

AD-A119 373

STANFORD UNIV CA DEPT OF STATISTICS

F/G 12/1

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER.(U)

MAY 82 T L LAI, D SIEGMUND

N00014-77-C-0306

NL

UNCLASSIFIED

TR-16

1-1

1-1



END
DATE
FILMED
10-82
DTIC

(12)

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER

BY

T. L. LAI and D. SIEGMUND

TECHNICAL REPORT NO. 16

MAY 1982

PREPARED UNDER CONTRACT

N00014-77-C-0306 (NR-042-373)

FOR THE OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited.

AD A119373

DTIC FILE COPY

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA



SEP 20 1982
A

82 09 20 00 3

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER

by

T.L. Lai
Columbia University

and

D. Siegmund
Stanford University

TECHNICAL REPORT NO. 16
May 1982

Prepared under the Auspices
of
Office of Naval Research Contract
N00014-77-C-0306 (NR-042-373)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A	



Also issued as Technical Report No. 189 under National Science
Foundation Grant MCS80-24649 Stanford University, Department
of Statistics.

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER

BY T. L. LAI and D. SIEGMUND

ABSTRACT

For a first order non-explosive autoregressive process with unknown parameter $\beta \in [-1, 1]$, it is shown that if data are collected according to a particular stopping rule, the least squares estimator of β is asymptotically normally distributed uniformly in β . In the case of normal residuals, the stopping rule may be interpreted as sampling until the observed Fisher information reaches a preassigned level. The situation is contrasted with the fixed sample size case, where the estimator has a non-normal limiting distribution when $|\beta| = 1$.

AMS 1970 Subject Classification: Primary 62L10

Key Words and Phrases. Stopping rule, fixed width confidence interval, uniform asymptotic normality.

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER

1. Introduction and Summary.

Consider the first order, non-explosive, autoregressive model

$$(1.1) \quad x_n = \beta x_{n-1} + \varepsilon_n \quad (n = 1, 2, \dots),$$

where $\varepsilon_1, \varepsilon_2, \dots$ are independent, identically, distributed random variables with $E\varepsilon_1 = 0$ and $0 < E\varepsilon_1^2 = \sigma^2 < \infty$. The initial state x_0 is a random variable (not depending on β) which is stochastically independent of $\{\varepsilon_n\}$. The constant $\beta \in [-1, 1]$ is an unknown parameter, which at stage n is customarily estimated by the least squares estimate,

$$(1.2) \quad b_n = \left(\sum_{i=1}^n x_{i-1} x_i \right) / \left(\sum_{i=1}^n x_{i-1}^2 \right) = \beta + \left(\sum_{i=1}^n x_{i-1} \varepsilon_i \right) / \left(\sum_{i=1}^n x_{i-1}^2 \right).$$

If the ε 's are normally distributed, then b_n is also the maximum likelihood estimator of β , and the observed Fisher information about β contained in $x_0, x_1, \dots, x_n$ is

$$(1.3) \quad - \frac{d^2}{d\beta^2} \left[\beta \sum_{i=1}^n x_{i-1} x_i - \beta^2 / 2 \sum_{i=1}^n x_{i-1}^2 \right] = \sum_{i=1}^n x_{i-1}^2.$$

It is well-known and easy to prove that for fixed $\beta \in (-1, 1)$, as $n \rightarrow \infty$

$$(1.4) \quad \left(\sum_{i=1}^n x_{i-1}^2 \right)^{1/2} (b_n - \beta) \xrightarrow{L} N(0, \sigma^2).$$

Here $N(\mu, \sigma^2)$ denotes a normal random variable with mean μ and variance σ^2 , and $\xrightarrow{L}$ indicates convergence in law. See, for example, Anderson (1959).

However, for $\beta = \pm 1$, an entirely different limiting distribution occurs (Rao, 1978). For example, if $\beta = 1$, $x_n - x_0$ is the sum of i.i.d. random variables $\epsilon_1 + \dots + \epsilon_n$, and summation by parts in (1.2) yields

$$\left(\sum_{i=1}^n x_{i-1}^2\right)^{1/2} (b_n - 1) = \frac{1}{2} n^{-1} (x_n^2 - x_0^2 - \sum_{i=1}^n \epsilon_i^2) / (n^{-2} \sum_{i=1}^n x_{i-1}^2)^{1/2}.$$

By Donsker's theorem this converges in law to

$$(1.5) \quad \frac{\sigma}{2} (W^2(1) - 1) / \left(\int_0^1 W^2(t) dt\right)^{1/2},$$

where $\{W(t), 0 \leq t \leq 1\}$ is a standard Brownian motion process. Of course, this result indicates that the asymptotic normality of (1.4) breaks down for β in a neighborhood of ± 1 , in the sense that for any given n , no matter how large, there will be a neighborhood of ± 1 in which one should not expect (1.4) to yield reasonable approximations.

Examples in econometrics having values of β close to 1 are cited by Evans and Savin (1981) and by Fuller (1981).

In this paper we consider the asymptotic behavior of $\{b_n\}$ under a sequential sampling scheme which measures time in terms of accumulated (observed) Fisher information. Define

$$(1.6) \quad N_c = \text{first } n \geq 1 \text{ such that } \sum_{i=1}^n x_{i-1}^2 \geq c\sigma^2.$$

Our principal result (Theorem 2.1) is that as $c \rightarrow \infty$

$$(1.7) \quad \left(\sum_{i=1}^{N_c} x_{i-1}^2\right)^{1/2} (b_{N_c} - \beta) \xrightarrow{d} N(0, \sigma^2)$$

uniformly in β for $-1 \leq \beta \leq 1$.

The sampling rule (1.6) is motivated by the theory of fixed width confidence intervals, cf. Anscombe (1953), Chow and Robbins (1965), or Grambsch (1982). Some precedent for the rather surprising uniformity in the convergence of (1.7) is found in the work of Siegmund (1981), although the underlying reasons are quite different in the present case.

When it is feasible to use the sampling rule (1.6), its advantages appear to be threefold: (i) the accuracy of b_{N_c} as an estimator of β (as measured by the variance of its asymptotic distribution) is approximately a small constant c^{-1} , rather than an uncontrolled random variable, (ii) the appropriate asymptotic distribution theory does not depend on the value of the unknown parameter β ; and (iii) the convergence to asymptotic normality is much more rapid, even when β is not near the values ± 1 . (See Section 3.)

The remainder of this paper is arranged as follows. Section 2 contains a proof of (1.7). In Section 3 we give the results of some simulations comparing confidence intervals obtained by indiscriminate use of (1.4) for fixed sample sizes with (1.7) for sequentially determined sample sizes. Section 4 is concerned with some related asymptotic results, and in particular the appropriate modification of (1.6) when σ is unknown.

Partial results for the model $x_n = \alpha + \beta x_{n-1} + \varepsilon_n$ indicate that the multiparameter case can be appreciably more complicated. We expect to consider this problem in the future.

2. Uniform Asymptotic Normality of b_{N_c}

The main result of this section is the proof of (1.7) (Theorem 2.1 below). Our approach is motivated by the observation that $b_n - \beta = \frac{\sum_{i=1}^n x_{i-1} \varepsilon_i}{\sum_{i=1}^n x_{i-1}^2}$ is of the form of a martingale divided by the sum of

the conditional variances of its increments. The novelty of our approach lies in systematic exploitation of the condition (2.6) and the simple identity (2.16). We begin with some preliminary probabilistic results.

Proposition 2.1. Let $x_n, \varepsilon_n, n = 0, 1, \dots$ be random variables adapted to the increasing sequence of σ -algebras $\mathcal{F}_n, n = 0, 1, \dots$. Let $\{P_\theta, \theta \in \Xi\}$ be a family of probability measures such that under every P_θ

$$(2.1) \quad \varepsilon_1, \varepsilon_2, \dots \text{ are i.i.d. with } E_\theta \varepsilon_1 = 0, E_\theta \varepsilon_1^2 = 1;$$

$$(2.2) \quad \sup_\theta E_\theta \{ \varepsilon_1^2; |\varepsilon_1| > a \} \rightarrow 0 \text{ as } a \rightarrow \infty;$$

$$(2.3) \quad \varepsilon_n \text{ is independent of } \mathcal{F}_{n-1} \text{ for each } n \geq 1;$$

$$(2.4) \quad P_\theta \{ \sum_{i=0}^{\infty} x_i^2 = \infty \} = 1;$$

$$(2.5) \quad \sup_\theta P_\theta \{ x_n^2 > a \} \rightarrow 0 \text{ as } a \rightarrow \infty \text{ for each } n \geq 0;$$

$$(2.6) \quad \lim_{n \rightarrow \infty} \sup_\theta P_\theta \{ x_n^2 \geq \delta \sum_{i=0}^{n-1} x_i^2 \text{ for some } n \geq n \} = 0.$$

For $c > 0$ let

$$(2.7) \quad T_c = \inf \{ n: \sum_{i=1}^n x_{i-1}^2 \geq c \} \quad (\inf \emptyset = +\infty).$$

Then uniformly in $\theta \in \Xi$ and $-\infty < t < \infty$

$$(2.8) \quad P_\theta \{ c^{-1/2} \sum_{i=1}^T x_{i-1} \varepsilon_i \leq t \} \rightarrow \Phi(t) \text{ as } c \rightarrow \infty.$$

Remark. For the autoregressive model (1.1), if we identify θ with the autoregressive parameter β , all the conditions of Proposition 2.1 are trivially verified, with the single exception of (2.6).

Proposition 2.1 is proved by reducing it to the following convenient martingale central limit theorem (cf. Freedman, 1971, pp. 90-92).

Lemma 2.1. Given $0 < \delta < 1$, there exists $\rho(\delta) > 0$ with $\lim_{\delta \rightarrow 0} \rho(\delta) = 0$ such that for any martingale difference sequence $\{u_n, \mathcal{F}_n, n \geq 0\}$ satisfying

$$(2.9) \quad |u_n| \leq \delta \text{ for all } n$$

and

$$(2.10) \quad \sum_{1}^{\infty} E(u_n^2 | \mathcal{F}_{n-1}) > r \quad \text{a.s.},$$

if $\tau = \inf\{n: \sum_{1}^n E(u_i^2 | \mathcal{F}_{i-1}) \geq r\}$, then $P\{\tau < \infty\} = 1$

and

$$\sup_x |P\{\sum_{1}^{\tau} u_i \leq x\} - \Phi(x/r^{1/2})| \leq \rho(\delta/r^{1/2}).$$

Proof of Proposition 2.1.

By (2.4) $P_{\theta}\{T_c < \infty\} = 1$, and obviously $P_{\theta}\{\lim_{c \rightarrow \infty} T_c = \infty\} = 1$ for all θ . Let $0 < \delta < 1$ and define $\tilde{x}_n = x_n$ if $x_n^2 \leq \delta^2 c$ and $\tilde{x}_n = \delta c^{1/2}$ otherwise. Then for all θ

$$\begin{aligned} & P_{\theta}\{x_n \neq \tilde{x}_n \text{ for some } n < T_c\} \\ & \leq \sum_{i=1}^m P_{\theta}\{x_{i-1}^2 > \delta^2 c\} + P_{\theta}\{T_c > m, x_n \neq \tilde{x}_n \text{ for some } m \leq n < T\} \\ & \leq \sum_{i=1}^m P_{\theta}\{x_{i-1}^2 > \delta^2 c\} + P_{\theta}\{x_n^2 \geq \delta^2 \sum_{0}^{n-1} x_i^2 \text{ for some } n \geq m\}, \end{aligned}$$

which is $\leq 2\delta$ if we first choose m large enough and use (2.6), then choose c large and use (2.5). Hence if

$$\Omega_c = \{x_n = \tilde{x}_n \text{ for all } n < T_c\} ,$$

then for all large c

$$(2.11) \quad P(\Omega_c) \geq 1 - 2\delta$$

Define $\tilde{\epsilon}_n = \epsilon_n I\{|\epsilon_n| \leq \delta^{-1/2}\}$ and $\tilde{\tilde{\epsilon}}_n = \epsilon_n - \tilde{\epsilon}_n$. Then under P_θ $\{c^{-1/2} \tilde{x}_{n-1}(\tilde{\epsilon}_n - E_\theta \tilde{\epsilon}_n), \mathcal{F}_n, 0 \leq n < \infty\}$ is a martingale difference sequence satisfying

$$|c^{-1/2} \tilde{x}_{n-1}(\tilde{\epsilon}_n - E_\theta \tilde{\epsilon}_n)| \leq 2\delta^{1/2} ,$$

and by (2.4)

$$P_\theta \left\{ \sum_0^\infty \tilde{x}_n^2 = \infty \right\} = 1$$

also by (2.2), as $\delta \rightarrow 0$

$$(2.12) \quad v_\theta(\delta) = \text{var}_\theta \tilde{\epsilon}_1 \rightarrow 1$$

uniformly in θ . Hence, if

$$\tau_c = \inf\{n: \sum_1^n \tilde{x}_{i-1}^2 \geq c\} ,$$

then by Lemma 2.1

$$(2.13) \quad |P_{\theta}\{c^{-1/2} \sum_1^{\tau} \tilde{x}_{i-1}(\tilde{\epsilon}_i - E_{\theta} \tilde{\epsilon}_i) \leq t\} - \Phi(t/(v_{\theta}(\delta))^{1/2})| \leq \rho[2(\delta/v_{\theta}(\delta))^{1/2}] .$$

On Ω_c , $\tau_c = T_c$ and

$$(2.14) \quad \left| \sum_1^{\tau} \tilde{x}_{i-1}(\tilde{\epsilon}_i - E_{\theta} \tilde{\epsilon}_i) - \sum_1^T x_{i-1} \epsilon_i \right| = \left| \sum_1^{\tau} \tilde{x}_{i-1}(\tilde{\epsilon}_i - E_{\theta} \tilde{\epsilon}_i) \right| .$$

By Wald's identity (cf. Chow, Robbins, and Siegmund, 1971, p. 23)

$$(2.15) \quad E_{\theta}[c^{-1/2} \sum_1^{\tau} \tilde{x}_{i-1}(\tilde{\epsilon}_i - E_{\theta} \tilde{\epsilon}_i)]^2 = c^{-1}(1-v_{\theta}(\delta))E_{\theta}(\sum_1^{\tau} \tilde{x}_{i-1}^2) \\ \leq (1-v_{\theta}(\delta))(1+\delta^2)$$

The Proposition follows from (2.11)-(2.15) by letting $\delta \rightarrow 0$.

For ease of reference we state without proof the following lemma.

It is related to the strong law of large numbers given e.g. by Neveu (1965, p. 148), whose method of proof can be adapted to the present purpose. Alternatively it follows easily from Proposition 2 of Robbins and Siegmund (1971) and a straightforward calculation along the lines of their Lemma 1.

Lemma 2.2. Under the measurability assumptions of Proposition 2.1, (2.1), and (2.3), for each $\gamma > \frac{1}{2}$ and increasing sequence of positive constants $c_n \rightarrow +\infty$, and for each $\delta > 0$

$$\sup_{\theta} P_{\theta}\{|\sum_1^n x_{i-1} \epsilon_i| \geq \delta \max(c_n, (\sum_1^n x_{i-1}^2)^{\gamma}) \text{ for some } n \geq m\} \rightarrow 0$$

as $m \rightarrow \infty$.

We now return to the autoregressive model (1.1) and write P_{β} to denote dependence of probabilities on the parameter β . (The joint distribution of

$x_0, \varepsilon_1, \varepsilon_2, \dots$, however, is assumed not to depend on β .) The principal result of this section is

Theorem 2.1. Define $\{b_n, n = 2, 3, \dots\}$ by (1.2) and N_c by (1.6).

If $\varepsilon_1, \varepsilon_2, \dots$ are i.i.d. with mean 0 and variance σ^2 , and are independent of x_0 , then

$$\lim_{c \rightarrow \infty} P_{\beta} \left\{ \left(\sum_{i=1}^{N_c} x_{i-1}^2 \right) (b_{N_c} - \beta) \leq t \right\} = \Phi(t/\sigma)$$

uniformly for $-1 \leq \beta \leq 1$ and $-\infty < t < \infty$.

Proof. Theorem 2.1 follows immediately from Proposition 2.1 once we have verified (2.6). To this end, note that squaring (1.1) and summing yields

$$(2.16) \quad x_n^2 + (1-\beta^2) \sum_{i=0}^{n-1} x_i^2 - x_0^2 = \sum_{i=1}^n \varepsilon_i^2 + 2\beta \sum_{i=1}^n x_{i-1} \varepsilon_i.$$

Let $|\beta| \leq 1$ and $0 < \lambda < \sigma^2/4$. Let

$$(2.17) \quad \Omega_{m,\lambda} = \left\{ \left| n^{-1} \sum_{i=1}^n \varepsilon_i^2 - \sigma^2 \right| < \lambda, \left| \sum_{i=1}^n x_{i-1} \varepsilon_i \right| < \max(\lambda n, \left(\sum_{i=1}^n x_{i-1}^2 \right)^{2/3}) \text{ for all } n \geq m \right\}.$$

On $\Omega_{m,\lambda}$, if $n \geq m$ and $x_n^2 \leq \lambda n$, then (2.16) implies

$$\sum_{i=0}^{n-1} x_i^2 \geq (\sigma^2 - \lambda)n - \lambda n - 2 \max(\lambda n, \left(\sum_{i=1}^n x_{i-1}^2 \right)^{2/3})$$

and hence for all m sufficiently large

$$(2.19) \quad 2 \sum_{i=0}^{n-1} x_i^2 \geq (\sigma^2 - 4\lambda)n \geq (\sigma^2 \lambda^{-1} - 4)x_n^2$$

On the other hand, since $|x_{n-1}| \geq |\beta x_{n-1}| \geq |x_n| - |\varepsilon_n|$, it follows that

$$\min_{1 \leq j \leq k} |x_{n-j}| \geq |x_n| - \sum_{j=0}^{k-1} |\varepsilon_{n-j}| \geq |x_n| \{1 - \sum_{j=0}^{k-1} |\varepsilon_{n-j}| / (\lambda n)^{1/2}\}$$

if $x_n^2 \geq \lambda n$. Hence $n \geq k$ and $x_n^2 \geq \lambda n$ imply

$$(2.19) \quad \sum_{i=1}^{n-1} x_{i-1}^2 / x_n^2 \geq k^2 \{1 - \sum_{j=0}^{k-1} |\varepsilon_{n-j}| / (\lambda n)^{1/2}\} ;$$

and since $\sum_{j=0}^{k-1} |\varepsilon_{n-j}|$, $n = k, k+1, \dots$ are i.i.d. with finite second moment, the right hand side of (2.19) converges to k^2 with probability one, at a rate which does not involve β . Hence by choosing λ so small that $2/(\sigma^2 \lambda^{-1} - 4) < \delta$ and $k^2 > 1/\delta$, we see by (2.18) and (2.19) that

$$\begin{aligned} \lim_{m \rightarrow \infty} \sup_{|\beta| \leq 1} P_{\beta} \{x_n^2 \geq \delta \sum_{i=1}^{n-1} x_i^2 \text{ for some } n \geq m\} \\ \leq \lim_{m \rightarrow \infty} \sup_{|\beta| \leq 1} P_{\beta}(\Omega_{m,\lambda}^c) , \end{aligned}$$

which equals 0 by Lemma 2.2 and the strong law of large numbers. This establishes (2.6) and hence the theorem.

3. Monte Carlo Results.

In this section we report the results of a Monte Carlo experiment to compare the fixed sample size and sequential asymptotic distributions. The basic experiment to assess the accuracy of the normal approximation indicated

by (1.4) and (1.7) consisted of a frequency count of the number of times the normalized estimator of β exceeded z or was less than $-z$ for commonly used quantiles z of the standard normal distribution. Since very similar results were obtained for various z , we report here only for $z = 1.28$, for which the (one-tailed) probabilities are nominally 0.10. For simplicity $x_0 = 0$, and $\epsilon_1, \epsilon_2, \dots$ were taken to be $N(0,1)$. In the fixed sample experiment, $n = 50$ observations were taken; and the results are reported in Table 1. In Table 2 observations are taken sequentially with c chosen so that $E_{\beta} N_c \approx 50$ for all β ; in Table 3 $c = 50$ and $E_{\beta} N_c$ varies with β .

For normally distributed ϵ_i , b_n (b_{N_c}) fails to be normally distributed only because of variability in $\sum_{i=1}^n x_{i-1}^2$ ($\sum_{i=1}^{N_c} x_{i-1}^2$). (See Dvoretzky, 1972, p. 520) Of course, the sequential experiment is designed to reduce this variability. The columns of Tables 1-3 below with the headings a_{xx} and s_{xx} report the observed average and standard deviation of $\sum_{i=1}^2 x_{i-1}^2$ respectively. Variability for $\sum_{i=1}^2 x_{i-1}^2$ in the fixed sample experiment goes into variability in N_c for the sequential experiment, so Tables 2 and 3 also report estimates of $E_{\beta} N_c$ and $(\text{var}_{\beta} N_c)^{1/2}$. Note that variability in $\sum_{i=1}^2 x_{i-1}^2$ for the fixed sample size case and $(\text{var}_{\beta} N_c)^{1/2} / E_{\beta} N_c$ for the sequential case increase dramatically as β approaches one.

The columns headed $\hat{p}_r$ and $\hat{p}_l$ report the percentage of excesses in the right and left tails of the distributions.

The figures in Tables 1 and 2 indicate that the fixed sample size asymptotic theory is not especially good when $n = 50$, even for small $|\beta|$,

TABLE 1

Fixed Sample Case

 $m = 50$, $z = 1.28$, p_r and p_l are nominally 0.10

β	$\hat{p}_r$	$\hat{p}_l$	a_{xx}	s_{xx}
0.1	.084	.108	49	10
0.5	.086	.118	65	18
0.9	.052	.136	232	136
1.0	.048	.166	1272	1505

TABLE 2

Sequential Case

 $z = 1.28$, p_r and p_l are nominally 0.10

β	$\hat{p}_r$	$\hat{p}_l$	c	a_{xx}	s_{xx}	$E_{\beta} N_c$	$(\text{var}_{\beta} N_c)^{1/2}$
0.1	.094	.105	50	52	1.7	52	10
0.5	.103	.096	65	67	2.2	52	13
0.9	.103	.099	200	207	8.0	51	21
1.0	.098	.101	500	524	22	48	24

TABLE 3

Sequential Case

 $c = 50$, $z = 1.28$, p_r and p_l are nominally 0.10

β	$\hat{p}_r$	$\hat{p}_l$	a_{xx}	s_{xx}	$E_{\beta}(N_c)$	$(\text{var}_{\beta} N_c)^{1/2}$
.5	.097	.099	52	2.4	41	11
.9	.092	.102	55	5.4	20	9.0
1.0	.108	.107	58	8.2	16	7.6
1.1	.096	.093	61	10	13	5.9

and it deteriorates quite noticeably for β near 1. In the sequential case the asymptotic theory is much better and shows no dependence upon the value of β .

4. Additional Asymptotic Theory.

Here we give some additional asymptotic results, which follow from the techniques developed in Section 2. Theorem 4.1 describes the asymptotic behavior of N_c , and to some extent explains the rather surprising differences between the fixed sample and sequential cases. Theorem 4.2 is concerned with the uniform strong consistency of b_n and $\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (x_i - b_n x_{i-1})^2$. It provides the foundation for consideration of the case of unknown σ .

Theorem 4.1. Under the conditions of Theorem 2.1,

(i) for each $\beta \in (-1, 1)$, $P_\beta \{ \lim_{c \rightarrow \infty} c^{-1} N_c = (1 - \beta^2) \} = 1$,

and

(ii) for $|\beta| = 1$, $c^{-1/2} N_c \stackrel{L}{\rightarrow} \inf \{ t : \int_0^t W^2(s) ds = 1 \}$,

where $W(t)$, $0 \leq t < \infty$, is a standard Brownian motion process.

Remark. In addition to containing information on the sample size of our sequential procedure, Theorem 4.1 has interesting connections with Theorem 2.1 and a theorem of Anscombe (1952). As generalized by Nagy (1962) Anscombe's theorem says that if $Y_n \stackrel{L}{\rightarrow} Y$, some additional technical conditions are satisfied, and if $v(c)$ are integer-valued random variables which can be normalized by constants $k(c) \rightarrow +\infty$ in such a way that $v(c)/k(c)$ converges in probability to a positive random variable, then $Y_{v(c)} \stackrel{L}{\rightarrow} Y$. Hence (1.4), Theorem 4.1 (i), and Anscombe's theorem show that $(\sum_{i=1}^{N_c} x_{i-1}^2)^{1/2} (b_{N_c} - \beta) \stackrel{L}{\rightarrow} N(0, 1)$

for each fixed $\beta \in (-1, 1)$. However, because the convergence of Theorem 4.1 (ii) is in law and not in probability, Anscombe's theorem is not applicable for $|\beta| = 1$, and in fact its conclusion would be incorrect.

Proof of Theorem 4.1.

On the event $\Omega_{m,\lambda}$ defined in (2.17), (2.16) implies for all $n \geq m$

$$(4.1) \quad |x_n^2 - x_0^2 + (1-\beta^2) \sum_{i=1}^n x_{i-1}^2 - n\sigma^2| \leq 3\lambda n + 2(\sum_{i=1}^n x_{i-1}^2)^{2/3}.$$

If $|\beta|$ is bounded away from 1, say $|\beta| \leq \rho < 1$, so $1 - \beta^2 \geq 1 - \rho^2 > 0$, then (2.6), (4.1), and $P_\beta(\Omega_{m,\lambda}) \rightarrow 1$ (uniformly in β) imply that

$$(4.2) \quad \lim_{m \rightarrow \infty} \sup_{|\beta| \leq \rho} P_\beta\{|(1-\beta^2)n^{-1} \sum_{i=1}^n x_{i-1}^2 - \sigma^2| \geq \delta \text{ for some } n \geq m\} = 0.$$

Theorem 4.1 (i) follows easily from (4.2).

If $\beta = 1$, then $x_n = x_0 + S_n$, where $S_n = \epsilon_1 + \dots + \epsilon_n$. Thus Theorem 4.1 (ii) follows from Donsker's Theorem, which implies that

$$\{c^{-1} \sum_{i=1}^{\lfloor c^{1/2} t \rfloor} S_i^2, t \geq 0\} \xrightarrow{d} \{\sigma^2 \int_0^t W^2(s) ds, t \geq 0\}.$$

A similar but slightly more complicated argument handles the case $\beta = -1$.

Theorem 4.2. Define the least squares estimate for β as in (1.2), and in the case of unknown σ , estimate σ^2 at stage n by

$$(4.3) \quad \hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (x_i - b_n x_{i-1})^2.$$

Then b_n and $\hat{\sigma}_n^2$ are uniformly consistent for $|\beta| \leq 1$ in the sense that for all $\delta > 0$

$$(4.4) \quad \lim_{m \rightarrow \infty} \sup_{|\beta| \leq 1} P_{\beta} \{ |b_n - \beta| \geq \delta \text{ for some } n \geq m \} = 0$$

and

$$(4.5) \quad \lim_{m \rightarrow \infty} \sup_{|\beta| \leq 1} P_{\beta} \{ |\hat{\sigma}_n^2 - \sigma^2| \geq \delta \text{ for some } n \geq m \} = 0 .$$

Proof. From (2.6), (2.16), and (4.2) it is easy to see that for all $0 < \lambda < 1$

$$\sup_{|\beta| \leq 1} P_{\beta} \{ n^{-1} \sum_{i=1}^n x_{i-1}^2 \leq \lambda \sigma^2 \text{ for some } n \geq m \} \rightarrow 0 \quad (m \rightarrow \infty) ,$$

which together with (1.2) and Lemma 2.2 proves (4.4).

To prove (4.5) note that

$$\begin{aligned} \hat{\sigma}_n^2 &= n^{-1} \sum_{i=1}^n (\epsilon_i + (\beta - b_n) x_{i-1})^2 \\ (4.6) \quad &= n^{-1} \sum_{i=1}^n \epsilon_i^2 - \left(\sum_{i=1}^n x_{i-1} \epsilon_i \right)^2 / n \sum_{i=1}^n x_{i-1}^2 . \end{aligned}$$

For $|\beta| \leq 1$, $|x_n| \leq |x_0| + \sum_{i=1}^n |\epsilon_i| = U_n$, say, where the distribution of U_n does not depend on β . Moreover, $\sum_{i=1}^n x_{i-1}^2 \leq \sum_{k=0}^{n-1} U_k^2 = o(n^4)$ with probability one. Hence by Lemma 2.2

$$\lim_{m \rightarrow \infty} \sup_{|\beta| \leq 1} P_{\beta} \{ \left(\sum_{i=1}^n x_{i-1} \epsilon_i \right)^2 \geq \delta_1 \sum_{i=1}^n x_{i-1}^2 \text{ for some } n \geq m \} = 0 ,$$

which along with (4.6) and the strong law of large numbers implies (4.5).

With the help of (4.5) it is possible to modify the definition of N_c to handle the case of unknown σ . The following result can be proved along

the lines of Theorem 2.1, although the details are considerably more complicated. The proof is omitted.

Theorem 4.3. Define b_n by (1.2) and $\hat{\sigma}_n^2$ by (4.3). Let $\{\delta_n\}$ be a sequence of positive constants with $\delta_n \rightarrow 0$. Let $s_n^2 = \max(\delta_n, \hat{\sigma}_n^2)$ and for $c > 0$ define

$$(4.7) \quad \hat{N}_c = \inf\{n: n \geq 2, \sum_{i=1}^n x_{i-1}^2 \geq cs_n^2\}.$$

Then as $c \rightarrow \infty$

$$P_{\beta}\left\{\left(\sum_{i=1}^{\hat{N}_c} x_{i-1}^2\right)^{1/2} (b_{\hat{N}_c} - \beta) / \hat{\sigma}_{\hat{N}_c} \leq t\right\} \rightarrow \Phi(t)$$

uniformly in $|\beta| \leq 1$ and $-\infty < t < \infty$.

REFERENCES

- Anscombe, F. J. (1952). Large sample theory of sequential estimation. Proc. Camb. Phil. Soc. 48, 600-607.
- Anscombe, F. J. (1953). Sequential estimation. J. R. Statist. Soc. B 15, 1-21.
- Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations, Ann. Math. Statist. 30, 676-687.
- Chow, Y. S. and Robbins, H. E. (1965). On the asymptotic theory of fixed width confidence intervals for the mean, Ann. Math. Statist. 36, 457-462.
- Chow, Y. S., Robbins, H. E. and Siegmund, D. (1971). Great Expectations: the Theory of Optimal Stopping, Houghton Mifflin, Boston.
- Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables, Proc. Sixth Berk. Symp. Math. Statist. and Prob., vol. II, Univ. of California Press, Berkeley, 513-535.
- Evans, G. B. A. and Savin, N. E. (1981). The calculation of the limiting distribution of the least squares estimator of the parameter in a random walk model, Ann. Statist. 9, 1114-1118.
- Freedman, D. (1971). Brownian Motion and Diffusion, Holden Day, San Francisco.
- Fuller, W. A. (1981). Stanford University Seminar Lecture.
- Grambsch, P. (1982). Sequential sampling based on the observed Fisher information to guarantee the accuracy of the maximum likelihood estimator, Bell Laboratories Technical Report.

- Mogyoródi, J. (1962). A central limit theorem for the sum of a random number of independent random variables, Publ. Math. Inst. Hung. Acad. Sci. 7, 409-424.
- Neveu, J. (1965). Mathematical Foundations of the Calculus of Probabilities, Holden Day, San Francisco.
- Rao, M. M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process, Ann. Statist. 6, 185-190.
- Robbins, H. and Siegmund, D. (1971). A convergence theorem for non-negative almost supermartingales and some applications, in Optimizing Methods in Statistics, ed. by J. S. Rustagi, Academic Press, New York, 233-258.
- Siegmund, D. (1981). A sequential confidence interval for the odds ratio, to appear in Probability and Mathematical Statistics.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 16	2. GOVT ACCESSION NO. A119323	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) T. L. LAI and D. SIEGMUND		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0306
9. PERFORMING ORGANIZATION NAME AND ADDRESS DEPT. OF STATISTICS STANFORD UNIVERSITY STANFORD, CALIFORNIA		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-373
11. CONTROLLING OFFICE NAME AND ADDRESS STATISTICS & PROBABILITY PROGRAM CODE 436 OFFICE OF NAVAL RESEARCH ARLINGTON VA 22217		12. REPORT DATE MAY 1982
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		16a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES ALSO ISSUED AS TECHNICAL REPORT NO. 189 under NATIONAL SCIENCE FOUNDATION GRANT MCS80-24649, STANFORD UNIVERSITY, DEPT. OF STATISTICS		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) STOPPING RULE, FIXED WIDTH CONFIDENCE INTERVAL, UNIFORM ASYMPTOTIC NORMALITY.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) PLEASE SEE OTHER SIDE		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-010-6001

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FIXED ACCURACY ESTIMATION OF AN AUTOREGRESSIVE PARAMETER

BY T. L. LAI and D. SIEGMUND

ABSTRACT

For a first order non-explosive autoregressive process with unknown parameter $\beta \in [-1, 1]$, it is shown that if data are collected according to a particular stopping rule, the least squares estimator of β is asymptotically normally distributed uniformly in β . In the case of normal residuals, the stopping rule may be interpreted as sampling until the observed Fisher information reaches a preassigned level. The situation is contrasted with the fixed sample size case, where the estimator has a non-normal limiting distribution when $|\beta| = 1$.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DA
FILM
0-8