $A \rightarrow \infty$.

For the Gamma scale and Poisson mean problems, bounded regret has been obtained by Starr and Woodroffe (1972) and Vardi (1979), respectively, using stopping rules which differ from the rule considered in the normal case and which only makes sense in these special cases. In the gamma case, Woodroffe (1977) has also obtained second order expansions to expected stopping time and to the regret.

For the general case, Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981) have proposed sequential procedures using a stopping rule of the type considered by Robbins (1959). They have shown, under certain growth rate conditions on the initial sample size, that their procedures are asymptotically risk efficient as $c \rightarrow 0$. Chow and Martinsek

(1982) have shown that the regret of these procedures is bounded as $A \rightarrow \infty$. Furthermore, Martinsek (1983) has obtained a second order expansion for the regret of these procedures. Martinsek (1983) has also constructed an example where the regret can take arbitrarily negative values, as the distribution of the observations varies (even among symmetric distributions). This shows that for some (nonnormal) distributions and large A , it is better to use sequential procedure when the variance is unknown than to use BFSP when the variance is known.. In effect, the sequential procedure does better by being sensitive to characteristics of the distribution other than the variance.

As mentioned above, all of the above papers use the sample mean to estimate the population mean, even though sample mean is not necessarily the best estimator. To improve things even further, Martinsek (1984) has proposed sequential procedures for choosing an estimator (from the class of trimmed means) as well as sample size, when the distribution of the observations is symmetric about the parameter of interest. At this point, it should be mentioned that results of asymptotic risk efficiency, in symmetric distribution case, have also been obtained by Sen (1981) and Jureckova and Sen (1982) using M -, L -, and R -estimators of location, rather than the sample mean. In an elegant work by Sen and Ghosh (1981), asymptotic risk efficiency has been established for sequential estimation of

estimable parametric functions using U-statistics based on i.i.d observations. This result, indeed, generalizes the results of Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981). In a paper by Finster (1983), independent observations are generated via a general linear model with normal errors and the slope parameter is estimated sequentially using the least squares estimator. Using the normality of errors and independence in the model, he shows that the regret of his sequential estimator is bounded. As a last reference we cite a paper by Lai and Siegmund (1983), where, for a first order, non-explosive, autoregressive process with unknown autoregressive parameter $\beta \in [-1, 1]$, it is shown that if the data are collected according to a particular stopping rule, then the least squares estimator of β is asymptotically normally distributed uniformly in β . It should, perhaps, be pointed out here that Lai and Siegmund (1983) do not discuss the sequential point estimation of β .

In this thesis we consider the problem of sequential estimation of parameters in a first order autoregressive model. More precisely, consider the first order, non-explosive, autoregressive process

$$(0.1) \quad X_1 - \mu = \beta(X_{1-1} - \mu) + \varepsilon_1, \quad i = 1, 2, \dots,$$

where $-\infty < \mu < \infty$, $|\beta| < 1$ and $\varepsilon_1, \varepsilon_2, \dots$ are i.i.d. according to some unknown d.f. F , with $E\varepsilon_1 = 0$ and $0 < E\varepsilon_1^2 = \sigma^2 < \infty$. The initial state X_0 is a random

variable (r.v.), which is assumed to be independent of $\{\varepsilon_i, i \geq 1\}$, with $0 < EX_0^2 < \infty$. The first problem considered in this thesis deals with sequential estimation of the location parameter μ . The second problem deals with the sequential estimation of the autoregressive parameter β when μ is known.

Chapter 1 discusses the sequential estimation of μ in (0.1). Section 1.1 discusses the sequential point estimation of μ , using the sample mean $\bar{X}$ and the loss function that is squared error plus cost per observation. Sequential fixed width confidence interval for μ is also given in this section. Theorem 1.1 gives the asymptotic risk efficiency of the sequential point estimator and the asymptotic normality of the standardized sequential point estimator. Theorem 1.2 gives the asymptotic consistency and efficiency of the sequential fixed width interval estimator. The proofs of these theorems are given in Section 1.2.

Chapter 2 discusses the sequential point estimation of β in (0.1) when μ is known. In Section 2.1, the parameter β is estimated sequentially using the least squares estimator subject to a loss function that is squared error plus sample size. Here the cost per observation is assumed to be one, but there is a penalty for estimation error. In Section 2.2, Theorem 2.1 gives the asymptotic

risk efficiency of the sequential estimator of β . In Section 2.3, Theorem 2.2 gives a second order expansion for the expected stopping time and Theorem 2.3 gives the asymptotic normality and the uniform integrability of the standardized stopping time. Section 2.4 contains the 'regret' analysis of the proposed sequential estimator of β .

Although the proofs of the above theorems use methodologies similar to those of Sen and Ghosh (1981), Chow and Yu (1981), Chow and Martinsek (1982) and Woodroffe (1982), it should be noted that, due to lack of independence between observations, none of the results in these papers can be directly applied to prove these theorems. However, in these models we can exploit the basic structure to provide us with some martingales and reverse martingales, which enable us to obtain uniform integrability of various stopped sums that are so crucial for the analysis. Here we also use some of the almost sure convergence results of Lai and Siegmund (1983).

Finally, we conclude this introduction by listing some of the intermediate results which we obtain as a by-product of our analysis and which are interesting in their own way:

(a) In Chapter 1, we obtain (non-sequential) L_p -rates of convergence for the sample mean and the autocorrelations, (see Lemma 1.2), which are extremely crucial for Theorem 1.1 and which, I believe, are not available in the literature.

(b) In Chapter 2, we obtain the uniform integrability of the standardized least squares estimator, which is one of the key tools for all of the theorems in this chapter (see Lemma 2.2). We also obtain some results which guarantee that the standardized least squares estimator satisfies the well known Anscombe's condition, which in turn enables us to establish the asymptotic normality of the standardized and stopped least squares estimator. (See Lemma 2.5, Corollary 2.1 and the equation (2.2.36)).

(c) In Section 2.3, we state and prove a lemma which is a slight extension of Theorem 2 of Hagwood and Woodroffe (1982) and Theorem 3 of Lai and Siegmund (1979) (see Lemma 2.7).

(d) In Section 2.4, we obtain L_p - boundedness of certain standardized and stopped martingale and stationary sum. (See Lemma 2.15 and 2.16). We also obtain the first and second moments of certain stopped martingales (see Lemma 2.17) which are similar to the Lemma in Chow and Martinsek (1982).

CHAPTER 1

SEQUENTIAL ESTIMATION OF THE MEAN OF A FIRST ORDER AUTOREGRESSIVE PROCESS

1.1 INTRODUCTION

Consider the first-order, non-explosive, autoregressive model

$$(1.1.1) \quad X_i - \mu = \beta(X_{i-1} - \mu) + \varepsilon_i, \quad i = 1, 2, \dots,$$

where $-\infty < \mu < \infty$, $|\beta| < 1$ and $\varepsilon_1, \varepsilon_2, \dots$ are i.i.d. according to some unknown d.f. F , with $E\varepsilon_1 = 0$ and $0 < E\varepsilon_1^2 = \sigma^2 < \infty$. The initial state X_0 is a random variable (r.v.), which is assumed to be independent of $\{\varepsilon_i, i \geq 1\}$, with $0 < EX_0^2 < \infty$.

Note that the model above can be written as

$$(1.1.2) \quad X_i - \mu = \beta^i(X_0 - \mu) + \sum_{j=1}^i \beta^{i-j} \varepsilon_j, \quad i = 1, 2, \dots$$

POINT ESTIMATION OF μ . Given a sample of size n , one wishes to estimate μ by the sample mean $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$, subject to the loss function,

$$(1.1.3) \quad L_n = a(\bar{X}_n - \mu)^2 + cn, \quad a > 0, \quad c > 0,$$

where c is the cost per observation. The object is to minimize the risk in estimation by choosing an appropriate sample size.

Now, by Proposition 1.1 proved below,

$$(1.1.4) \quad E(\bar{X}_n - \mu)^2 = n^{-1} \sigma^2 / (1 - \beta)^2 + o(n^{-1}), \quad \text{as } n \rightarrow \infty.$$

Therefore, if σ and β are known, the risk,

$$(1.1.5) \quad R_n = E L_n = n^{-1} a \sigma^2 / (1-\beta)^2 + cn + o(n^{-1})$$

is approximately minimized by the BFSP

$$(1.1.6) \quad n_0 \simeq c^{-1/2} a^{1/2} \sigma / (1-\beta)$$

with corresponding minimum risk

$$(1.1.7) \quad R_{n_0} \simeq 2cn_0.$$

However, if either σ or β are unknown the BFSP cannot be used, and there is no fixed sample size procedure that will achieve the risk R_{n_0} . For this case, we now describe a sequential procedure for choosing a sample size whose risk will be close to R_{n_0} for small c .

Let $m (\geq 3)$ be an initial sample size, $h(>0)$ be a suitable constant to be defined later on and define the stopping rule N , in analogy with n_0 , by

$$(1.1.8) \quad N = \inf \{n \geq m: n \geq c^{-1/2} a^{1/2} [\hat{\sigma}_n / |1-\hat{\beta}_n| + n^{-h}]\},$$

where,

$$(1.1.9) \quad \hat{\beta}_n = \frac{n-1}{\sum_{i=1}^{n-1} (X_i - \bar{X}_n)(X_{i+1} - \bar{X}_n)} / \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad \text{and}$$

$$\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 [1 - \hat{\beta}_n^2].$$

The proposed sequential point estimator of μ is $\bar{X}_N$ and its risk is

$$(1.1.10) \quad R_N = a E(\bar{X}_N - \mu)^2 + c E N.$$

It will be proved in Section 1.2 that this sequential procedure is "asymptotically risk efficient" with respect to the procedure that uses the best fixed sample size n_0 , i.e., as $c \rightarrow 0$, $R_N/R_{n_0} \rightarrow 1$.

INTERVAL ESTIMATION OF μ . The problem here is to find a confidence interval for μ of prescribed width $2d$ and coverage probability $1 - \alpha$ ($0 < \alpha < 1$). In order to motivate our procedure, observe from (1.2.1) below that,

$$(1.1.11) \quad \sqrt{n}(\bar{X}_n - \mu) \rightarrow_d N(0, \sigma^2/(1-\beta)^2) \quad \text{as } n \rightarrow \infty,$$

by CLT for i.i.d r.v.'s and the Slutsky theorem. Based on this, when σ and β are known, if one uses the confidence interval

$$(1.1.12) \quad I_{k_0} = [\bar{X}_{k_0} - d, \bar{X}_{k_0} + d],$$

where

$$k_0 \simeq d^{-2} Z_{1-\alpha}^2 \sigma^2 / (1-\beta)^2$$

and $Z_{1-\alpha}$ satisfies

$$(2\pi)^{-1/2} \int_{-Z_{1-\alpha}}^{Z_{1-\alpha}} \exp(-u^2/2) du = 1 - \alpha,$$

then as $d \rightarrow 0$,

$$k_0 \rightarrow \infty,$$

and

$$\begin{aligned} P[\mu \in I_{k_0}] &= P[\sqrt{k_0}|\bar{X}_{k_0} - \mu|/\gamma \leq d\sqrt{k_0}/\gamma] \\ &\rightarrow 1 - \alpha \quad (\text{asymptotic consistency}), \end{aligned}$$

where $\gamma = \sigma/(1-\beta)$. Furthermore, k_0 is (asymptotically) the smallest sample size that is asymptotically consistent.

As in the point estimation case, in practice one does not know σ or β . However, for

$$(1.1.13) \quad T = \inf \{n \geq m: n \geq d^{-2} Z_{1-\alpha}^2 \left[\hat{\sigma}_n^2 / (1 - \hat{\beta}_n)^2 + n^{-h} \right] \}$$

it will be proved in Section 1.2 that, as $d \rightarrow 0$,

$$P[\mu \in I_T] \rightarrow 1 - \alpha$$

and

$$E(T/k_0) \rightarrow 1 \quad (\text{asymptotic efficiency}).$$

The following theorems assess the performance of the above sequential procedures.

THEOREM 1.1. (POINT ESTIMATION). Suppose that
 $E|\xi_1|^{2p} < \infty$, $E|X_0|^{2p} < \infty$ for $p > 2$, and that
 $h \in (0, \frac{(p-2)^2}{2p})$ where h is as in (1.1.8). Then, as
 $c \rightarrow 0$,

$$(1.1.14) \quad N/n_0 \rightarrow 1 \quad \text{a.s.},$$

$$(1.1.15) \quad E|N/n_0 - 1| \rightarrow 0$$

$$(1.1.16) \quad \sqrt{N} (\bar{X}_N - \mu) \rightarrow_d \mathcal{N}(0, \sigma^2 / (1 - \beta)^2)$$

and

$$(1.1.17) \quad R_N / R_{n_0} \rightarrow 1.$$

THEOREM 1.2. (FIXED-WIDTH INTERVAL ESTIMATION). Under
the conditions of Theorem 1.1, as $d \rightarrow 0$

$$(1.1.18) \quad T/k_0 \rightarrow 1 \quad \text{a.s.},$$

$$(1.1.19) \quad P[\mu \in I_T] \rightarrow 1 - \alpha$$

and

$$(1.1.20) \quad E(T/k_0) \rightarrow 1,$$

where k_0 is as in (1.1.12).

REMARK: The stopping time T also appears in Subramanyam (1984), where (1.1.8) and (1.1.19) are proved, under somewhat different conditions. However, (1.1.20) is a new result.

1.2. PROOFS. The following representation and observation will be repeatedly used in the proofs below. First, from (1.1.2),

$$(1.2.1) \quad \bar{X}_n - \mu = [(X_0 - \mu)\beta(1 - \beta^n) + \sum_{j=1}^n \varepsilon_j(1 - \beta^{n-j+1})]/n(1 - \beta).$$

Secondly, (1.1.2), $E|\varepsilon_1|^p < \infty$, $E|X_0|^p < \infty$ for $p > 1$, and the Minkowski inequality yield

$$(1.2.2) \quad \|X_1 - \mu\|_p \leq \|X_0 - \mu\|_p + \|\varepsilon_1\|_p/(1 - |\beta|) < \infty, \quad i \geq 1.$$

Before proving Theorems 1.1 and 1.2 we state and prove a proposition which justifies (1.1.4).

PROPOSITION 1.1. For all $|\beta| < 1$, all real μ and
all $n \geq 1$.

$$\begin{aligned} n(1-\beta)^2 E(\bar{X}_n - \mu)^2 &= \sigma^2 + n^{-1}\beta^2(1-\beta^n)^2 E(X_0 - \mu)^2 \\ &\quad + n^{-1}\sigma^2(1-\beta^2)^{-1}(1-\beta^n)[\beta^{n+2} - 2\beta - \beta^2]. \end{aligned}$$

Proof. Square the representation (1.2.1) for $\bar{X}_n - \mu$ and use the independence of X_0 and $\{f_i, i \geq 1\}$ to get the required result. $\square$

For the sake of completeness, we state the following inequalities and a result which will be repeatedly used in this thesis.

A. Burkholder inequality. (See Chow and Teicher, 1978, Theorem 11.2.1.)

If $p \in (1, \infty)$ and $f = \{f_n; n \geq 1\}$ is an L_p -martingale, then with $A_p = [18p^{3/2}/(p-1)]^{-1}$ and $B_p = 18p^{3/2}/(p-1)^{1/2}$

$$A_p \|S_n(f)\|_p \leq \|f_n\|_p \leq B_p \|S_n(f)\|_p.$$

where $S_n(f) = \left[\sum_{i=1}^n (f_i - f_{i-1})^2 \right]^{1/2}$ and $f_0 = 0$. $\square$

B. Marcinkiewicz-Zygmund inequality. (See Chow and Teicher, 1978, Corollary 10.3.2.)

If $\{X_n; n \geq 1\}$ are i.i.d. with $EX_1 = 0$, $E|X_1|^p < \infty$ and $p \geq 2$, then $E \left| \sum_{i=1}^n X_i \right|^p = O(n^{p/2})$ as $n \rightarrow \infty$. $\square$

C. Doob's maximal inequality. (See Stout (1974) Corollary 3.3.2.)

Let $\{T_k; k \geq 1\}$ be a martingale or a nonnegative submartingale. If $p > 1$ and $E|T_k|^p < \infty$ for each

$k \geq 1$. then

$$E \max_{1 \leq k \leq n} |T_k|^p \leq \left(\frac{p}{p-1}\right)^p E |T_n|^p \quad \text{for each } n \geq 1. \quad \square$$

D. Result. (See Chow, Robbins, Siegmund (1970).

Example 2.1(d).)

Let $Y_1, Y_2, \dots$ be i.i.d. with $E|Y_1| < \infty$. Then

$\bar{Y}_n = \sum_{i=1}^n Y_i/n$ is a reverse martingale with respect to

$\mathcal{G}_n = \sigma\{\bar{Y}_n, \bar{Y}_{n+1}, \dots\}$. $\square$

E. Maximal inequality for the reverse martingale

$\{\bar{Y}_k; k \geq n\}$ (See Chow, Robbins and Siegmund (1970), display (4.39) and Stout (1974), (second) Lemma 3.3.1.)

Let $Y_1, Y_2, \dots$ be i.i.d. with $E|Y_1|^p < \infty$ for some $p > 1$. Then

$$E \sup_{k \geq n} |\bar{Y}_k|^p \leq \left(\frac{p}{p-1}\right)^p E |\bar{Y}_n|^p \quad \text{for each } n \geq 1. \quad \square$$

The proofs of Theorems 1.1 and 1.2 depend on a series of four lemmas, the first of which will be stated without proof. In the first three lemmas all the limits are as $n \rightarrow \infty$.

LEMMA 1.1. If f maps R^k into R^m and f is continuous at a , $s \in (0, \infty)$ and $\{X_n\}$ is a sequence of random k -vectors such that

$$P[\|X_n - a\| \geq \delta] = O(n^{-s}) \quad \text{for every } \delta > 0,$$

then

$$P[\|f(X_n) - f(a)\| \geq \epsilon] = O(n^{-s}) \quad \text{for every } \epsilon > 0. \quad \square$$

In order to state the next two lemmas, we need to introduce the following notations. Let $\bar{\epsilon}_n = n^{-1} \sum_{j=1}^n \epsilon_j$ and $\mathfrak{F}_n = \sigma\{X_0, \epsilon_1, \epsilon_2, \dots, \epsilon_n\}$ for $n \geq 1$. Define for $k = 0, 1$,

$$C_{k,n} = n^{-1} \sum_{i=1}^{n-k} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n), \quad \text{and} \quad (1.2.3)$$

$$C_{k,n}^* = n^{-1} \sum_{i=1}^{n-k} (X_i - \mu)(X_{i+k} - \mu).$$

Note that

$$(1.2.4) \quad \hat{\beta}_n = C_{1,n}/C_{0,n} \quad \text{and} \quad \hat{\sigma}_n^2 = C_{0,n}(1 - \hat{\beta}_n^2).$$

Lemma 1.2 below gives L_p rates of convergence of $\bar{X}_n$ to μ , and of the autocovariances $C_{0,n}$ and $C_{1,n}$ to $\sigma^2/(1-\beta^2)$ and $\beta\sigma^2/(1-\beta^2)$, respectively. Lemma 1.3 below deals with rates of convergence in probability of $\hat{\beta}_n$ to β and $\hat{\sigma}_n^2$ to σ^2 .

LEMMA 1.2. If $E|\epsilon_1|^{2p} < \infty$ and $E|X_0|^{2p} < \infty$, then

$$(1.2.5) \quad \|\bar{X}_n - \mu\|_{2p} = O(n^{-1/2}), \quad \text{if } p \geq 1,$$

$$(1.2.6) \quad \|C_{0,n} - \sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}), \quad \text{if } p \geq 2,$$

and

$$(1.2.7) \quad \|C_{1,n} - \beta\sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}), \quad \text{if } p \geq 2.$$

Proof: Assume w.l.o.g. that $\mu = 0$ and $\sigma^2 = 1$. From (1.2.1) and the Minkowski inequality

$$(1-\beta)\|\bar{X}_n\|_{2p} \leq n^{-1}\|X_0\|_{2p} + \|\bar{\varepsilon}_n\|_{2p} + \|\varepsilon_1\|_{2p} n^{-1} \sum_{j=1}^n |\beta|^{n-j+1} = O(n^{-1/2}),$$

by the Marcinkiewicz-Zygmund (M-Z) inequality (see Inequality B above) and $|\beta| < 1$.

As for the assertion (1.2.6), use the definitions (1.2.3) of $C_{0,n}$ and $C_{0,n}^*$ to write

$$(1.2.8) \quad C_{0,n} - \sigma^2/(1-\beta^2) = C_{0,n}^* - \sigma^2/(1-\beta^2) - \bar{X}_n^2.$$

Now, use (1.1.2) to write

$$(1.2.9) \quad C_{0,n}^* - \sigma^2/(1-\beta^2) = n^{-1} \sum_{i=1}^n [\beta^{2i} X_0^2 + \sum_{j=1}^i \beta^{2(i-j)} \varepsilon_j^2 - \sigma^2(1-\beta^2)] + 2 \sum_{1 \leq j < k \leq i} \beta^{2i-j-k} \varepsilon_j \varepsilon_k + 2X_0 \beta^i \left(\sum_{j=1}^i \beta^{i-j} \varepsilon_j \right).$$

As for the first term in (1.2.9), use $|\beta| < 1$ to get

$$(1.2.10) \quad n^{-1} \left\| \sum_{i=1}^n \beta^{2i} X_0^2 \right\|_p = O(n^{-1/2}).$$

As for the second and third terms in (1.2.9), interchange of the order of summation and $|\beta| < 1$ yields,

$$\begin{aligned} n^{-1} \sum_{i=1}^n \sum_{j=1}^i \beta^{2(i-j)} \varepsilon_j^2 - \sigma^2/(1-\beta^2) &= n^{-1} \sum_{j=1}^n \varepsilon_j^2 (1-\beta^{2(n-j+1)})/(1-\beta^2) - \sigma^2/(1-\beta^2) \\ &= (1-\beta^2)^{-1} n^{-1} \left[\sum_{j=1}^n (\varepsilon_j^2 - \sigma^2) - \sum_{j=1}^n \beta^{2(n-j+1)} \varepsilon_j^2 \right]. \end{aligned}$$

An application of the M - Z inequality yields

$$(1.2.11) \quad n^{-1} \left\| \sum_{j=1}^n (\varepsilon_j^2 - \sigma^2) \right\|_p = O(n^{-1/2}).$$

and the Minkowski inequality and $|\beta| < 1$ yields

$$(1.2.12) \quad n^{-1} \left\| \sum_{j=1}^n \beta^{2(n-j+1)} \varepsilon_j^2 \right\|_p = O(n^{-1}).$$

As for the fourth term in (1.2.9), interchange of the order of summation and change of indices yields

$$2n^{-1} \sum_{i=1}^n \sum_{1 \leq j < k \leq i} \beta^{2i-j-k} \varepsilon_j \varepsilon_k = 2n^{-1} \sum_{u=0}^{n-2} \sum_{v=u+1}^{n-1} \beta^{u+v} \sum_{i=v+1}^n \varepsilon_{i-u} \varepsilon_{i-v}.$$

Now, define for each u and v fixed, where $u < v \leq n-1$,

$$f_n = \sum_{i=v+1}^n \varepsilon_{i-u} \varepsilon_{i-v} \quad \text{and} \quad \mathcal{G}_n = \sigma\{\varepsilon_k; k \leq n-u\}.$$

Then the independence of $\{\varepsilon_i\}$ and the assumption that $E\varepsilon_1 = 0$ yield that $\{f_n, \mathcal{G}_n; n \geq v+1\}$ is a mean zero martingale. By the Burkholder inequality (see Inequality A above), the moment inequality and the independence of ε_{i-u} and ε_{i-v}

$$\begin{aligned} (n-v)^{-1/2} B_p^{-1} \|f_n\|_p &\leq \|[(n-v)^{-1} \sum_{i=v+1}^n \varepsilon_{i-u}^2 \varepsilon_{i-v}^2]^{1/2}\|_p \\ &\leq \|((n-v)^{-1} \sum_{i=v+1}^n |\varepsilon_{i-u} \varepsilon_{i-v}|^p)^{1/p}\|_p \\ &= \|\varepsilon_1\|_p^2. \end{aligned}$$

Therefore,

$$(1.2.13) \quad 2n^{-1} \left\| \sum_{i=1}^n \sum_{1 \leq j < k \leq i} \beta^{2i-j-k} \varepsilon_j \varepsilon_k \right\|_p = O(n^{-1/2}).$$

As for the fifth term in (1.2.9), interchange the order of summation and use $|\beta| < 1$, the Minkowski inequality and the independence of X_0 and $\{\varepsilon_i\}$ to get

$$\begin{aligned} 2n^{-1} \|X_0 \sum_{i=1}^n \sum_{j=1}^i \beta^{2i-j} \varepsilon_j\|_p &\leq 2(1-\beta^2)^{-1} (1-|\beta|)^{-1} n^{-1} \|X_0\|_p \|\varepsilon_1\|_p \\ (1.2.14) \quad &= O(n^{-1}). \end{aligned}$$

Therefore, from (1.2.9) to (1.2.14) and the Minkowski inequality it follows that

$$(1.2.15) \quad \|C_{0,n}^* - \sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}).$$

Hence the required result follows from (1.2.8), the Minkowski inequality and (1.2.5).

Finally, for the assertion (1.2.7), algebraic manipulations yield

$$(1.2.16) \quad C_{1,n} = C_{1,n}^* - \bar{X}_n^2 + n^{-1} \bar{X}_n [X_1 + X_n - \bar{X}_n].$$

Now, use (1.1.1) to write

$$\begin{aligned} C_{1,n}^* - \beta \sigma^2/(1-\beta^2) &= \beta [C_{0,n}^* - \sigma^2/(1-\beta^2)] + n^{-1} \sum_{i=1}^n X_{i-1} \varepsilon_i \\ &\quad - n^{-1} X_0 \varepsilon_1 + n^{-1} X_n \varepsilon_{n+1} - n^{-1} X_{n+1} X_n. \end{aligned}$$

Let $D_n = \sum_{i=1}^n X_{i-1} \varepsilon_i$. The independence of X_{i-1} and ε_i for each $i \geq 1$ and the assumption $E\varepsilon_i = 0$ imply that

$$(1.2.17) \quad \{D_n, \mathcal{F}_n; n \geq 1\} \text{ is mean zero martingale.}$$

By the Burkholder and the moment inequalities, the independence of ε_i and X_{i-1} and (1.2.2)

$$\begin{aligned} n^{-1/2} B_p^{-1} \|D_n\|_p &\leq \|(n^{-1} \sum_{i=1}^n X_{i-1}^2 \varepsilon_i^2)^{1/2}\|_p \\ &\leq \|(n^{-1} \sum_{i=1}^n |X_{i-1} \varepsilon_i|^p)^{1/p}\|_p \\ (1.2.18) \quad &= O(1). \end{aligned}$$

Also by the Minkowski inequality, the Schwarz inequality, the independence of X_{i-1} and ε_i and (1.2.2)

$$\|X_0 \varepsilon_1 + X_n \varepsilon_{n+1} - X_{n+1} X_n\|_p = O(1).$$

Therefore by the Minkowski inequality and (1.2.15)

$$(1.2.19) \quad \|C_{1,n}^* - \beta \sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}).$$

Hence the required result follows from (1.2.16), the Minkowski inequality, the Schwarz inequality, (1.2.2) (1.2.5) and (1.2.19). $\square$

LEMMA 1.3. If $E|\varepsilon_1|^{2p} < \infty$ $E|X_0|^{2p} < \infty$ for $p \geq 2$, then for every $\epsilon > 0$,

$$(1.2.20) \quad P[|C_{0,n} - \sigma^2/(1-\beta^2)| > \epsilon] = O(n^{-p/2}).$$

$$(1.2.21) \quad P[|C_{1,n} - \beta \sigma^2/(1-\beta^2)| > \epsilon] = O(n^{-p/2}),$$

and

$$(1.2.22) \quad P[|\hat{\sigma}_n^2/(1-\hat{\beta}_n)^2 - \sigma^2/(1-\beta)^2| > \epsilon] = O(n^{-p/2}),$$

where $\hat{\sigma}_n^2$ and $\hat{\beta}_n$ are as in (1.2.4).

Proof: Lemma 1.2 and the Markov inequality yield (1.2.20) and (1.2.21). By (1.2.4),

$$\hat{\sigma}_n^2(1-\hat{\beta}_n)^{-2} = C_{0,n}[1-(C_{1,n}/C_{0,n})^2]/[1-(C_{1,n}/C_{0,n})]^2.$$

Hence (1.2.22) follows from (1.2.20), (1.2.21) and Lemma 1.1 applied to the function $f(x,y) = x[1-(y/x)^2]/[1-(y/x)]^2$ continuous at $(\sigma^2/(1-\beta^2), \beta\sigma^2/(1-\beta^2))$. $\square$

The next lemma gives a rate on the tail behavior of the stopping rule N , which is crucial for the proof of Theorems 1.1 and 1.2.

Before we state the lemma we need to introduce the following notations. Let

$$(1.2.23) \quad b = (a/c)^{1/2}, \quad n_1 = [b^{1/(1+h)}],$$

$$n_2 = [n_0(1-\epsilon)] \quad \text{and} \quad n_3 = [n_0(1+\epsilon)] \quad \text{for} \quad 0 < \epsilon < 1.$$

Henceforth all unidentified limits are taken as $c \rightarrow 0$.

LEMMA 1.4. Assume the conditions of Theorem 1.1. Then for every $\epsilon > 0$ and for $s = p/2$.

$$(1.2.24) \quad P[N \leq n_2] = O(c^{(s-1)/[2(1+h)]}),$$

and

$$(1.2.25) \quad \sum_{n \geq n_3} P[N > n] = O(c^{(s-1)/2}).$$

Proof. From (1.1.8), $N \geq n_1$. Moreover,

$$P[N \leq n_2] \leq P[\hat{\sigma}_n / |1 - \hat{\beta}_n| \leq b^{-1}n \text{ for some } n_1 \leq n \leq n_2]$$

$$\leq P[\hat{\sigma}_n^2 / (1 - \hat{\beta}_n)^2 \leq b^{-2}n_2^2 \text{ for some } n_1 \leq n \leq n_2]$$

$$\leq P[\hat{\sigma}_n^2 / (1 - \hat{\beta}_n)^2 \leq (1 - \epsilon)^2 \sigma^2 / (1 - \beta)^2 \text{ for some } n_1 \leq n \leq n_2] \quad (\text{by (1.1.6)})$$

$$(1.2.26) \quad \leq \sum_{n=n_1}^{\infty} P[|\hat{\sigma}_n^2 / (1 - \hat{\beta}_n)^2 - \sigma^2 / (1 - \beta)^2| \geq \epsilon(2 - \epsilon)\sigma^2 / (1 - \beta)^2].$$

Now use Lemma 1.3 and (1.2.23) to get (1.2.24).

As for (1.2.25), it follows from (1.1.8) that, for

$$n \geq n_3$$

$$P[N > n] \leq P[\hat{\sigma}_n / |1 - \hat{\beta}_n| > b^{-1}n - n^{-h}]$$

$$(1.2.27) \quad \leq P[\hat{\sigma}_n / |1 - \hat{\beta}_n| - \sigma / |1 - \beta| > b^{-1}(n_3 - n_0) - n_3^{-h}]$$

Choose 'c' small enough so that

$$\epsilon \sigma / (1 - \beta) - \{(1 - \beta)c^{1/2} / [\sigma a^{1/2}(1 + \epsilon)]\}^h > \sigma \epsilon / 2(1 - \beta).$$

Then

$$P[N > n] \leq P[\hat{\sigma}_n / |1 - \hat{\beta}_n| - \sigma / |1 - \beta| > \sigma \epsilon / 2(1 - \beta)]$$

$$\leq P[|\hat{\sigma}_n^2 / |1 - \hat{\beta}_n|^2 - \sigma^2 / (1 - \beta)^2| > \sigma^2 \epsilon^2 / 4(1 - \beta)^2].$$

Now argue as for (1.2.24) to conclude (1.2.25). $\square$

Before we proceed to prove Theorem 1.1 we quote a definition and a lemma from Woodroffe (1982), which are not only useful here but also in Chapter 2.

DEFINITION 1.2. A sequence ξ_n , $n \geq 1$ of random variables is said to be uniformly continuous in probability (u.c.i.p.) if for every $\epsilon > 0$ there is a $\delta > 0$ for which

$$(i) \quad P\left\{\max_{0 \leq k \leq n\delta} |\xi_{n+k} - \xi_n| > \epsilon\right\} < \epsilon \quad \text{for all } n \geq 1.$$

The u.c.i.p. condition is often called Anscombe's condition.

Note 1. $\xi_n \rightarrow \ell(<\infty)$ a.s., as $n \rightarrow \infty$, implies (i).

LEMMA 1.4W. (See Woodroffe (1982), p. 10).

If Y_n , $n \geq 1$, and Z_n , $n \geq 1$, 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 (tight), and if ϕ is any continuous function on $\mathbb{R}^2$, then $\phi(Y_n, Z_n)$, $n \geq 1$, is u.c.i.p..

In what follows, $A = [n_2 < N < n_3]$, $B = [N \leq n_2]$ and $D = [N \geq n_3]$. Also, 1_F and $\bar{F}$ denote the indicator and the complement of a set F , respectively.

PROOF OF THEOREM 1.1. From Lemma 1.3 and the Borel-Cantelli Lemma, $C_{1,n} \rightarrow \beta\sigma^2/(1-\beta^2)$ a.s., and

$C_{0,n} \rightarrow \sigma^2/(1-\beta^2)$ a.s., as $n \rightarrow \infty$. From this and (1.2.4)

$$(1.2.28) \quad \hat{\beta}_n \rightarrow \beta \text{ a.s., and } \hat{\sigma}_n^2 \rightarrow \sigma^2 \text{ a.s., as } n \rightarrow \infty.$$

Therefore it follows that $N < \infty$ a.s. Also, since

$N \geq n_1$, $N \rightarrow \infty$ a.s.. Hence from (1.2.28)

$$(1.2.29) \quad \hat{\beta}_N \rightarrow \beta \text{ a.s. and } \hat{\sigma}_N^2 \rightarrow \sigma^2 \text{ a.s..}$$

But by definition

$$c^{-1/2} a^{1/2} \hat{\sigma}_N / |1 - \hat{\beta}_N| \leq N,$$

and

$$N \leq c^{-1/2} a^{1/2} \left[\hat{\sigma}_{N-1} / |1 - \hat{\beta}_{N-1}| + (N-1)^{-h} \right] + m.$$

Therefore from (1.2.29) and (1.1.6)

$$N/n_0 \rightarrow 1 \text{ a.s..}$$

As for the L_1 -convergence (1.1.15), observe that $(N/n_0 - 1)^- \leq 1$. Therefore by the dominated convergence theorem and (1.1.14)

$$E(N/n_0 - 1)^- \rightarrow 0.$$

Moreover, since $(N - n_0)^+ 1_B = 0$ we have

$$\begin{aligned} E(N/n_0 - 1)^+ &= E(N/n_0 - 1)^+ 1_A + \frac{1}{n_0} E(N - n_3)^+ + \frac{(n_3 - n_0)}{n_0} P(D) \\ &\leq 2\epsilon + o(1), \end{aligned}$$

by (1.2.25) and (1.2.23). Since ϵ is arbitrary, (1.1.15) follows.

As for (1.1.16), recall (1.2.1). By example 1.8 of Woodroffe (1982)

$$\{\sqrt{n} \bar{\epsilon}_n, n \geq 1\} \text{ is u.c.i.p.}$$

By the Markov inequality

$$\begin{aligned} \sum_{n=1}^{\infty} P(|n^{-1/2} \sum_{j=1}^n \beta^{n-j+1} \epsilon_j| > \epsilon) &\leq \epsilon^{-p} \sum_{n=1}^{\infty} E \left| \sum_{j=1}^n \beta^{n-j+1} \epsilon_j \right|^p / n^{p/2} \\ &< \infty, \end{aligned}$$

since $\| \sum_{j=1}^n \beta^{n-j+1} \varepsilon_j \|_p \leq K$ (constant), by the Minkowski inequality. Therefore by the Borel-Cantelli lemma

$$n^{-1/2} \sum_{j=1}^n \beta^{n-j+1} \varepsilon_j \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Also,

$$\beta(1-\beta^n)(1-\beta)^{-1}(X_0 - \mu)/\sqrt{n} \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

From these, Note 1 following Definition 1.2 and Lemma 1.4W it follows that

$$(1.2.30) \quad \{\sqrt{n}(\bar{X}_n - \mu), n \geq 1\} \text{ is u.c.i.p..}$$

The required result now follows from (1.1.14), (1.1.11) and the Anscombe's theorem (see Woodroffe (1982), Theorem 1.4).

It remains to show that N is asymptotically risk efficient. Assume without loss of generality that $\mu = 0$ and $\sigma^2 = 1$.

Write

$$R_N/R_{n_0} \simeq \{a E \bar{X}_N^2 + c E N\}/2cn_0.$$

From (1.1.15), it suffices to show that,

$$(1.2.31) \quad a E \bar{X}_N^2/cn_0 \rightarrow 1.$$

In order to establish (1.2.31) it suffices to show that

$$(1.2.32) \quad a E \bar{X}_N^2 I_B /cn_0 \rightarrow 0$$

and

$$(1.2.33) \quad a E \bar{X}_N^2 1_{AUD}/cn_0 \rightarrow 1.$$

In what follows we will often use the fact that $(cn_0)^{-1} = O(c^{-1/2})$ and $cn_0^2 = O(1)$.

Now, clearly from (1.2.1)

$$(1.2.34) \quad \bar{X}_n^2 \leq 4(1-\beta)^{-2} [(X_0/n)^2 + \bar{\varepsilon}_n^2 + (\sum_{j=1}^n \beta^{n-j+1} \varepsilon_j/n)^2].$$

As for (1.2.32), by the Holder inequality and $N \geq n_1$ we have

$$\begin{aligned} (cn_0)^{-1} E(X_0/N)^2 1_B &\leq (cn_0)^{-1} n_1^{-2} \|X_0\|_p^2 \{P(B)\}^{1-2/p} \\ &= \|X_0\|_p^2 O(c^{-1/2}) c^{1/(1+h)} [O(c^{(p-2)/4(1+h)})]^{(p-2)/p} \\ (1.2.35) \quad &\rightarrow 0, \end{aligned}$$

since $h < [(p-2)^2/2p]$. By the Holder inequality, the maximal inequality for the reverse martingale $\{\bar{\varepsilon}_n, n \geq n_1\}$ (see Result D and E above), the M-Z inequality and (1.2.24),

$$\begin{aligned} (cn_0)^{-1} E \max_{n_1 \leq n \leq n_2} \bar{\varepsilon}_n^2 1_B &\leq (cn_0)^{-1} \max_{n_1 \leq n \leq n_2} \bar{\varepsilon}_n^2 \cdot \{P(B)\}^{1-2/p} \\ &\leq (cn_0)^{-1} \left(\frac{p}{p-1}\right)^2 \|\bar{\varepsilon}_{n_1}\|_p^2 \{P(B)\}^{(p-2)/p} \\ &= O(c^{-1/2}) O(n_1^{-1}) \cdot O(c^{(p-2)^2/4p(1+h)}) \\ (1.2.36) \quad &\rightarrow 0. \end{aligned}$$

Furthermore by the Schwarz inequality and $|\beta| < 1$, for all $n \geq 1$,

$$\begin{aligned} |n^{-1} \sum_{j=1}^n \beta^{n-j+1} \varepsilon_j|^2 &\leq (n^{-1} \sum_{j=1}^n \beta^{2(n-j+1)}) (n^{-1} \sum_{j=1}^n \varepsilon_j^2). \\ (1.2.37) \quad &\leq n^{-1} (1-\beta^2)^{-1} (n^{-1} \sum_{j=1}^n \varepsilon_j^2). \end{aligned}$$

Therefore,

$$\begin{aligned} (1-\beta^2)(cn_0)^{-1} E(N^{-1} \sum_{j=1}^N \beta^{N-j+1} \varepsilon_j)^2 1_B \\ \leq (cn_0)^{-1} n_1^{-1} E \max_{n_1 \leq n \leq n_2} (n^{-1} \sum_{j=1}^n \varepsilon_j^2) 1_B \end{aligned}$$

$$\begin{aligned}
&\leq (cn_0)^{-1} n_1^{-1} \max_{n_1 \leq n \leq n_2} n^{-1} \sum_{j=1}^n \varepsilon_j^2 \|_{p/2} \{P(B)\}^{1-2/p} \\
&\leq \left(\frac{p}{p-2}\right) O(c^{-1/2}) c^{1/2(1+h)} \|_{n_1}^{-1} \sum_{j=1}^{n_1} \varepsilon_j^2 \|_{p/2} O(c^{(p-2)^2/4p(1+h)}) \\
(1.2.38) \quad &\rightarrow 0.
\end{aligned}$$

by the Holder inequality, the maximal inequality for the reverse martingale $\{n^{-1} \sum_{j=1}^n \varepsilon_j^2, n \geq n_1\}$, the Minkowski inequality, (1.2.24). Combining (1.2.34) with (1.2.35), (1.2.36) and (1.2.38) proves (1.2.32).

As for (1.2.33), since $n_0^2 \approx (a/c)(1-\beta)^{-2}$ (by (1.1.6)) and $P(AUD) \rightarrow 1$ (by Lemma 1.4), it follows from (1.1.16) and (1.1.14) that

$$(1.2.39) \quad a \bar{X}_N^2 1_{AUD}/cn_0 \rightarrow_d \chi_1^2.$$

where χ_1^2 is a chi-square r.v. with one degree of freedom. Furthermore,

$$\begin{aligned}
&(cn_0)^{-1} E(X_0/N)^2 1_{AUD} \leq (cn_0)^{-1} n_2^{-2} EX_0^2 \\
(1.2.40) \quad &\rightarrow 0.
\end{aligned}$$

Also, since

$$\begin{aligned}
\bar{\varepsilon}_N^2 1_{AUD}/cn_0 &\leq (cn_2^2)^{-1} \left(\sum_{i=1}^N \varepsilon_i / \sqrt{n_0} \right)^2 \\
&= O(1) \left(\sum_{i=1}^N \varepsilon_i / \sqrt{n_0} \right)^2.
\end{aligned}$$

we have that

$$(1.2.41) \quad \{\bar{\varepsilon}_N^2 1_{AUD}/cn_0\} \text{ is U.I.}$$

by Lemma 5 of Chow and Yu since $\{N/n_0\}$ is U.I. (by (1.1.15)). Finally, by (1.2.37) and arguments similar to

(1.2.38) we have, for $s > 1$,

$$\begin{aligned}
 (1-\beta^2)^{-s} (cn_0)^{-s} E(N^{-1} \sum_{j=1}^N \beta^{N-j+1} \xi_j)^{2s} 1_{\text{AUD}} \\
 \leq \left(\frac{s}{s-1}\right) E|\xi_1|^{2s} (cn_0)^{-s} n_2^{-s} \\
 (1.2.42) \qquad \qquad \qquad = O(1).
 \end{aligned}$$

Combining (1.2.34) with (1.2.40), (1.2.41) and (1.2.42) we have that

$$(1.2.43) \qquad \{a \bar{X}_N^2 1_{\text{AUD}} / cn_0\} \text{ is U.I.}$$

The required result now follows from (1.2.39) and (1.2.43).

Hence the theorem. $\square$

PROOF OF THEOREM 1.2. Since the stopping times N and T are similar, arguments given for assertions (1.1.14), (1.1.15) and (1.1.16) also show that

$$T/k_0 \rightarrow 1 \text{ a.s.},$$

$$E(T/k_0) \rightarrow 1 \text{ a.s.}, \text{ as } d \rightarrow 0.$$

and

$$\sqrt{T}(\bar{X}_T - \mu)/\gamma \rightarrow_d N(0,1), \text{ as } d \rightarrow 0,$$

where $\gamma = \sigma/(1-\beta)$. Therefore,

$$\begin{aligned}
 P[\mu \in I_T] &= P[\sqrt{T}|\bar{X}_T - \mu|/\gamma \leq d\sqrt{T}/\gamma] \\
 &\rightarrow 1 - \alpha, \text{ as } d \rightarrow 0.
 \end{aligned}$$

Hence the Theorem. $\square$

CHAPTER 2

SEQUENTIAL ESTIMATION OF THE AUTOREGRESSIVE PARAMETER IN A FIRST ORDER AUTOREGRESSIVE PROCESS

2.1 INTRODUCTION. In this chapter we consider the sequential estimation of β in the model (1.1.1) when μ is assumed to be known, and hence taken to be 0. For convenient reference later on, we restate this modified model here. Thus one observes a process $\{X_i\}$ such that

$$(2.1.1) \quad X_i = \beta X_{i-1} + \varepsilon_i, \quad i = 1, 2, \dots,$$

where β , X_0 and $\varepsilon_1, \varepsilon_2, \dots$ are as in (1.1.1).

Given a sample of size n , one wishes to estimate β by the least squares estimator

$$(2.1.2) \quad \begin{aligned} \hat{\beta}_n &= \left(\sum_{i=1}^n X_{i-1} X_i \right) / \left(\sum_{i=1}^n X_{i-1}^2 \right) \\ &= \beta + \left(\sum_{i=1}^n X_{i-1} \varepsilon_i \right) / \left(\sum_{i=1}^n X_{i-1}^2 \right), \end{aligned}$$

subject to the loss function

$$(2.1.3) \quad L_n = A n^{-1} \sum_{i=1}^n X_{i-1}^2 (\hat{\beta}_n - \beta)^2 + n,$$

where $A (>0)$ is the cost of estimation error. The object is to minimize the risk in estimation by choosing an appropriate sample size.

Recall from Theorems 4.2 and 4.3 of Anderson (1959) that for

$$(2.1.4) \quad \sqrt{\sum_{i=1}^n X_{i-1}^2} (\hat{\beta}_n - \beta) \rightarrow_d \mathcal{N}(0, \sigma^2), \quad \text{as } n \rightarrow \infty.$$

Moreover, by Lemma 2.2 below, under certain regularity conditions, $\{ \sum_{i=1}^n X_{i-1}^2 (\hat{\beta}_n - \beta)^2; n \geq 1 \}$ is uniformly integrable (U.I). Therefore, from this and (2.1.4)

$$(2.1.5) \quad E \sum_{i=1}^n X_{i-1}^2 (\hat{\beta}_n - \beta)^2 = \sigma^2 + o(1), \quad \text{as } n \rightarrow \infty.$$

Now, if σ and β are known, then the risk

$$(2.1.6) \quad R_n = E L_n = \Lambda n^{-1} \sigma^2 + n + o(n^{-1})$$

is approximately minimized by the BFSP

$$(2.1.7) \quad n_0 \simeq \Lambda^{1/2} \sigma$$

with the corresponding minimum risk

$$(2.1.8) \quad R_{n_0} \simeq 2\Lambda^{1/2} \sigma.$$

However, if σ is unknown then the BFSP cannot be used, and there is no fixed sample size procedure that will achieve the risk R_{n_0} .

For this case, we define a stopping rule t , in analogy with n_0 , by

$$(2.1.9) \quad t = \inf\{n \geq n_A: n \geq \Lambda^{1/2} \hat{\sigma}_n\}$$

where n_A is an initial sample size depending on Λ and

$\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (X_i - \hat{\beta}_n X_{i-1})^2$. The proposed sequential point estimator of β is $\hat{\beta}_t$ and its risk is $E L_t$.

It will be proved in Section 2.2 that this sequential procedure is asymptotically risk efficient. In Section 2.3 a second order approximation for the expected stopping time is obtained. In Section 2.4 the regret $R_t - R_{n_0}$ is analysed.

2.2 RISK EFFICIENCY.

THEOREM 2.1. Assume, for $s > 2$, that $E|\varepsilon_1|^{4s} < \infty$, $E|X_0|^{4s} < \infty$ and $E|1/(\varepsilon_1^2 + \dots + \varepsilon_M^2)|^{2s} < \infty$ for some positive integer M . In addition, if n_A in (2.1.9) is such that $A^{1/2(1+\eta)} \leq n_A = o(A^{1/2})$ with $\eta \in (0, s-2)$, then, as $A \rightarrow \infty$,

$$(2.2.1) \quad t/A^{1/2}\sigma \rightarrow 1, \quad \text{a.s.}$$

$$(2.2.2) \quad E \, t/A^{1/2}\sigma \rightarrow 1$$

and

$$(2.2.3) \quad R_t/R_{n_0} \rightarrow 1.$$

The proof of Theorem 2.1 depends on a series of lemmas, the first two of which give the uniform integrability of $\{\sum_{i=1}^n X_{i-1}^2 (\hat{\beta}_n - \beta)^2; n \geq 1\}$, which was extremely crucial to compute $E L_n$. Before we state the lemmas, we need to introduce the following notations. Let,

$$(2.2.4) \quad D_n = \sum_{i=1}^n X_{i-1} \varepsilon_i, \quad I_n = \sum_{i=1}^n X_{i-1}^2$$

$$\bar{I}_n = I_n/n \quad \text{and} \quad C_n = D_n^2/(nI_n).$$

Note that

$$(2.2.5) \quad \hat{\beta}_n - \beta = D_n/I_n, \quad n^{-1}I_n(\hat{\beta}_n - \beta)^2 = C_n, \quad \text{and}$$

$$\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n \varepsilon_i^2 - C_n.$$

Henceforth all unidentified limits are taken as

$A \rightarrow \infty$.

LEMMA 2.1. If $E|\varepsilon_1|^p < \infty$ and $E|X_0|^p < \infty$ for
 $p \geq 2$. then.

$$(2.2.6) \quad \|D_n\|_p = O(n^{1/2}) \text{ as } n \rightarrow \infty.$$

Proof. The required result follows from (1.2.17) and (1.2.18). $\square$

LEMMA 2.2. If $s \geq 1$, and $E|\varepsilon_1|^{4s} < \infty$, $E|X_0|^{4s} < \infty$
and $E|1/(\varepsilon_1^2 + \dots + \varepsilon_M^2)|^{2s} < \infty$ for some positive integer M ,
then

$$(2.2.7) \quad \sup_{n \geq 1} E|I_n(\hat{\beta}_n - \beta)^2|^s < \infty.$$

Hence, $\{|I_n(\hat{\beta}_n - \beta)^2|^q, n \geq 1\}$ is U.I. for all $q < s$.

Proof. From (2.2.5),

$$(2.2.8) \quad I_n(\hat{\beta}_n - \beta)^2 = D_n^2 / I_n.$$

Now, use (2.1.1) to get, for any $\ell > 1$,

$$\sum_{i=1}^{\ell-1} \varepsilon_i^2 = \sum_{i=1}^{\ell-1} (X_i - \beta X_{i-1})^2$$

$$(2.2.9) \quad \leq (1+|\beta|)^2 I_\ell$$

Using the above inequality, we will first show, for any fixed $m \geq 1$, that

$$(2.2.10) \quad E \sup_{k \geq 1} |k/I_{mk+1}|^{2s} < \infty.$$

Then use this to show that for M in the assumption,

$$(2.2.11) \quad E \sup_{n \geq M+1} |n/I_n|^{2s} < \infty.$$

By the Schwarz inequality

$$(2.2.12) \quad \|D_n^2/I_n\|_s \leq \|D_n^2/n\|_{2s} \|n/I_n\|_{2s}.$$

From Lemma 2.1,

$$(2.2.13) \quad \|D_n^2/n\|_{2s} = O(1), \text{ as } n \rightarrow \infty.$$

Therefore, the required result will follow from (2.2.11).

As for (2.2.10), it follows from (2.2.9) and the Jensen's inequality that for each $m \geq 1$,

$$(2.2.14) \quad \begin{aligned} k/I_{mk+1} &\leq (1+|\beta|)^2 k / \left(\sum_{i=1}^k B_{i,m} \right) \\ &\leq (1+|\beta|)^2 k^{-1} \sum_{i=1}^k (1/B_{i,m}), \end{aligned}$$

$$\text{where } B_{i,m} = \sum_{j=(i-1)m+1}^{im} \varepsilon_j^2.$$

Observe that for each fixed $m \geq 1$, $\{1/B_{i,m}; i \geq 1\}$ forms an i.i.d sequence. Now, by Result D (see Section 1.2)

$$(2.2.15) \quad \{k^{-1} \sum_{i=1}^k 1/B_{i,m}, \mathcal{G}_{k,m}; k \geq 1\} \text{ is a reverse martingale,}$$

$$\text{where } \mathcal{G}_{k,m} = \sigma\{\ell^{-1} \sum_{i=1}^{\ell} 1/B_{i,m}; \ell \geq k\}.$$

Therefore, an application of maximal inequality for reverse martingale (see Inequality E Section 1.2) yields,

$$(2.2.16) \quad E \sup_{k \geq 1} \left| k^{-1} \sum_{i=1}^k 1/B_{i,m} \right|^{2s} \leq \left(\frac{2s}{2s-1} \right)^{2s} E |1/B_{1,m}|^{2s} < \infty,$$

for $m \geq M$, by assumption.

As for (2.2.11), let $n \geq M + 1$. Then there exists $k \geq 1$ such that $kM + 1 \leq n \leq (k+1)M$. Now, use the fact that I_n is $\uparrow$ in n to get

$$(2.2.17) \quad \begin{aligned} n/I_n &\leq (k+1)M/I_{kM+1} \\ &\leq 2M \sup_{k \geq 1} (k/I_{kM+1}). \end{aligned}$$

Use this along with (2.2.14) and (2.2.16) to get (2.2.11).

Hence the lemma. $\square$

LEMMA 2.3. Assume that $E|\xi_1|^{2s} < \infty$ with $s \geq 1$ and
that $n_A = O(A^{1/2})$. Then

$$(2.2.18) \quad \{ |t/A^{1/2}\sigma|^s; A \geq 1 \} \text{ is U.I..}$$

Proof. Assume w.l.o.g. $\sigma^2 = 1$. Define another
 stopping time t' by

$$(2.2.19) \quad t' = \inf \{ n \geq n_A: A^{-1}n^2 \geq n^{-1} \sum_{i=1}^n \xi_i^2 \}.$$

Then, by Lemma 2 of Chow and Yu (1981) applied to
 $\{\xi_i^2; i \geq 1\}$, with a_n, α, b_n and λ in Chow and Yu (1981)
 replaced by $n^2, 2, 0$ and A^{-1} respectively, we have that

$$(2.2.20) \quad \{ |t'/A^{1/2}|^s; A \geq 1 \} \text{ is U.I..}$$

Now, since $t' \geq t$ we have the required uniform integra-
 bility. $\square$

The next lemma gives a rate of convergence on the tail
 behavior of the stopping time t .

LEMMA 2.4. Assume the moment condition of Lemma 2.2.
Then for $A^{1/2(1+\eta)} \leq n_A$ and $\epsilon > 0$,

$$(2.2.21) \quad P(t < (1-\epsilon)A^{1/2}\sigma) = O(A^{-(s-1)/2(1+\eta)}),$$

and

$$(2.2.22) \quad P(t > [(1+\epsilon)A^{1/2}\sigma] + 1) = o(A^{-(s-1)/2(1+\eta)}).$$

Proof. Assume w.l.o.g. $\sigma^2 = 1$. Let $s > 1, \epsilon < 1$ and
 $\delta = 1 - (1-\epsilon)^2$. Then by the definition of t in (2.1.9),

$$P(t < (1-\epsilon)A^{1/2}) = P(n^{-1} \sum_{i=1}^n \xi_i^2 - C_n \leq A^{-1}n^2 \\ \text{for some } n_A \leq n < (1-\epsilon)A^{1/2})$$

$$\begin{aligned}
&\leq P(n^{-1} \sum_{i=1}^n \xi_i^2 - C_n \leq (1-\epsilon)^2 \text{ for some } n_A \leq n) \\
&= P(n^{-1} \sum_{i=1}^n (1-\xi_i^2) + C_n \geq \delta \text{ for some } n_A \leq n) \\
&\leq P(n^{-1} \sum_{i=1}^n (1-\xi_i^2) \geq \delta/2 \text{ for some } n_A \leq n) \\
&\quad + P(C_n \geq \delta/2 \text{ for some } n_A \leq n)
\end{aligned}$$

$$(2.2.23) \quad \leq P(V \geq n_A \delta/4) + \sum_{n \geq n_A} P(C_n \geq \delta/2)$$

with $V = \sup_{n \geq 1} \left[\sum_{i=1}^n (1-\xi_i^2) - \delta/4 \right]^+$. Then by Theorem 4.13 of Chow, Robbins and Siegmund (1971),

$$E V^{s-1} < \infty.$$

Therefore by the moment inequality and $n_A \geq A^{1/2(1+\eta)}$

$$P(V \geq n_A \delta/4) \leq \left(\frac{4}{\delta}\right)^{(s-1)} n_A^{-(s-1)} \int_{\{V > n_A \delta/4\}} V^{(s-1)} dP$$

$$(2.2.24) \quad = o(A^{-(s-1)/2(1+\eta)}).$$

An application of the moment inequality yields

$$\begin{aligned}
(2.2.25) \quad \sum_{n \geq n_A} P(C_n \geq \delta/2) &\leq \left(\frac{2}{\delta}\right)^s \sum_{n \geq n_A} E C_n^s \\
&\leq \left(\frac{2}{\delta}\right)^s \left(\sum_{n \geq n_A} \frac{1}{n^s} \right) \sup_{n \geq 1} E(n C_n)^s \\
&= o(n_A^{-(s-1)}).
\end{aligned}$$

by Lemma 2.2. The required result now follows from (2.2.23), (2.2.24) and (2.2.25).

As for (2.2.22), recall the definition of t' in (2.2.19) and the observation that $t' \geq t$ a.s.. Let $N'' = [(1+\epsilon)A^{1/2}] + 1$. Then

$$P(t' > N'') \leq P((N'')^{-1} \sum_{i=1}^{N''} \xi_i^2 > A^{-1}(N'')^2)$$

$$\leq P((N'')^{-1} \sum_{i=1}^{N''} \xi_i^2 > (1+\epsilon)^2)$$

$$\leq P((N'')^{-1} \sum_{i=1}^{N''} (\xi_i^2 - 1) > (1+\epsilon)^2 - 1).$$

Let $h = (1+\epsilon)^2 - 1$. Then, for V defined above with ' δ ' replaced by ' $2h$ ' and $1 - \xi_i^2$ replaced by $\xi_i^2 - 1$, the moment inequality implies

$$P\left(\sum_{i=1}^{N''} (\xi_i^2 - 1 - h/2) > N''h/2\right) \leq P(V > N''h/2)$$

$$\begin{aligned} &\leq \left(\frac{2}{h}\right)^{s-1} (N'')^{-(s-1)} \int_{\{V > N''h/2\}} V^{s-1} dP \\ &= o(A^{-(s-1)/2}), \end{aligned}$$

by another usage of the Chow, Robbins and Siegmund theorem.

□

The following lemma and corollary deal with the well-known Anscombe's condition. These are not only used in the proof of Theorem 2.1 but also used in the next section for a deeper analysis of our stopping time t .

LEMMA 2.5. Assume the model (2.1.1). Then
 $\{D_n/\sqrt{n}, n \geq 1\}$. is stochastically bounded and u.c.i.p..

Proof. Stochastically boundedness follows from Lemma 2.1 and the Markov Inequality. To prove u.c.i.p., observe that for $k \geq 0, n \geq 1$

$$\left| \frac{D_{n+k}}{\sqrt{n+k}} - \frac{D_n}{\sqrt{n}} \right| \leq \frac{1}{\sqrt{n}} \left| D_{n+k} - D_n \right| + \left| (1 - \sqrt{n/(n+k)}) \right| \left| \frac{D_n}{\sqrt{n}} \right|$$

If $k \leq n\delta$, then the L_1 -norm of the second term on the right is bounded by

$$[1 - (1+\delta)^{-1/2}] \left\| \frac{D_n}{\sqrt{n}} \right\|_1.$$

which tends to zero as $\delta \rightarrow 0$, uniformly in $n \geq 1$, by Lemma 2.1.

For the first term, the Doob's maximal inequality applied to the martingale $\{D_n\}$ (see Inequality C, Section 1.2.) and the independence of X_j and ε_k for $j < k$ yield

$$P\left\{ \max_{0 \leq k \leq n\delta} |D_{n+k} - D_n| > \epsilon \frac{\sqrt{n}}{2} \right\} \leq \frac{16}{\epsilon^2 n} E |D_{n+[n\delta]} - D_n|^2$$

$$= \frac{16}{\epsilon^2 n} \sigma^2 \sum_{i=n+1}^{n+[n\delta]} EX_{i-1}^2$$

$$= O(\delta), \text{ as } \delta \rightarrow 0,$$

uniformly in $n \geq 1$, by (1.2.2). That $D_n/\sqrt{n}, n \geq 1$ is u.c.i.p. now follows easily. $\square$

COROLLARY 2.1. Assume the model (2.1.1). Then
 $\{D_n/I_n^{1/2}, n \geq 1\}$, is u.c.i.p. and stochastically bounded.

Proof. From Lemma 2.5 above, we have $D_n/\sqrt{n}, n \geq 1$,
 is u.c.i.p.. Moreover, from the proof of Theorem 4.1 of Lai
 and Siegmund (1983) we have

$$(2.2.26) \quad \bar{I}_n \rightarrow \sigma^2/(1-\beta^2) \quad \text{a.s., as } n \rightarrow \infty,$$

and therefore is u.c.i.p. and stochastically bounded. Hence
 the corollary follows from the Lemma 1.4W of Chapter 1. $\square$

Before we proceed to prove Theorem 2.1 we need to
 introduce the following notations. Let

$$(2.2.27) \quad N' = [(1-\epsilon)A^{1/2}\sigma], N'' = [(1+\epsilon)A^{1/2}\sigma] + 1$$

$$F = \{N' \leq t \leq N''\}.$$

PROOF OF THEOREM 2.1. Assume w.l.o.g that $\sigma^2 = 1$.
 From Theorem 4.2 of Lai and Siegmund (1983) we have that

$$\hat{\sigma}_n^2 \rightarrow 1 \quad \text{a.s., as } n \rightarrow \infty.$$

Use this, the definition of the stopping time t in
 (2.1.9), the fact that $t \geq A^{1/2(1+\eta)}$ and similar arguments
 as in the proof of Theorem 1.1 to get

$$(2.2.28) \quad \begin{aligned} t &< \infty \quad \text{a.s.} \\ t &\rightarrow \infty \quad \text{a.s.} \quad \text{and} \\ t/A^{1/2} &\rightarrow 1 \quad \text{a.s..} \end{aligned}$$

Moreover, because $\{t/A^{1/2}; A \geq 1\}$ is U.I (by Lemma 2.3),
 we have that

$$E(t/A^{1/2}) \rightarrow 1.$$

Therefore it only remains to show that t is asymptotically risk efficient.

For this write,

$$R_t / R_{n_0} \simeq \{A E C_t + Et\} / 2A^{1/2}.$$

From (2.2.2) it suffices to show

$$(2.2.29) \quad A^{1/2} E C_t \rightarrow 1.$$

For this, clearly it suffices to show that,

$$(2.2.30) \quad A^{1/2} E C_t 1_{\bar{F}_A} \rightarrow 0,$$

and

$$(2.2.31) \quad A^{1/2} E C_t 1_F \rightarrow 1.$$

Consider (2.2.30). Now,

$$\begin{aligned} A^{1/2} E C_t 1_{\{t < N'\}} &\leq A^{1/2} \sum_{n=n_A}^{N'-1} E C_n 1_{\{t = n\}} \\ &\leq A^{1/2} \sum_{n=n_A}^{N'-1} \|C_n\|_s P^{1-1/s} \{t=n\} \\ &\leq A^{1/2} \left(\sum_{n=n_A}^{N'} E C_n^s \right)^{1/s} P^{1-1/s} \{t < (1-\epsilon)A^{1/2}\} \\ &\leq A^{1/2} \sup_{n \geq 1} \|nC_n\|_s \left(\sum_{n=n_A}^{\infty} n^{-s} \right)^{1/s} P^{1-1/s} \{t < (1-\epsilon)A^{1/2}\} \\ &= A^{1/2} O(n_A^{-(s-1)/s}) O(A^{-(s-1)^2/2s(1+\eta)}) \quad (\text{by Lemma 2.2}) \\ &\rightarrow 0, \quad \text{since } 0 < \eta < s - 2. \end{aligned}$$

Similar but simpler arguments yield

$$(2.2.32) \quad A^{1/2} E C_t 1_{\{t > N''\}} \rightarrow 0.$$

As for (2.2.31), it suffices to show

$$(2.2.33) \quad \{A^{1/2} C_t 1_F, A \geq 1\} \text{ is U.I.}$$

and

$$(2.2.34) \quad A^{1/2} C_t 1_F \rightarrow_d \mathfrak{A}_1^2.$$

For (2.2.33), use I_n is $\uparrow$ in n , the Holder inequality and choose A large so that $n_A \geq M + 1$ to get

$$A^{s/2} E C_t^s 1_F \leq A^{s/2} E \max_{N' \leq n \leq N''} C_n^s$$

$$\leq \frac{A^{s/2}}{(N')^{2s}} E |N'/I_{N'}|^s \max_{N' \leq n \leq N''} D_n^{2s}$$

$$\leq \frac{A^{s/2}}{(N')^{2s}} \| \sup_{n \geq M} (n/I_n) \|_{2s}^s \| \max_{N' \leq n \leq N''} D_n \|_{4s}^{2s}$$

$$\leq \frac{A^{s/2}}{(N')^{2s}} O[(N'')^s]$$

$$(2.2.35) \quad = O(1),$$

by the Doob's maximal inequality for the martingale $\{D_n\}$ (see Inequality C, Section 1.2), Lemma 2.1 and (2.2.11).

As for (2.2.34), it follows from (2.1.4), Corollary 2.1, (2.2.1) and the Anscombe's theorem that

$$(2.3.36) \quad \sqrt{t} C_t \rightarrow_d |Z|,$$

where $Z \sim N(0,1)$. Therefore by (2.2.1), the Slutsky theorem and the fact that $1_F \rightarrow 1$ a.s. we have the required result. Hence the theorem. $\square$

2.3. A Second order expansion for the expected stopping time. Asymptotic normality and Uniform integrability of standardized stopping time.

One of the main results of this section is Theorem 2.2, where we obtain a second order expansion to the expected stopping time Et_A . Since the proof of this theorem has its roots in the non-linear renewal theory developed by Lai and Siegmund (1977, 1979) and Hagwood and Woodroffe (1982), we shall give a brief description of their papers.

In the i.i.d set up, Lai and Siegmund (1977, 1979) have developed the non-linear theory to analyse stopping times that stop the first time a perturbed random walk exceeds a boundary. Under the condition that the perturbation sequence is slowly changing (see below), besides some other regularity conditions, Lai and Siegmund (1979) have obtained a second order expansion to the expected stopping time, as the boundary tends to infinity. Later, Hagwood and Woodroffe (1982) have shown that the second order expansion obtained by Lai and Siegmund (1979) holds under an alternative set of regularity conditions. For an excellent survey of these results, see Woodroffe (1982).

In view of the existing literature mentioned above, we first show that our stopping time t also stops the first time a perturbed random walk exceeds a boundary. Here the perturbation sequence is also shown to be a slowly changing sequence. Once this is done, although many of the results

and methodologies developed by Lai and Siegmund (1977, 1979) and Hagwood and Woodroffe (1982) can be used, yet their main theorems which guarantee a second order expansion cannot be directly applied. However, a careful analysis of their proofs yields an extension of their main theorems. This extension is stated and proved as a Lemma below. Using this Lemma we prove Theorem 2.2.

Before we proceed with the analysis of stopping time t we shall give a definition of a slowly changing sequence (see Woodroffe (1982), Chapter 4).

DEFINITION 2.1. A sequence $\{F_n, n \geq 1\}$ is said to be slowly changing if:

(i) $\{F_n; n \geq 1\}$ is u.c.i.p (see Definition 1.2)

and

(ii) $n^{-1} \max_{1 \leq i \leq n} |F_i| \rightarrow 0$ in probability, as $n \rightarrow \infty$.

Note 2. $F_n/n \rightarrow 0$ a.s. as $n \rightarrow \infty$ implies (ii). See the Definition 1.2 for a sufficient condition for (i).

Now, for a deeper analysis of our stopping rule t defined in (2.1.9), write

$$(2.3.1) \quad t = \inf\{n \geq n_A: Z_n \geq A^{1/2}\sigma\}$$

where $Z_n = n(\sigma/\hat{\sigma}_n)$.

By Taylor's theorem

$$(2.3.2) \quad (\sigma^2/\hat{\sigma}_n^2)^{1/2} = 1 - (1/2)(\hat{\sigma}_n^2/\sigma^2 - 1) + (3/8)\lambda_n^{-5/2}(\hat{\sigma}_n^2/\sigma^2 - 1)^2,$$

where λ_n is a random variable lying between 1 and $\hat{\sigma}_n^2/\sigma^2$. Now substitute the representation (2.2.5) for $\hat{\sigma}_n^2$ in the second term of the R.H.S of (2.3.2) to get

$$(2.3.3) \quad Z_n = n - (n/2) \left[n^{-1} \sum_{i=1}^n (\xi_i^2/\sigma^2 - 1) - C_n/\sigma^2 \right] + (3/8)\lambda_n^{-5/2} n(\hat{\sigma}_n^2/\sigma^2 - 1)^2$$

$$= S_n + F_n.$$

where $S_n = \sum_{i=1}^n [1 - (1/2)(\xi_i^2/\sigma^2 - 1)],$

and

$$F_n = (nC_n/2\sigma^2) + (3/8)\lambda_n^{-5/2} n(\hat{\sigma}_n^2/\sigma^2 - 1)^2.$$

Next lemma shows that the $F_n, n \geq 1,$ is slowly changing. This fact enables us to use the nonlinear renewal theory to analyse our stopping time t .

Before we state the Lemma we need to introduce some notations. Let

$$(2.3.4) \quad L_{1,n} = nC_n/2\sigma^2, \quad L_{2,n} = (3/8)\lambda_n^{-5/2} n(\hat{\sigma}_n^2/\sigma^2 - 1)^2.$$

Note that $F_n = L_{1,n} + L_{2,n}.$

LEMMA 2.6. Assume the model in (2.1.1). Then $\{L_{1,n}, n \geq 1\}$ and $\{L_{2,n}, n \geq 1\}$ are slowly changing. Consequently, $\{F_n, n \geq 1\}$ is also slowly changing.

Proof. From Lai and Siegmund (1982) (see Theorem 4.2 and its proof) $C_n \rightarrow 0$ and $\hat{\sigma}_n^2 \rightarrow \sigma^2$ a.s., as $n \rightarrow \infty$. Therefore,

$$n^{-1} L_{1,n} \rightarrow 0 \quad \text{and} \quad n^{-1} L_{2,n} \rightarrow 0 \quad \text{a.s., as } n \rightarrow \infty,$$

since $\lambda_n \rightarrow 1$ a.s.

By Corollary 2.1 and Lemma 1.4W

(2.3.5) $\{L_{1,n}, n \geq 1\}$ is stochastically bounded and
u.c.i.p..

From the representation (2.2.5) for $\hat{\sigma}_n^2$, example 1.8 of Woodroffe (1982), (2.3.5), $\lambda_n \rightarrow 1$ a.s. and Lemma 1.4W it follows that

(2.3.6) $\{L_{2,n}, n \geq 1\}$ is stochastically bounded and
u.c.i.p..

The required result now follows immediately. $\square$

REMARK. Consider the linear model,

$$Y_i = \beta X_i + \varepsilon_i, \quad i \geq 1,$$

where $\{X_i, i \geq 1\}$ are i.i.d., $\{\varepsilon_i, i \geq 1\}$ are i.i.d. $N(0, \sigma^2)$, $\{X_i, i \geq 1\}$ is independent of $\{\varepsilon_i, i \geq 1\}$. Using the least squares estimator of β , the normality of $\{\varepsilon_i\}$ and the given independence, one can write $\sum_{i=1}^n (Y_i - \hat{\beta} X_i)^2$ as a sum of $(n-1)$ i.i.d Chi-square r.v.'s. Using this representation, Finster (1983) investigated his sequential point estimator of β by expressing the defining variable of his stopping time as a sum of a random walk and a slowly changing sequence of r.v.'s.

In our problem, we neither need to assume any form of the underlying distribution nor do we have the independence between $\{X_{i-1}, i > 1\}$ and $\{\varepsilon_i, i \geq 1\}$. Nevertheless, by exploiting the model (2.1.1) and the representation (2.2.5)

for $\hat{\sigma}_n^2$ we are able to express the defining variable Z_n in (2.3.1) as a sum of a random walk and a slowly changing sequence of r.v.'s as evidenced by Lemma 2.5, Corollary 2.1 and Lemma 2.6. This indicates that the normality assumption in Finster (1983) was unnecessary to obtain a second order expansion for the expected stopping time. $\square$

In what follows, references to Lemmas, Theorems and assumptions in Woodroffe (1982) will be followed by 'W'.

Before we state the main theorem we need to introduce some more notations. Let

$$\tau = \inf\{n \geq n_A: S_n \geq A^{1/2}\sigma\} \quad \text{and} \quad (2.3.7)$$

$$\rho = E S_{\tau_0}^2 / (2 E S_{\tau_0}), \quad \text{where } \tau_0 \text{ is same as } \tau \text{ with } 'A^{1/2}\sigma' \text{ replaced by } '0'.$$

Note 3. The finiteness of ρ follows from Theorems 2.3W, 2.4W and Corollary 2.2W, if $E\epsilon_1^4 < \infty$.

Theorem 2.2. (Second order expansion) Assume the error d.f. F is non arithmetic. Under the conditions of Theorem 2.1.

$$(2.3.8) \quad E t_A - A^{1/2}\sigma = \rho - (1/2) - (3/8)E(\epsilon_1^2/\sigma^2 - 1)^2 + o(1),$$

as $A \rightarrow \infty$.

The proof of this theorem will follow from a series of lemmas.

The following lemma is a slight extension of Theorem 2 of Hagwood and Woodroffe (1982), which appears as Theorem 4.5 in Woodroffe (1982). For convenience, we will refer to Theorem 4.5 of Woodroffe (1982) and assume all of the relevant material from Chapter 4 of Woodroffe (1982). In order not to confuse with our notations we will add a ' $\sim$ ' to all the notations used in Woodroffe (1982).

LEMMA 2.7. Assume the conditions (4.10)W, (4.12)W and (4.16)W of Theorem 4.5W. Assume, instead of (4.11)W, that

(2.3.9) $\xi'_n = \tilde{\ell}_n + L'_{1,n} + L'_{2,n}$ on $\tilde{\Lambda}_n$, $n \geq 1$,
where the sets $\tilde{\Lambda}_n$, $n \geq 1$ and constants $\tilde{\ell}_n$, $n \geq 1$, are as in Theorem 4.5W, and $L'_{1,n}$ and $L'_{2,n}$ for $n \geq 1$ are such that for each $i = 1, 2$

$$(2.3.10) \quad \{ \max_{0 \leq k \leq n} |L'_{i,n}|; n \geq 1 \} \text{ is U.I.}$$

$$(2.3.11) \quad \sum_{n=1}^{\infty} P(L'_{i,n} < -n\epsilon) < \infty \text{ for some } \epsilon, 0 < \epsilon < \tilde{\mu}$$

$$(2.3.12) \quad L'_{i,n} \rightarrow_d L'_i \text{ as } n \rightarrow \infty.$$

If $L'_{1,n}$ and $L'_{2,n}$, $n \geq 1$ are slowly changing, $\tilde{F}$ the d.f. of the summands in $\tilde{S}_n$ is non-arithmetic, and $\tilde{N}_a = [a/\tilde{\mu}]$, then as $a \rightarrow \infty$

$$(2.3.13) \quad E \tilde{t}_a = (1/\mu)(a + \tilde{\rho} - \tilde{\ell}_{N_a} - EL'_1 - EL'_2) + O(1).$$

Proof. A careful analysis of the proof of Theorem 4.5W shows that we only need to prove analogues of Lemmas 4.3W and 4.4W in order to conclude (2.3.13). That is, we need to show that

$$\int_{\tilde{B}_a} \tilde{f}_{t_a}' dP = \tilde{\ell}_{N_a} + E(L_1') + E(L_2') + o(1), \quad \text{and}$$

(2.3.14)

$$\int_{\tilde{B}_a} \tilde{R}_a dP = \tilde{\rho} + o(1), \quad \text{as } a \rightarrow \infty,$$

where $\tilde{B}_a = \{\tilde{N}' \leq \tilde{t}_a \leq \tilde{N}'', \tilde{\tau}_b > \tilde{N}'\} \cap \bigcap_{k=\tilde{N}'}^{\infty} \tilde{A}_k$ with $\tilde{N}' = [\epsilon_1 \tilde{N}_a]$, $\tilde{N}'' = [\epsilon_2 \tilde{N}_a] + 1$ and $\tilde{b} = \tilde{a} - \tilde{\ell}_{N_a}$.

As for (2.3.14), write

$$(\tilde{f}_{t_a}' - \tilde{\ell}_{N_a}) 1_{\tilde{B}_a} = (\tilde{f}_{t_a}' - \tilde{f}_{N_a}') 1_{\tilde{B}_a} + L_{1, \tilde{N}_a}' 1_{\tilde{B}_a} + L_{2, \tilde{N}_a}' 1_{\tilde{B}_a}.$$

Since $\tilde{f}_n'$, $n \geq 1$ is slowly changing, the Anscombe's theorem and $\tilde{t}_a/\tilde{N}_a \rightarrow 1$ a.s. implies $(\tilde{f}_{t_a}' - \tilde{f}_{N_a}') 1_{\tilde{B}_a} \rightarrow 0$ in

Probability, as $a \rightarrow \infty$. Also, $(\tilde{f}_{t_a}' - \tilde{\ell}_{N_a}) 1_{\tilde{B}_a}$ is dominated

by $M'_a = \max_{\tilde{N}' \leq k \leq \tilde{N}''} |\tilde{f}_k' - \tilde{\ell}_{N_a}|$, which is U.I. by condition

(2.3.10). Moreover, (2.3.10) implies $\{L_{i, \tilde{N}_a}' 1_{\tilde{B}_a}; a \geq 0\}$ is

U.I. for each $i = 1, 2$. Therefore, $\{(\tilde{f}_{t_a}' - \tilde{f}_{N_a}') 1_{\tilde{B}_a}; a \geq 0\}$

is U.I. Consequently, since $P(\tilde{B}_a) = o(1/\tilde{N}_a)$ we have the first assertion (2.3.14), by (2.3.12). The other assertion in (2.3.14) uses the uniform integrability of M'_a and follows by arguments similar to the proof of Lemma 4.4W. $\square$

Note 4. In the application of Lemma 2.7 we will take the set $\tilde{A}_k = \Omega$, the sample space, for all $k \geq 1$, $\tilde{\mu} = 1$, $a = A^{1/2}\sigma$, $\tilde{N}' = N'$, $\tilde{N}'' = N''$, $\tilde{\ell}_{N_a} = 0$, $L_{1, n}' = L_{1, n}$.

$$L'_{2,n} = L_{2,n}, \quad \tilde{S}_n = S_n, \quad \xi'_n = \xi_n, \quad \tilde{t}_a = t, \quad \tilde{R}_a = U, \quad \tilde{\rho} = \rho, \\ L'_1 = L \quad \text{and} \quad L'_2 = L_2.$$

REMARK. Had we assumed that $(L_{1,n}, L_{2,n}) \rightarrow_d (L_1, L_2)$, Lemma 2.7 would have directly followed from Theorem 4.5W. Thus the main observation here is that one needs only the weak convergence of the marginals, and not those of the joint distributions, besides of course the other assumptions, in order for (2.3.13) to be valid.

We use this Lemma to prove Theorem 2.2 below where we readily have the weak convergence of the marginals and where the joint convergence is not at all apparent. Finally, this Lemma could be further extended to the case where $\xi'_n = \tilde{\ell}_n + L'_{1n} + L'_{2n} + \dots + L'_{kn}$ on $\tilde{\Lambda}_n$, $n \geq 1$, for $k > 1$ and finite, and where each L'_{in} satisfies conditions similar to (2.3.10) - (2.3.12) of Lemma 2.7. $\square$

Note that the proof of Lemma 2.7 assumes the results such as Lemma 4.5W, and weak and the moment convergences of $\tilde{R}_a$. Therefore, in order to prove Theorem 2.2 we need to establish these in the present context.

The next lemma is similar to Lemma 4.5W, and therefore only a sketch of its proof will be given. In what follows all unidentified limits are taken as $A \rightarrow \infty$.

LEMMA 2.8. Under the conditions of Theorem 2.2,

$$\int_F S_t dP = Et + o(1),$$

where F is as in (2.2.27).

Proof. Since $\xi_n \geq 0$ we have that $t \leq \tau$ (see (2.3.7)). Therefore, by Note 4 the set F is same as the set $\tilde{B}_a$ defined in Woodroffe (1982) (see also (2.3.14)).

Also, by Lemma 2.4

$$(2.3.15) \quad P(\bar{F}) = o(A^{-1/2}).$$

Use this, (2.2.2) and arguments similar to the proof of Lemma 4.5W to get the required result. $\square$

Before we state the next lemma, let

$$(2.3.16) \quad W_n = n\hat{\sigma}_n^2$$

and note that the relation

$$\hat{\beta}_n = \hat{\beta}_{n-1} + (X_n - \hat{\beta}_{n-1}X_{n-1})^2 X_{n-1}^2 / I_n$$

yields $W_n = W_{n-1} + (1 - X_{n-1}^2 / I_n)(X_n - \hat{\beta}_{n-1}X_{n-1})^2$. Therefore

$$(2.3.17) \quad W_n \leq W_{n+1} \text{ for all } n \geq 1.$$

The next lemma essentially verifies (2.3.10) of Lemma 2.7.

LEMMA 2.9. Assume that $E|\xi_1|^8 < \infty$, $E|X_0|^8 < \infty$ and $E|1/(\xi_1^2 + \dots + \xi_M^2)|^4 < \infty$ for some positive integer M. Then

$$(2.3.18) \quad \{ \max_{N' \leq n \leq N''} L_{1,n}^1 F; A \geq 1 \} \text{ is U.I.}$$

and

$$(2.3.19) \quad \{ \max_{N' \leq n \leq N''} L_{2,n}^1 F; A \geq 1 \} \text{ is U.I.}$$

where F is as in (2.2.27). Consequently

$$\left\{ \max_{N' \leq n \leq N''} F_n 1_F; A \geq 1 \right\} \text{ is U.I.}$$

Proof. Assume w.l.o.g. $\sigma^2 = 1$. Argue as for (2.2.33) to get (2.3.18).

Now for (2.3.19),

$$\begin{aligned} \max_{N' \leq n \leq N''} L_{2,n} 1_F &\leq (3/8) \max_{N' \leq n \leq N''} \lambda_n^{-5/2} 1_F \\ &\quad \cdot \max_{N' \leq n \leq N''} n(\hat{\sigma}_n^2 / \sigma^2 - 1)^2 1_F. \end{aligned}$$

where, by the definition of λ_n (see (2.3.2)) and (2.3.17)

$$\begin{aligned} \max_{N' \leq n \leq N''} \lambda_n^{-5/2} 1_F &\leq 1 + \max_{N' \leq n \leq N''} (n/W_n)^{5/2} 1_F \\ &\leq 1 + 2^{5/2} \max_{N' \leq n \leq N''} (n-1/W_{n-1})^{5/2} 1_F \\ &\leq 1 + O(1) [(N'-1)/W_{N'-1}]^{5/2} 1_F. \end{aligned}$$

By the definition of t

$$F \subset \{ (N'-1) < A^{1/2} [W_{N'-1} / (N'-1)]^{1/2} \} = G.$$

Therefore,

$$[(N'-1)/W_{N'-1}]^{5/2} 1_F \leq [A^{1/2} / (N'-1)]^5 1_G$$

$$(2.3.20) \quad = O(1).$$

Therefore, in order to show (2.3.19), it suffices to show

$$(2.3.21) \quad \left\{ \max_{N' \leq n \leq N''} n(\hat{\sigma}_n^2/\sigma^2 - 1)^2 1_F; A \geq 1 \right\} \text{ is U.I.}$$

By example 4.3 of Woodroffe (1982) applied to $\{\varepsilon_1^2/\sigma^2\}$ we have

$$\left\{ \max_{N' \leq n \leq N''} n^{-1} \left[\sum_{i=1}^n (\varepsilon_i^2/\sigma^2 - 1) \right]^2; A \geq 1 \right\} \text{ is U.I.}$$

Moreover, by the Schwarz inequality

$$E \max_{N' \leq n \leq N''} n C_n^2 \leq (N')^{-3} \left\| \sup_{n \geq M} (n/I_n) \right\|_4^2 \left\| \max_{N' \leq n \leq N''} D_n \right\|_8^4$$

$$= O(A^{-3/2}) O(N'')^2$$

$$\rightarrow 0.$$

by the Doob's maximal inequality (see Inequality C, Sec. 1.2) and Lemma 2.1. Use the above arguments and (2.2.5) to get (2.3.21). $\square$

In order to state the next two lemmas, denote the "overshoot" of the stopping time t (see (2.3.1)) by

$$(2.3.22) \quad U = Z_t - A^{1/2} \sigma.$$

The next two lemmas give the asymptotic distribution of U and the moment convergence of $U 1_F$.

LEMMA 2.10. Suppose that the error distribution F is non-arithmetic. Then U has a limiting distribution H , where

$$H(dr) = \frac{1}{ES_{\tau_0}} P(S_{\tau_0} \geq r) dr, \quad r > 0.$$

where τ_0 is as in (2.3.7).

Proof. Follows from Theorem 4.1W. $\square$

LEMMA 2.11. Under the assumptions of Lemmas 2.9 and 2.10,

$$\int_F U \, dP = \rho + o(1),$$

where ρ is as in (2.3.7).

Proof. The proof uses Lemma 2.9 and 2.10 and follows from similar arguments given in the proof of Lemma 4.4W. $\square$

PROOF OF THEOREM 2.2. Assume w.l.o.g. $\sigma^2 = 1$. The proof will follow from Lemma 2.7, once we verify all the regularity conditions in it.

By (2.1.4)

$$(2.3.23) \quad L_{1,n} \rightarrow_d (1/2)\mathfrak{L}_1^2 = L_1.$$

Therefore, by CLT, the Slutsky Theorem and $\lambda_n \rightarrow 1$ a.s. we have

$$(2.3.24) \quad L_{2,n} \rightarrow_d (3/8)\mathfrak{L}_1^2 E(\mathfrak{L}_1^2 - 1)^2 = L_2.$$

The verification of rest of the conditions of Lemma 2.7 follows from Note 4, (2.3.15), Lemma 2.9 and Lemma 2.6.

Hence the theorem. $\square$

The next theorem gives asymptotic normality and uniform integrability of standardized stopping time t .

Before we state the next theorem let

$$(2.3.25) \quad a = \Lambda^{1/2} \sigma \quad \text{and} \quad t^* = (t-a)/\sqrt{a}.$$

This notation will be repeatedly used in Section 2.4.

THEOREM 2.3. (Asymptotic normality and Uniform integrability)

If $E|\xi_1|^4 < \infty$, then

$$(2.3.26) \quad t^* \rightarrow_d N(0, \alpha^2),$$

where $\alpha^2 = (1/4) V(\xi_1^2/\sigma^2)$.

Assume the moment conditions of Theorem 2.1. If n_A in

(2.1.9) is such that $\Lambda^{1/2(1+\eta)} \leq n_A = o(\Lambda^{1/2})$, with

$\eta \in (0, (s-2)/s)$. Then

$$(2.3.27) \quad \{|t^*|^s; \Lambda \geq 1\} \text{ is U.I.}$$

The proof of Theorem 2.3 depends on a technical lemma.

LEMMA 2.12. Let $b > 0$, define for $\gamma > 0$, $x \geq 0$,

$g_\gamma(x) = x^3/b^2 - x - \gamma(x-b)$. Then

(i) $g_2(x) \geq 0$ for all $x \geq 0$.

(ii) If $\gamma \in (0, 2)$ then there exists $0 < \epsilon < 1$ such that for all $x \in [(1-\epsilon)b, b]$, $g_\gamma(x) \leq 0$.

Proof. Observe that $g_\gamma(b) = 0$ for every $\gamma > 0$.

Moreover, $g'_\gamma(x) = 3x^2/b^2 - 1 - \gamma$. From this it readily

follows that g_γ is strictly convex on $(0, \infty)$ and it has a unique minimum at $x = b\sqrt{(1+\gamma)/3}$.

Assertion (i) follows since $g_2(x) \geq g_2(b) = 0$ for all $x \geq 0$. For (ii), note that $g_\gamma(b\sqrt{(1+\gamma)/3}) < g_\gamma(b) = 0$. Now let $\epsilon = 1 - \sqrt{(1+\gamma)/3}$ to get the required result. $\square$

PROOF OF THEOREM 2.3. Assume $\sigma^2 = 1$ w.l.o.g.. The assertion (2.3.26) follows from Lemma 4.2W, since $\{F_n, n \geq 1\}$ is slowly changing and $F_n/\sqrt{n} \rightarrow 0$ in probability (see (2.3.5) and (2.3.6)).

As for (2.3.27), it suffices to show

$$(2.3.28) \quad \{ [(t^*)^+]^s; A \geq 1 \} \text{ is U.I.}$$

and

$$(2.3.29) \quad \{ [(t^*)^-]^s; A \geq 1 \} \text{ is U.I.}$$

Consider (2.3.28). By Lemma 2.12 (i) applied with $x = t - 1$ and $b = a$ we have

$$2(t - 1 - a) \leq (t - 1)^3/a^2 - (t - 1)$$

$$\leq \sum_{i=1}^{t-1} (\xi_i^2 - 1).$$

by (2.1.9), (2.2.5) and since $C_n \geq 0$. By the c_s -inequality

$$[(t^*)^+]^s \leq c_s 2^{-s} \left[\left(\sum_{i=1}^{t-1} (\xi_i^2 - 1)/\sqrt{a} \right)^+ \right]^s + c_s (1/\sqrt{a})^s.$$

Now, by Lemma 2.3 and Lemma 5 of Chow and Yu (1981)

$$(2.3.30) \quad \{ \left| \sum_{i=1}^t (\xi_i^2 - 1)/\sqrt{a} \right|^s; A \geq 1 \} \text{ is U.I.}$$

Moreover, $E|g_t^2 - 1|^s = o(Et)$ (see Chow and Teicher (1978), p. 148, Lemma 2). Therefore by (2.2.2)

$$(2.3.31) \quad E|g_t^2 - 1|^s / (\sqrt{a})^s \rightarrow 0.$$

Hence, (2.3.28) follows from (2.3.30) and (2.3.31).

It only remains to show (2.3.29). Now,

$$E[(t^*)^-]^s 1_{[t < (1-\epsilon)a]} = \int_0^\infty s x^{s-1} P\{t^* \leq -x, t < (1-\epsilon)a\} dx.$$

Observe that

$$P(t^* \leq -x, t < (1-\epsilon)a) \begin{cases} = P(t < (1-\epsilon)a) & \text{if } x \leq \epsilon\sqrt{a} \\ \leq P(t < (1-\epsilon)a) & \text{if } \epsilon\sqrt{a} < x \leq \sqrt{a} \\ = 0 & \text{if } x > \sqrt{a} \end{cases}$$

Now use (2.2.21) to conclude

$$(2.3.32) \quad \int_0^{\sqrt{a}} s x^{s-1} P\{t \leq (1-\epsilon)a\} dx = a^{s/2} \cdot O(A^{-(s-1)/2(1+\eta)}) \rightarrow 0 \quad (\text{since } \eta < [(s-2)/s]).$$

From these it follows that $[(t^*)^-]^s 1_{[t < (1-\epsilon)a]}$ is U.I.

As for the Uniform integrability of $[(t^*)^-]^s$ on the set $[t \geq (1-\epsilon)a]$, it suffices to show that $[(t^*)^-]^s$ on the set $[(1-\epsilon)a \leq t \leq a]$ is U.I., because $(t^*)^- = 0$ for $t > a$.

Let $\gamma \in (0, 2)$. Then by Lemma 2.12 (ii)

$$\gamma(a - t)1_{F'} \leq (t - t^{3/a^2})1_{F'},$$

where $F' = \{(1-\epsilon)a \leq t \leq a\}$. Therefore,

$$(2.3.33) \quad (t - a)^- 1_{F'} \leq \gamma^{-1} (t - t^{3/a^2}) 1_{F'} \leq \gamma^{-1} [tC_t - \sum_{i=1}^t (g_i^2 - 1)] 1_{F'}.$$

By the Schwarz inequality

$$E(tC_t/\sqrt{a})^s 1_F \leq A^{-s/4} (N')^{-s} \left\| \sup_{n \geq M} (n/I_n) \right\|_{2s}^s \left\| \max_{n \leq N''_A} D_n \right\|_{4s}^{2s}$$

$$(2.3.34) \quad \rightarrow 0,$$

by the Doob's maximal inequality (see Inequality C, Section 1.2) and Lemma 2.1. The required result now follows from (2.3.33), (2.3.34) and (2.3.30). Hence the theorem. $\square$

2.4. Regret Analysis.

The main purpose of this section is to analyse the regret $R_t - R_{n_0}$ of the sequential procedure $\hat{\beta}_t$. To that effect, use (2.3.25) to write

$$\begin{aligned} R_t - R_{n_0} &= A E C_t + E t - 2a \\ &= A E (C_t - \sigma^2 t^{-1}) + E(t-a)^2 t^{-1} \\ (2.4.1) \quad &= A E t^{-1} I_t^{-1} (D_t^2 - \sigma^2 I_t) + E a t^{-1} (t^*)^2. \end{aligned}$$

As for the second term above, it follows from (2.3.26), (2.2.1) and the Slutsky theorem that

$$a t^{-1} (t^*)^2 \rightarrow_d a^2 g_1^2, \quad \text{as } A \rightarrow \infty.$$

Moreover, by the Holder inequality, (2.1.9) (2.3.27) and (2.2.21)

$$\begin{aligned} a E t^{-1} (t^*)^2 1_{[t < N']} &\leq a n_A \|t^*\|_{2p}^2 \{P(t < N')\}^{1-1/p} \\ &= O(A^{1/2}) A^{-1/2(1+\eta)} O(A^{-(p-1)^2/2p(1+\eta)}) \\ &= o(1), \quad \text{as } A \rightarrow \infty. \end{aligned}$$

Also, by (2.3.27)

$$\{a t^{-1} (t^*)^2 1_{[t > N']}, A \geq 1\} \text{ is U.I..}$$

Therefore,

$$E a t^{-1} (t^*)^2 = \alpha^2 + o(1), \text{ as } A \rightarrow \infty.$$

Of course, all this is valid under the conditions of Theorem 2.1 with the proviso that $\eta < [(p-1)^2/p]$.

As for the first term in (2.4.1), use the Taylor's theorem to write

$$\begin{aligned} A E t^{-1} I_t^{-1} (D_t^2 - \sigma^2 I_t) &= \theta^{-1} E(D_t^2 - \sigma^2 I_t) \\ &\quad - \theta^{-2} E(D_t^2 - \sigma^2 I_t)(A^{-1} t I_t - \theta) \\ &\quad + E(D_t^2 - \sigma^2 I_t) \lambda_t^{-3} (A^{-1} t I_t - \theta)^2 \\ (2.4.2) \quad &= (a) + (b) + (c), \text{ say.} \end{aligned}$$

where λ_t is a r.v. lying between θ and $A^{-1} t I_t$ with $\theta = \sigma^4/(1-\beta^2)$.

To analyse (a), (b) and (c) further, we will prove a series of lemmas. These lemmas are of interest in their own way. Moreover, these will yield, under enough moment conditions, that

$$\begin{aligned} (a) &= 0, \quad (c) = O(1) \text{ and} \\ (b) &= -A^{1/2} E(D_t^2 - \sigma^2 I_t) X_t^2 + O(1), \text{ as } A \rightarrow \infty. \end{aligned}$$

In what follows assume further that the r.v.'s X_0 and X_1 have the same distribution. Note that this assumption implies $\{X_i, i \geq 0\}$ is stationary.

LEMMA 2.13. Assume the moment conditions of Lemma 2.2.

If n_A in (2.1.9) is such that $A^{1/2(1+\eta)} \leq n_A$ with $\eta \in (0, (s-1)/s)$ and $s > 1$, then

$$\{(t/a)^{-s}; A \geq 1\} \text{ is U.I.}$$

Proof. Observe that $(a/t)^s 1_{[t \geq N']} = O(1)$.

Moreover,

$$\begin{aligned} E(a/t)^s 1_{[t < N']} &\leq a^s n_A^{-s} P(t < N') \\ &\leq O(A^{s/2}) A^{-s/2(1+\eta)} O(A^{-(s-1)/2(1+\eta)}) \\ &\rightarrow 0 \quad \text{as } A \rightarrow \infty, \end{aligned}$$

by (2.2.21) and since $\eta < (s-1)/s$. Hence the lemma. $\square$

In fact the extra assumption of stationarity will be needed only for the next lemma and Lemma 2.16 below.

LEMMA 2.14. Assume the model in (2.1.1). If $r \geq 1$ and $E|\varepsilon_1|^r < \infty$ and $E|X_0|^r < \infty$, then for every $\beta > 0$

$$P\left(\sup_{n \geq 1} n^{-1} \sum_{i=1}^n |X_{i-1}|^r > \beta\right) \leq \beta^{-1} \int_{\left\{\sup_{n \geq 1} n^{-1} \sum_{i=1}^n |X_{i-1}|^r > \beta\right\}} |X_0|^r dP.$$

Moreover if $p > 1$ and $E|\varepsilon_1|^{rp} < \infty$ and $E|X_0|^{rp} < \infty$, then

$$(2.4.3) \quad E \sup_{n \geq 1} \left(\frac{1}{n} \sum_{i=1}^n |X_{i-1}|^r \right)^p \leq \left(\frac{p}{p-1} \right)^p E|X_0|^{rp}.$$

Proof. Since $\{|X_i|^r, i \geq 0\}$ is also stationary, the assertions follow from the Doob's maximal inequality for stationary averages (see Doob (1953), Chapter X, Lemma 2.2, p. 466) and (second) Lemma 3.3.1 of Stout (1974). $\square$

The following lemma is a sequential analogue of Lemma 2.1. Henceforth all unidentified limits are taken as $A \rightarrow \infty$.

LEMMA 2.15. If $p \geq 1$ and $E|X_0|^{4p} < \infty$ and
 $E|\varepsilon_1|^{4p} < \infty$, then

$$\| a^{-1/2} D_t \|_p = O(1),$$

where D_n is as in (2.2.4).

Proof. Assume w.l.o.g. $\sigma^2 = 1$. Let $B = \{t \leq N''\}$.
 By the Doob's maximal inequality for the martingale $\{D_n\}$
 (see Inequality C, Section 1.2) and Lemma 2.1 we get

$$\begin{aligned} a^{-1/2} \| D_t 1_B \|_p &\leq a^{-1/2} \| \max_{n \leq N''} D_n \|_p \\ &= \left(\frac{p}{p-1}\right) a^{-1/2} O[(N'')^{1/2}] \\ (2.4.4) \qquad \qquad &= O(1). \end{aligned}$$

By the Schwarz inequality

$$(2.4.5) \quad a^{-1/2} \| D_t 1_{\bar{B}} \|_p \leq \| a^{-1/2} t 1_{\bar{B}} \|_{2p} \| (D_t/t) 1_{\bar{B}} \|_{2p}.$$

Now, by the Minkowski inequality

$$\begin{aligned} \| a^{-1/2} t 1_{\bar{B}} \|_{2p} &\leq \| (t^*)^+ 1_{\bar{B}} \|_{2p} + \sqrt{a} \{P(\bar{B})\}^{1/2p} \\ &\leq (1 + \epsilon^{-1}) \| (t^*)^+ 1_{\bar{B}} \|_{2p} \quad (\text{since on } \bar{B}, t - a > \epsilon a) \\ (2.4.6) \qquad \qquad &\rightarrow 0, \end{aligned}$$

by (2.3.28) and because $P(\bar{B}) \rightarrow 0$ (by (2.2.22)).

By the Schwarz inequality and the Minkowski inequality

$$\begin{aligned} \| (D_t/t) 1_{\bar{B}} \|_{2p} &\leq \| \sup_{n \geq N''} |D_n/n| \|_{2p} \\ &\leq \| \sup_{n \geq N''} \bar{I}_n^{1/2} \cdot \sup_{n \geq N''} (n^{-1} \sum_{i=1}^n \varepsilon_i^2)^{1/2} \|_{2p} \\ &\leq 2^{-1} \| \sup_{n \geq N''} \bar{I}_n \|_{2p} + \sup_{n \geq N''} (n^{-1} \sum_{i=1}^n \varepsilon_i^2) \|_{2p} \\ &\leq 2^{-1} \| \sup_{n \geq N''} \bar{I}_n \|_{2p} + 2^{-1} \| \sup_{n \geq N''} (n^{-1} \sum_{i=1}^n \varepsilon_i^2) \|_{2p} \\ &\leq 2^{-1} \left(\frac{2p}{2p-1}\right) [\|X_0\|_{4p}^2 + \|\varepsilon_1\|_{4p}^2] \\ (2.4.7) \qquad \qquad &< \infty, \end{aligned}$$

by Lemma 2.14 and the maximal inequality for the reverse martingale $\{n^{-1} \sum_{i=1}^n \varepsilon_i^2, n \geq 1\}$ (see Inequality E, Section 1.2). The required result follows from (2.4.4), (2.4.5), (2.4.6) and (2.4.7). $\square$

LEMMA 2.16. If $p \geq 2$, and $E|X_0|^{4p} < \infty$, $E|\varepsilon_1|^{4p} < \infty$, then,

$$a^{-1/2} \|(1-\beta^2) I_t - t\sigma^2\|_p = o(1),$$

where I_n is as in (2.2.4).

Proof. Assume w.l.o.g. $\sigma^2 = 1$. Squaring the model in (2.1.1) and summing over i yields

$$(2.4.8) \quad (1-\beta^2)I_n = X_0^2 - X_n^2 + 2\beta D_n + \sum_{i=1}^n \varepsilon_i^2.$$

Now, clearly

$$(2.4.9) \quad a^{-1/2} \|X_0^2\|_p \rightarrow 0.$$

Therefore, in view of (2.3.30) and Lemma 2.15 it only remains to analyse $a^{-1/2} \|X_t^2\|_p$. For this, first observe that

$$\begin{aligned} a^{-1} |X_t|^{2p} &= a^{-1} \left[\sum_{i=1}^{t+1} |X_{i-1}|^{2p} - \sum_{i=1}^t |X_{i-1}|^{2p} \right] \\ &\rightarrow 0 \quad \text{a.s.}, \end{aligned}$$

by (2.2.1) and $t^{-1} \sum_{i=1}^t X_{i-1}^{2p} \rightarrow E(X_0^{2p} | \mathcal{I})$ a.s., where $\mathcal{I}$ is a σ -field of invariant events (see Breiman (1968), Theorem 6.28 for non-sequential a.s. convergence). Moreover,

$$a^{-1} |X_t|^{2p} \leq a^{-1} \sum_{i=1}^{t+1} |X_{i-1}|^{2p}$$

$$\leq ((t+1)/a) \sup_{n \geq 1} n^{-1} \sum_{i=1}^n |X_{i-1}|^{2p}$$

$$\leq 2^{-1}((t+1)/a)^2 + 2^{-1} \sup_{n \geq 1} \left[n^{-1} \sum_{i=1}^n |X_{i-1}|^{2p} \right]^2$$

and therefore U.I. by Lemmas 2.3 and 2.14. From these we have

$$(2.4.10) \quad E |X_t|^{2p} = o(a).$$

Thus, for $p \geq 2$

$$(2.4.11) \quad a^{-1/2} \|X_t\|_p^2 = o(1).$$

The lemma now follows from (2.4.8) and the Minkowski inequality. $\square$

The next lemma is similar to Lemma in Chow and Martinsek (1982).

LEMMA 2.17. Assume the model (2.1.1). Let t be any $\mathcal{F}_n$ -adapted stopping time such that $E t < \infty$ and $E |D_t| < \infty$.

Then

- (i) $E D_t = 0.$
- (ii) $E D_t^2 = \sigma^2 E I_t$
- (iii) $\{D_n^2 - \sigma^2 I_n, \mathcal{F}_n, n \geq 1\}$ is a mean zero martingale.

If, in addition, $E(\varepsilon_1^4) < \infty$ then

$$(iv) \quad E(D_t^2 - \sigma^2 I_t)^2 = \sigma_1^2 E \sum_{j=1}^t X_{j-1}^4 + 4\sigma^2 E \sum_{j=1}^t X_{j-1}^2 D_{j-1}^2 \\ + 4\mu_3 E \sum_{j=1}^t X_{j-1}^3 D_{j-1},$$

where $\sigma_1^2 = E(\xi_1^2 - \sigma^2)^2$ and $\mu_3 = E(\xi_1^3)$.

$$(v) \quad E(D_t^2 - \sigma^2 I_t) \left(\sum_{i=1}^t \xi_i^2 - t\sigma^2 \right) = 2\mu_3 E \sum_{i=1}^t X_{i-1} D_{i-1} \\ + \sigma_1^2 E I_t.$$

$$(vi) \quad E(D_t^2 - \sigma^2 I_t) D_t = \mu_3 E \sum_{i=1}^t X_{i-1}^3 + 2\sigma^2 E \sum_{i=1}^t X_{i-1}^2 D_{i-1}.$$

Proof. Recall that $\{D_n, \mathcal{F}_n\}$ is a mean zero martingale. Moreover, by the Schwarz inequality and the Markov inequality

$$E |D_n| 1_{\{t > n\}} \leq \|D_n\|_2 P^{1/2}(t > n) \\ \leq \|D_n\|_2 n^{-1/2} E^{1/2} t 1_{\{t > n\}}$$

$$(2.4.12) \quad \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

by Lemma 2.1, and the Dominated Convergence Theorem.

Therefore, by the optional stopping theorem

$$(2.4.13) \quad E\{D_t | \mathcal{F}_n\} = D_n \quad \text{if } t \geq n$$

and hence $E D_t = 0$.

As for (ii), it follows from (2.4.12), Theorem 1 and Lemma 6 of CRT (1965) and the independence of X_{i-1} and ξ_i

that
$$ED_t^2 = E \sum_{i=1}^t X_{i-1}^2 \varepsilon_i^2 = E \sum_{i=1}^t E(X_{i-1}^2 \varepsilon_i^2 | \mathcal{F}_{i-1})$$

$$= \sigma^2 E I_t.$$

As for (iii), let $M_n = D_n^2 - \sigma^2 I_n$. From (2.1.1) we have that

(2.4.14) $\{M_n, \mathcal{F}_n; n \geq 1\}$ is a mean zero martingale, with $X_{n-1}^2 \varepsilon_n^2 + 2X_{n-1} \varepsilon_n D_{n-1} - \sigma^2 X_{n-1}^2$ as the martingale difference sequence.

As for (iv), it follows from (2.4.8), (i), and the Wald's lemma that $(1-\beta^2)EI_t \leq EX_0^2 + \sigma^2 Et < \infty$. Therefore from (ii)

(2.4.15) $E D_t^2 < \infty$.

Furthermore,

$$\begin{aligned} E|M_n|1_{\{t > n\}} &\leq ED_n^2 1_{\{t > n\}} + \sigma^2 E I_n 1_{\{t > n\}} \\ &= E D_{t \wedge n}^2 - ED_t^2 1_{\{t \leq n\}} + \sigma^2 E I_n 1_{\{t > n\}} \\ (2.4.16) \quad &\rightarrow 0 \text{ as } n \rightarrow \infty, \end{aligned}$$

because

$$\begin{aligned} E D_{t \wedge n}^2 &= \sigma^2 E I_{t \wedge n} \rightarrow \sigma^2 E I_t, \quad E D_t^2 1_{\{t \leq n\}} \rightarrow E D_t^2 \\ \text{and } E I_n 1_{\{t > n\}} &\leq E I_t 1_{\{t > n\}} \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

Assertion (iv) now follows from (iii) and Theorem 1 and Lemma 6 of CRT (1965).

As for (v), use $-2xy = (x-y)^2 - x^2 - y^2$ to write

$$(2.4.17) \quad -M_t \left(\sum_{i=1}^t \varepsilon_i^2 - t\sigma^2 \right) = \frac{1}{2} \left[\left(M_t - \sum_{i=1}^t \varepsilon_i^2 + t\sigma^2 \right)^2 - M_t^2 - \left(\sum_{i=1}^t \varepsilon_i^2 - t\sigma^2 \right)^2 \right].$$

By the Wald's lemma

$$(2.4.18) \quad E \left[\sum_{i=1}^t (\xi_i^2 - \sigma^2) \right]^2 = E(\xi_1^2 - \sigma^2)^2 E t.$$

Moreover, using arguments in the proof of the Wald's lemma (see Woodroffe (1982), p. 9) it can be shown that

$$(2.4.19) \quad E \left| \sum_{i=1}^n \xi_i^2 - n\sigma^2 \right| 1_{\{t > n\}} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Therefore, the fact that M_n and $\sum_1^n (\xi_i^2 - \sigma^2)$ are martingales, (2.4.16), (2.4.19) and another application of Theorem 1 of CRT (1965) yields,

$$\begin{aligned} E(M_t - \sum_{i=1}^t \xi_i^2 + t\sigma^2)^2 \\ &= E \sum_{i=1}^t E[\{X_{i-1}^2(\xi_i^2 - \sigma^2) + 2X_{i-1}\xi_i D_{i-1} - (\xi_i^2 - \sigma^2)\}^2 | \mathcal{F}_{i-1}] \\ (2.4.20) \quad &= E(\xi_1^2 - \sigma^2)^2 E \sum_{i=1}^t (X_{i-1}^2 - 1)^2 + 4 E \xi_1^2 E \sum_{i=1}^t X_{i-1}^2 D_{i-1}^2 \\ &\quad + 4 E(\xi_1^2 - \sigma^2) \xi_1 E \sum_{i=1}^t (X_{i-1}^2 - 1) X_{i-1} D_{i-1}. \end{aligned}$$

Now take expectation on both sides of (2.4.17) and use (iv), (2.4.18) and (2.4.20) to get the required result.

The proof of (vi) follows from (2.4.12), (2.4.16),

(ii), (iv) and arguments similar to the ones in the proof of (v). Hence the lemma. $\square$

Using the above Lemmas and the results in Sections 2.2 and 2.3 we will now carry out the analysis of terms (a), (b)

and (c) in (2.4.2). In what follows we will assume as many moments as we need.

Assume w.l.o.g that $\sigma^2 = 1$. Now, by Lemma 2.17 (ii) it readily follows that

$$(2.4.21) \quad (a) = 0.$$

As for (c), write

$$(2.4.22) \quad \begin{aligned} A^{-1} t I_t - \theta &= a^{-2} t (I_t - \theta t) + \theta (a^{-2} t^2 - 1) \\ &= (a^{-1} t) a^{-1} [I_t - \theta t] + \theta (a^{-1} t - 1)^2 + 2\theta (a^{-1} t - 1). \end{aligned}$$

Since by definition $\lambda_t^{-3} \leq \theta^{-3} + (A t^{-1} I_t^{-1})^3$ we have

$$\begin{aligned} E |(D_t^2 - I_t) \lambda_t^{-3} (A^{-1} t I_t - \theta)^2| \\ \leq \theta^{-3} E |(D_t^2 - I_t) (A^{-1} t I_t - \theta)^2| + E |(D_t^2 - I_t) \cdot \\ \cdot (A t^{-1} I_t^{-1})^3 (A^{-1} t I_t - \theta)^2| \\ = (c_1) + (c_2), \quad \text{say..} \end{aligned}$$

We will show that both (c_1) and (c_2) are bounded as $A \rightarrow \infty$, by using the decomposition in (2.4.22).

Consider (c_1) . Now,

$$\begin{aligned} E |(D_t^2 - I_t) (a^{-1} t)^2 a^{-2} [I_t - \theta t]^2| \\ \leq E (a^{-1} D_t^2 + a^{-1} I_t) (a^{-1} t)^2 \{a^{-1} [I_t - \theta t]^2\} \end{aligned}$$

$$(2.4.23) \quad = O(1).$$

by the Holder inequality and the Lemmas 2.15, 2.14, 2.3, 2.16. Moreover, the Holder inequality, (2.3.27), and the

Lemmas 2.3, 2.14 and 2.15 yield

$$\begin{aligned} E|D_t^2 - I_t|(a^{-1}t - 1)^4 &\leq E(a^{-1}D_t^2 + a^{-1}I_t)a^{-1}(t^*)^4 \\ (2.4.24) \qquad \qquad \qquad &\rightarrow 0, \end{aligned}$$

and also

$$(2.4.25) \quad E|D_t^2 - I_t|(a^{-1}t - 1)^2 = o(1).$$

Hence

$$(2.4.26) \qquad \qquad \qquad (c_1) = o(1).$$

Consider (c_2) . Now,

$$(c_2) \leq E(a^{-1}D_t^2 + a^{-1}I_t)(a^2t^{-2})^3(\bar{I}_t)^{-3}.$$

$$\cdot [(a^{-1}t)^2 a^{-1}\{I_t - \theta t\}^2 + a^{-1}(t^*)^4 + (t^*)^2]$$

$$(2.4.27) \qquad = o(1),$$

by similar reasons given for the boundedness of (c_1) and by Lemma 2.13. Hence, from (2.4.26) and (2.4.27) we have that

$$(2.4.28) \qquad \qquad \qquad (c) = o(1).$$

Consider (b) of (2.4.2). Use (2.4.22) to write.

$$(b) = -\theta^{-2} E(D_t^2 - I_t)(a^{-1}t)a^{-1}[I_t - \theta t]$$

$$- \theta^{-1} E(D_t^2 - I_t)(a^{-1}t - 1)^2$$

$$- 2\theta^{-1} E(D_t^2 - I_t)(a^{-1}t - 1)$$

$$(2.4.29) \qquad = (b_1) + (b_2) + (b_3).$$

By (2.4.25) we have $(b_2) = O(1)$. Consider (b_1) .

Write

$$\begin{aligned} (b_1) &= -\theta^{-2} E(D_t^2 - I_t)(a^{-1}t - 1) a^{-1}[I_t - \theta t] \\ &\quad - \theta^{-2} E(D_t^2 - I_t)a^{-1}[I_t - \theta t] \\ &= (b_{11}) + (b_{12}). \end{aligned}$$

Now,

$$\begin{aligned} (b_{11}) &\leq E(a^{-1}D_t^2 + a^{-1}I_t)(t^*)a^{-1/2}|I_t - \theta t| \\ &= O(1), \end{aligned}$$

by similar arguments as before.

In the entire "regret" analysis the main point therefore is to analyse (b_{12}) and (b_3) . We will show that $(b_3) = O(1)$, and $(b_{12}) = -A^{-1/2} E(D_t^2 - I_t)X_t^2 + O(1)$.

Consider (b_3) . Recall the definition of W_n from (2.3.16), and the fact in (2.3.17). We will first show that

$$(2.4.30) \quad E(D_t^2 - I_t)(a^{-1}t - 1) = E(D_t^2 - I_t)[(t^{-1}W_t)^{1/2} - 1] + O(1)$$

(see t in (2.1.9) for motivation).

To do this, observe that by (2.1.9)

$$a^{-1}t \geq (t^{-1}W_t)^{1/2} \quad \text{and} \quad (t-1)a^{-1} \leq [(t-1)^{-1}W_{t-1}]^{1/2}.$$

From these,

$$\begin{aligned} ED_t^2(t^{-1}W_t)^{1/2} &\leq ED_t^2 a^{-1}t \leq ED_t^2 a^{-1}(t-1) \\ &\quad + E a^{-1} D_t^2 \\ &\leq E D_t^2 [(t-1)^{-1}W_{t-1}]^{1/2} + O(1) \quad [\text{see Lemma 2.15}] \\ &\leq E D_t^2 [(t-1)^{-1}W_t]^{1/2} + O(1) \end{aligned}$$

$$= E D_t^2 (t^{-1} w_t)^{1/2} + E D_t^2 (t^{-1} w_t)^{1/2} [\{t/(t-1)\}^{1/2} - 1] + O(1)$$

$$\leq E D_t^2 [t^{-1} w_t]^{1/2} + E D_t^2 a^{-1} t [\{t/(t-1)\}^{1/2} - 1] + O(1)$$

$$(2.4.31) \quad = E D_t^2 [t^{-1} w_t]^{1/2} + E D_t^2 a^{-1} t [(1/2)(t-1)^{-1} - (1/8)(t-1)^{-2} (\lambda'_t)^{-3/2}] + O(1),$$

by the Taylor's theorem, where λ'_t is a r.v. lying between 1 and $[t/(t-1)]$.

Using Lemma 2.15 and $(\lambda'_t)^{-3/2} \leq 1 + [t/(t-1)]^{-3/2}$, the second expression in the R.H.S of (2.4.31) can be shown to be $O(1)$. Therefore,

$$E D_t^2 (a^{-1} t - 1) = E D_t^2 [(t^{-1} w_t)^{1/2} - 1] + O(1).$$

Arguments similar to the ones above uses Lemmas 2.14 and 2.3 to get

$$E I_t (a^{-1} t - 1) = E I_t [(t^{-1} w_t)^{1/2} - 1] + O(1).$$

Therefore (2.4.30) follows.

Now, again by the Taylor's theorem

$$E(D_t^2 - I_t) [(t^{-1} w_t)^{1/2} - 1] = (1/2) E(D_t^2 - I_t) [t^{-1} w_t - 1]$$

$$(2.4.32) \quad - (1/8) E (\lambda''_t)^{-3/2} (D_t^2 - I_t) [t^{-1} w_t - 1]^2 + O(1),$$

where λ''_t is a r.v. lying between $t^{-1} w_t$ and 1.

Using similar arguments as before we can show that

$$E (\lambda''_t)^{-3/2} (D_t^2 - I_t) [t^{-1} w_t - 1]^2 = O(1).$$

As for the first expression in the R.H.S of (2.4.32), from (2.3.16) and (2.2.5) we have

$$\begin{aligned}
E(D_t^2 - I_t)[t^{-1}W_t - 1] &= E(D_t^2 - I_t)t^{-1}\left(\sum_{i=1}^t \varepsilon_i^2 - t\right) \\
&\quad - E(D_t^2 - I_t)C_t \\
&= E(D_t^2 - I_t)(t^{-1} - a^{-1})\left(\sum_{i=1}^t \varepsilon_i^2 - t\right) - E(D_t^2 - I_t)a^{-1}\left(\sum_{i=1}^t \varepsilon_i^2 - t\right) \\
&\quad - E(D_t^2 - I_t)C_t \\
&= (b_{31}) + (b_{32}) + (b_{33}).
\end{aligned}$$

Now,

$$\begin{aligned}
|(b_{31})| &\leq E(a^{-1}D_t^2 + a^{-1}I_t)|t^*|(a t^{-1})a^{-1/2}\left|\sum_{i=1}^t \varepsilon_i^2 - t\right| \\
&= O(1),
\end{aligned}$$

by similar arguments as before and by Lemma 5 of Chow and Yu (1981).

As for (b_{33}) it follows from the similar arguments given earlier that,

$$\begin{aligned}
|(b_{33})| &\leq E(a^{-1}D_t^2 + a^{-1}I_t)aC_t \\
&= E(a^{-1}D_t^2 + a^{-1}I_t)(a t^{-1})^2(a^{-1}D_t^2)(\bar{I}_t)^{-1} \\
&= O(1).
\end{aligned}$$

Consider (b_{32}) . For this use Lemma 2.17 (v) to get

$$(2.4.33) \quad (b_{32}) = -2\mu_3 a^{-1}E \sum_{i=1}^t X_{i-1}D_{i-1} - \sigma_1^2 a^{-1}E I_t.$$

By Lemmas 2.3 and 2.14 we have

$$\begin{aligned}
(2.4.34) \quad a^{-1}E I_t &= E(a^{-1}t) \bar{I}_t \\
&= O(1).
\end{aligned}$$

As for the first expression in the R.H.S. of (2.4.33), write

$$\begin{aligned}
a^{-1} E \sum_{i=1}^t X_{i-1} D_{i-1} &= a^{-1} E \sum_{i=1}^t X_{i-1} (D_t - \sum_{j=1}^t X_{j-1} \varepsilon_j) \\
&= a^{-1} E D_t \sum_{i=1}^t X_{i-1} - a^{-1} E \sum_{i=1}^t X_{i-1} \left(\sum_{j=1}^t X_{j-1} \varepsilon_j \right) \\
&= (b_{321}) + (b_{322}).
\end{aligned}$$

Consider (b_{322}) . If $E \sum_{i=1}^t \sum_{j=1}^t |X_{i-1} X_{j-1} \varepsilon_j| < \infty$, then,

$$\begin{aligned}
(b_{322}) &= E \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} X_{i-1} X_{j-1} \varepsilon_j 1_{[t \geq j]} \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} E X_{i-1} X_{j-1} 1_{[t \geq j]} E \varepsilon_j \\
&= 0,
\end{aligned}$$

because the event $[t \geq j]$, $X_{i-1} X_{j-1}$ are all measurable w.r.t to $\mathcal{F}_{j-1}$ and independent of ε_j , for all $i < j$.

For (b_{321}) , use (1.1.2) to write

$$\begin{aligned}
a^{-1} E D_t \sum_{i=1}^t X_{i-1} &= a^{-1} E D_t X_0 + a^{-1} E D_t \left\{ \sum_{j=1}^{t-1} (\beta^j X_0 \right. \\
&\quad \left. + \sum_{k=1}^j \beta^{j-k} \varepsilon_k) \right\}
\end{aligned}$$

By Lemma 2.15, and the Schwarz inequality we have

$$a^{-1} E D_t X_0 \rightarrow 0.$$

Also by similar arguments

$$\begin{aligned}
a^{-1} E D_t X_0 \sum_{j=1}^{t-1} \beta^j &= a^{-1} E D_t X_0 \beta(1-\beta^{t-1})(1-\beta)^{-1} \\
&= O(1).
\end{aligned}$$

By the interchange of the order of summation

$$\begin{aligned} a^{-1} ED_t \sum_{j=1}^{t-1} \sum_{k=1}^j \beta^{j-k} \xi_k &= a^{-1} E D_t \sum_{k=1}^{t-1} \xi_k \sum_{j=k}^{t-1} \beta^{j-k} \\ &= a^{-1} ED_t \sum_{k=1}^{t-1} \xi_k (1-\beta^{t-k})(1-\beta)^{-1}. \end{aligned}$$

$$= (1-\beta)^{-1} a^{-1} ED_t \sum_{k=1}^{t-1} \xi_k - (1-\beta)^{-1} a^{-1} ED_t \sum_{k=1}^{t-1} \beta^{t-k} \xi_k.$$

By Lemma 5 of Chow and Yu (1981) and Lemma 2.15 and the Schwarz inequality

$$a^{-1} ED_t \sum_{k=1}^{t-1} \xi_k = E(a^{-1/2} D_t) \cdot (a^{-1/2} \sum_{k=1}^{t-1} \xi_k) = o(1).$$

By (1.1.2)

$$a^{-1} ED_t \sum_{k=1}^{t-1} \beta^{t-k} \xi_k = a^{-1} ED_t (X_t - \beta^t X_0 - \xi_t) = o(1),$$

by Lemma 2.15, since $E|X_t|^2 = o(a)$ (see (2.4.10)) and

$E|\xi_t|^2 = o(a)$ (see Chow and Teicher (1978), Lemma 2,

p. 148). Thus from the above arguments we have that

$(b_{32}) = o(1)$. Hence $(b_3) = o(1)$.

It only remains to analyse (b_{12}) . For this, use (2.4.8) to get

$$\begin{aligned} (b_{12}) &= a^{-1} E(D_t^2 - I_t) X_0^2 - a^{-1} E(D_t^2 - I_t) X_t^2 \\ &\quad + 2\beta a^{-1} E(D_t^2 - I_t) D_t \\ &\quad + a^{-1} E(D_t^2 - I_t) \left(\sum_{i=1}^t \xi_i^2 - t \right) \\ &= (b_{121}) + (b_{122}) + (b_{123}) + (b_{124}). \end{aligned}$$

By the Schwarz inequality, and Lemmas 2.14, 2.15 and 2.3 we have

$$(b_{121}) = 0(1).$$

Now for (b_{123}) , use Lemma 2.17 (vi) to get

$$(b_{123}) = a^{-1} \mu_3 E \sum_{i=1}^t X_{i-1}^3 + 2a^{-1} E \sum_{i=1}^t X_{i-1}^2 D_{i-1}.$$

It follows from Lemmas 2.3 and 2.14 and the Schwarz inequality that

$$a^{-1} \mu_3 E \sum_{i=1}^t X_{i-1}^3 = E (a^{-1} t) (t^{-1} \sum_{i=1}^t X_{i-1}^3) = 0(1).$$

From CRT (1965) [see proof of Lemma 8, display (27) p. 794-795) by taking

$$u_{2,i} = E(X_{i-1}^2 \epsilon_i^2 | \mathcal{F}_{i-1}) = X_{i-1}^2; \quad S_{i-1} = D_{i-1}$$

(The u 's and S 's are defined in pages 793 and 789 in CRT (1965)) we get

$$\begin{aligned} a^{-1} E \sum_{i=1}^t X_{i-1}^2 D_{i-1} &= a^{-1} E D_t I_t \\ &= a^{-1} E D_t [I_t - \theta t] + \theta a^{-1} E D_t (t - a) + \theta E D_t \\ &= E a^{-1/2} D_t a^{-1/2} [I_t - \theta t] + \theta E (a^{-1/2} D_t) t^* \\ &= 0(1), \end{aligned}$$

by Lemmas 2.15, 2.16, (2.17)(i) and (2.3.27).

From the arguments above we have

$$(b) = (b_{12}) + 0(1) = -a^{-1} E(D_t^2 - I_t) X_t^2 + 0(1). \quad \square$$

BIBLIOGRAPHY

- Anderson, T.W. (1959). On asymptotic distribution of estimates of parameters of Stochastic difference equations. Ann. Math. Statist. 30 676-687.
- Anderson, T.W. (1971). The Statistical Analysis of Time Series. Wiley, New York.
- Breiman, L. (1968). Probability. Addison-Wesley, Reading, Massachusetts.
- Chow, Y.S. and Martinsek, A.T. (1982). Bounded regret of a sequential procedure for estimation of the mean. Ann. Statist. 10 909-914.
- Chow, Y.S. and Robbins, H.E. (1965). On the asymptotic theory of fixed width confidence interval for the mean. Ann. Math. Statist. 36 457-462.
- Chow, Y.S., Robbins, H., and Siegmund, D. (1971). Great Expectations: The theory of optimal stopping. Houghton Mifflin. Boston.
- Chow, Y.S., Robbins, H., and Teicher, H. (1965). Moments of randomly stopped sums. Ann. Math. Statist. 36 789-799.
- Chow, Y.S. and Teicher, H. (1978). Probability theory: Independence, Interchangeability, Martingales. Springer-Verlag, New York.
- Chow, Y.S. and Yu, K.F. (1981). The performance of a sequential procedure for the estimation of the mean. Ann. Statist. 9 184-189.
- Doob, J.L. (1953). Stochastic Processes. Wiley. New York.
- Finster, M. (1983). A frequentistic approach to sequential estimation in the general linear model. Jour. Amer. Statist. Assoc. 78 403-407.
- Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. Commun. Statist.-Theor. Meth. A8 637-651.
- Hagwood, C., and Woodroffe, M. (1982). On the expansion for expected sample size in non-linear renewal theory. Ann. Prob. 10 844-848.
- Jureckova, J. and Sen, P.K. (1982). M-estimators and L-estimators of location: Uniform integrability and asymptotic risk-efficient sequential version. Commun. Statist.-Sequential Analysis 1 27-56.

- Lai, T.L., and Siegmund, D. (1977). A non-linear renewal theory with application to sequential analysis, I. Ann. Statist. 5 946-954.
- Lai, T.L., and Siegmund, D. (1979). Non-linear renewal theory with application to sequential analysis II. Ann. Statist. 7 60-76.
- Lai, T.L., and Siegmund, D. (1983). Fixed accuracy estimation of an autoregressive parameter. Ann. Statist. 11 478-485.
- Martinsek, A.T. (1983). Second order approximation to the risk of a sequential procedure. Ann. Statist. 11 827-836.
- Martinsek, A.T. (1984). Sequential determination of estimator as well as sample size. Ann. Statist. 12 533-550.
- Martinsek, A.T. (1985). Comparison of some sequential procedures with related optimal stopping rules. Zeit. Wahr. 70 411-416.
- Robbins, H.E. (1959). Sequential estimation of the mean of a normal population. Probability and Statistics (Harald Cramer volume). Almqvist-Uppsala., 235-245.
- Sen, P.K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U-Statistics. Sankhya, Ser. A 43, 331-344.
- Sen, P.K. (1981). Sequential Nonparametrics: Invariance Principles and Statistical Inference. Wiley, New York.
- Starr, N. (1966b). On the asymptotic efficiency of a sequential procedure for estimating the mean. Ann. Math. Statist. 37 1173-1185
- Starr, N. and Woodroffe, M. (1969). Remarks on sequential point estimation. Proc. Nat. Acad. Sci. U.S.A. 63 285-288.
- Starr, N. and Woodroffe, M. (1972). Further remarks on sequential point estimation: the exponential case. Ann. Math. Statist. 43 1147-1154.
- Stout, W.F. (1974). Almost sure convergence. Academic Press. New York.
- Subramanyam, A. (1984). Contributions to the asymptotic theory of sequential procedures. University of Poona Ph.D. Dissertation.

Vardi, Y. (1979). Asymptotic optimal sequential estimation: the poisson case. Ann. Statist. 7 1040-1051.

Woodroffe, M. (1977). Second order approximations for sequential point and interval estimation. Ann. Statist. 5 984-995.

Woodroffe, M. (1982). Nonlinear Renewal Theory in Sequential Analysis. Society for Industrial and Applied Mathematics, Philadelphia.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03177 5939