The Annals of Statistics
1987, Vol. 15, No. 3, 1079–1090

SEQUENTIAL ESTIMATION OF THE MEAN OF A FIRST-ORDER STATIONARY AUTOREGRESSIVE PROCESS

BY T. N. SRIRAM

Michigan State University

This paper considers the problem of sequential point and fixed-width confidence interval estimation of the location parameter when the errors form an autoregressive process with unknown scale and autoregressive parameters. The sequential point estimator considered here is based on sample mean and 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.

1. Introduction. 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 Woodroofe (1982) and detailed references given in Chow and Martinsek (1982) and Martinsek (1983, 1984, 1985).

Two basic problems of estimation are the following. The first problem is to determine the sample size that minimizes the risk for a suitably defined loss function. The second problem is to construct a confidence interval of prescribed width and coverage probability for a parameter. In either case, the best fixed sample size procedure (BFSP), possessing the desired properties, 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 stopping rule of the type considered by Robbins (1959) for sequential estimation of the mean of a normal population in the presence of a nuisance parameter (the variance). The previously mentioned problems are then solved using sequential procedures.

The problem of estimating the mean of independent, identically distributed (i.i.d.) observations from an unknown population distribution, with the loss function equal to the squared error plus the cost per observation, has been considered by (among others) Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981). They have proposed sequential procedures using a stopping rule of the type introduced by Robbins (1959) and have shown, under certain growth rate assumptions on the initial sample size, that their procedures are "asymptotically risk efficient" (defined later) as the cost per observation tends to zero. For the same problem, Chow and Martinsek (1982) have shown that the "regret" of these procedures is bounded, as the cost of estimation error becomes infinite. Recently, Martinsek (1983) has also obtained a second-order approximation to the regret.

Received June 1985; revised November 1986.
AMS 1980 *subject classifications*. Primary 62L12; secondary 60G40, 62M10.
Key words and phrases. Asymptotic risk efficiency, asymptotic consistency, asymptotic efficiency, Burkholder inequality, Marcinkiewicz–Zygmund inequality, reverse martingale.

In all of these papers the sample mean is used to estimate the population mean and the latter two papers assume that the unit cost is chosen so that each observation costs one unit. Results of asymptotic risk efficiency have also been obtained by Sen and Ghosh (1981) for sequential estimation of estimable parametric functions using U-statistics based on i.i.d. observations.

While most of the available literature in this area deals with the situation when observations are independent, there is very little known when the observations are dependent. This paper considers the problem of sequential point and interval estimation of the location parameter μ, when the errors form a first-order autoregressive process with unknown scale and autoregressive parameters. More precisely, consider a stationary process $\{X_i;\ i \geq 0\}$ of the form

$$X_i - \mu = \sum_{j=0}^{\infty} \beta^j \varepsilon_{i-j}, \qquad |\beta| < 1, \tag{1.1}$$

where $\{\ldots, \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \ldots\}$ are i.i.d. according to some unknown distribution function F, with $E\varepsilon_1 = 0$ and $E\varepsilon_1 = \sigma^2 \in (0, \infty)$.

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

$$L_n = a\left(\overline{X}_n - \mu\right)^2 + cn, \qquad a > 0,\ c > 0, \tag{1.2}$$

where c is the cost per observation. The object is to minimize the risk in estimation by choosing an appropriate sample size. Using the model (1.1) and the independence of $\{\varepsilon_j;\ j \geq 1\}$ and X_0, it can be shown that

$$E\left(\overline{X}_n - \mu\right)^2 = n^{-1}\sigma^2/(1-\beta)^2 + o(n^{-1}), \quad \text{as } n \to \infty. \tag{1.3}$$

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

$$R_n = EL_n = n^{-1}a\sigma^2/(1-\beta)^2 + cn + o(n^{-1}) \tag{1.4}$$

is approximately minimized by the BFSP

$$n_0 \simeq c^{-1/2}a^{1/2}\sigma/(1-\beta), \tag{1.5}$$

with corresponding minimum risk

$$R_{n_0} \simeq 2cn_0. \tag{1.6}$$

However, if either σ or β is 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 (≥ 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

$$N = \inf\left\{n \geq m\colon n \geq c^{-1/2}a^{1/2}\left[\hat{\sigma}_n/|1 - \hat{\beta}_n| + n^{-h}\right]\right\}, \tag{1.7}$$

where

$$\hat{\beta}_n = \sum_{i=1}^{n-1} (X_i - \overline{X}_n)(X_{i+1} - \overline{X}_n) \Big/ \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \tag{1.8}$$

and

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

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

$$R_N = aE(\overline{X}_N - \mu)^2 + cEN. \tag{1.9}$$

It will be proved in Section 2 that this sequential procedure is "asymptotically risk efficient," i.e., $R_N/R_{n_0} \to 1$, as $c \to 0$.

Interval estimation. 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, recall from Anderson (1971), Theorem 8.4.1, that,

$$\sqrt{n}\,(\overline{X}_n - \mu) \to_{\mathscr{D}} N(0, \sigma^2/(1 - \beta)^2), \quad \text{as } n \to \infty. \tag{1.10}$$

Based on this, when σ and β are known, if one uses the confidence interval

$$I_{k_0} = [\overline{X}_{k_0} - d, \overline{X}_{k_0} + d], \tag{1.11}$$

where

$$k_0 = [d^{-2} z_{1-\alpha}^2 \sigma^2/(1 - \beta)^2] + 1$$

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 $d \to 0$,

$$k_0 \to \infty$$

and

$$\begin{aligned} P[\mu \in I_{k_0}] &= P[\sqrt{k_0}|\overline{X}_{k_0} - \mu|/\gamma \le d\sqrt{k_0}/\gamma] \\ &\to 1 - \alpha \qquad \text{(asymptotic consistency)}, \end{aligned} \tag{1.12}$$

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

$$T = \inf\{n \ge m\colon n \ge d^{-2} z_{1-\alpha}^2 [\hat{\sigma}_n^2/|1 - \hat{\beta}_n|^2 + n^{-h}]\}, \tag{1.13}$$

it will be proved in Section 2 that, as $d \to 0$,

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

and

$$E(T/k_0) \to 1 \qquad \text{(asymptotic efficiency)}.$$

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

THEOREM 1 (Risk efficiency). *Suppose* $E|\varepsilon_1|^{2p} < \infty$ *for* $p > 2$, *and that* $h \in (0, (p-2)/4)$, *where* h *is as in* (1.7). *Then as* $c \to 0$,

(1.14) $$N/n_0 \to 1 \quad a.s.,$$

(1.15) $$E|N/n_0 - 1| \to 0$$

and

(1.16) $$R_N/R_{n_0} \to 1.$$

THEOREM 2. *Under the conditions of Theorem* 1, *as* $c \to 0$,

(1.17) $$\sqrt{N}(\overline{X}_N - \mu) \to_{\mathscr{D}} N\big(0, \sigma^2/(1-\beta)^2\big).$$

THEOREM 3 (Fixed-width interval estimation). *Under the conditions of Theorem* 1, *as* $d \to 0$,

(1.18) $$T/k_0 \to 1 \quad a.s.,$$

(1.19) $$P[\mu \in I_T] \to 1 - \alpha$$

and

(1.20) $$E(T/k_0) \to 1,$$

where k_0 *is as in* (1.11).

REMARK. The stopping time T also appears in Subramanyam (1984), where (1.18) and (1.19) are proved, under somewhat different conditions. However, (1.20) is a new result.

The proofs of Theorems 1 and 2 are given in Section 2. The proof of Theorem 3 will be omitted, as the techniques used are similar to the ones that will be used for proving Theorem 1.

2. Proofs. The proof of Theorem 1 depends on a series of four lemmas, the first of which is an elementary lemma and therefore it will be stated without proof. In the first three lemmas all the limits are taken as $n \to \infty$.

LEMMA 1. *Let* Y_n, Z_n *be any sequence of random variables and* a, $b \neq 0$ *and* $s > 0$ *be real numbers. If*

$$P[|Y_n - a| > \varepsilon] = O(n^{-s}) = P[|Z_n - b| > \varepsilon], \quad \text{for every } \varepsilon > 0,$$

then

(2.1) $$P[|Y_n/Z_n - a/b| > \varepsilon] = O(n^{-s}), \quad \text{for every } \varepsilon > 0.$$

In order to state the next two lemmas, define for $k = 0, 1$

$$C_{k,n} = n^{-1} \sum_{i=1}^{n-k} (X_i - \overline{X}_n)(X_{i+k} - \overline{X}_n) \tag{2.2}$$

and

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

Note that

$$\hat{\beta}_n = C_{1,n}/C_{0,n} \quad \text{and} \quad \hat{\sigma}_n^2 = C_{0,n}(1 - \hat{\beta}_n)^2. \tag{2.3}$$

Lemma 2 gives L_p rates of convergence of $\overline{X}_n$ 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 3 deals with rates of convergence in probability of $\hat{\beta}_n$ to β and $\hat{\sigma}_n^2$ to σ^2.

LEMMA 2. *If* $E|\varepsilon_1|^{2p} < \infty$, *then*

$$\|\overline{X}_n - \mu\|_{2p} = O(n^{-1/2}), \qquad \textit{if } p \geq 1, \tag{2.4}$$

$$\|C_{0,n} - \sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}), \quad \textit{if } p \geq 2 \tag{2.5}$$

and

$$\|C_{1,n} - \beta\sigma^2/(1-\beta^2)\|_p = O(n^{-1/2}), \quad \textit{if } p \geq 2. \tag{2.6}$$

PROOF. Assume without loss of generality that $\mu = 0$. Using (1.1),

$$\begin{aligned} \|\overline{X}_n\|_{2p} &= n^{-1}\left\| \sum_{i=1}^{n} \sum_{u=0}^{\infty} \beta^u \varepsilon_{i-u} \right\|_{2p} \\ &\leq n^{-1}\left\| \sum_{u=0}^{\infty} |\beta|^u \left| \sum_{i=1}^{n} \varepsilon_{i-u} \right| \right\|_{2p} \\ &\leq n^{-1} \sum_{u=0}^{\infty} |\beta|^u \left\| \sum_{i=1}^{n} \varepsilon_{i-u} \right\|_{2p}, \end{aligned} \tag{2.7}$$

where the monotone convergence theorem and the Minkowski inequality were used to get the last inequality. Now, observe that for each fixed $u \geq 0$, the joint distributions of

$$\{\varepsilon_{1-u}, \ldots, \varepsilon_{n-u}\} \quad \text{and} \quad \{\varepsilon_1, \ldots, \varepsilon_n\}$$

are the same. By the Marcinkiewicz–Zygmund (M-Z) inequality [see Chow and Teicher (1978), Corollary 10.3.2],

$$\left\| \sum_{i=1}^{n} \varepsilon_{i-u} \right\|_{2p} = O(n^{1/2}). \tag{2.8}$$

Hence, (2.4) follows from (2.7) and (2.8). As for (2.5), use (2.2) to write

$$
\begin{aligned}
C_{0,n} - \sigma^2/(1-\beta^2) &= C^*_{0,n} - \sigma^2/(1-\beta^2) - \bar{X}_n^2 \\
&= n^{-1}\sum_{i=1}^{n}\left[\left(\sum_{u=0}^{\infty}\beta^u\varepsilon_{i-u}\right)^2 - \sigma^2/(1-\beta^2)\right] - \bar{X}_n^2 \\
&= \sum_{u=0}^{\infty}\beta^{2u}n^{-1}\sum_{i=1}^{n}\left(\varepsilon_{i-u}^2 - \sigma^2\right) \\
&\quad +2\sum_{u<v}\beta^{u+v}n^{-1}\sum_{i=1}^{n}\varepsilon_{i-u}\varepsilon_{i-v} - \bar{X}_n^2.
\end{aligned} \tag{2.9}
$$

As before

$$
\begin{aligned}
\left\|C_{0,n} - \sigma^2/(1-\beta^2)\right\|_p &\le \sum_{u=0}^{\infty}|\beta|^{2u}n^{-1}\left\|\sum_{i=1}^{n}\left(\varepsilon_{i-u}^2 - \sigma^2\right)\right\|_p \\
&\quad +2\sum_{u<v}|\beta|^{u+v}n^{-1}\left\|\sum_{i=1}^{n}\varepsilon_{i-u}\varepsilon_{i-v}\right\|_p + \|\bar{X}_n^2\|_p.
\end{aligned} \tag{2.10}
$$

Application of the M-Z inequality yields

$$
n^{-1}\left\|\sum_{i=1}^{n}\left(\varepsilon_{i-u}^2 - \sigma^2\right)\right\|_p = O(n^{-1/2}). \tag{2.11}
$$

As for the second expression on the r.h.s. of (2.10), define for each u and v fixed, where $u < v$,

$$
f_{n,u,v} = \sum_{i=1}^{n}\varepsilon_{i-u}\varepsilon_{i-v} \quad \text{and} \quad \mathscr{F}_{n-u} = \sigma\{\varepsilon_k;\ k \le n-u\}.
$$

Then, the independence of $\{\varepsilon_i\}$ and the assumption that $E\varepsilon_1 = 0$ yield that $\{f_{n,u,v}, \mathscr{F}_{n-u};\ n \ge 1\}$ is a mean zero martingale. By the Burkholder inequality [see Chow and Teicher (1978), Theorem 11.2.1], the moment inequality and the independence of ε_{i-u} and ε_{i-v}

$$
\begin{aligned}
B_p^{-1}n^{-1/2}\|f_{n,u,v}\|_p &\le \left\|\left(n^{-1}\sum_{i=1}^{n}\varepsilon_{i-u}^2\varepsilon_{i-v}^2\right)^{1/2}\right\|_p \\
&\le \left\|\left(n^{-1}\sum_{i=1}^{n}|\varepsilon_{i-u}\varepsilon_{i-v}|^p\right)^{1/p}\right\|_p \\
&= \|\varepsilon_1\|_p^2.
\end{aligned} \tag{2.12}
$$

The required result now follows from (2.10)–(2.12) and (2.4). In fact, (2.11) and (2.12) also imply that

$$
\left\|C^*_{0,n} - \sigma^2/(1-\beta^2)\right\|_p = O(n^{-1/2}). \tag{2.13}
$$

Finally, for (2.6), algebraic manipulations yield

$$C_{1,n} = C_{1,n}^* - \overline{X}_n^2 + n^{-1}\overline{X}_n\left[X_n + X_1 - \overline{X}_n\right]. \tag{2.14}$$

Now use (1.1) to write

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

Define $D_n = \sum_{i=1}^n X_{i-1}\varepsilon_i$ and $\mathscr{F}_n = \sigma\{\varepsilon_k;\ k \le n\}$. The independence of X_{i-1} and ε_i for each $i \ge 1$ and $E\varepsilon_1 = 0$ imply that $\{D_n, \mathscr{F}_n;\ n \ge 1\}$ is a mean zero martingale. By the Burkholder and the moment inequalities and the independence of ε_i and X_{i-1},

$$\begin{aligned} B_p^{-1}n^{-1/2}\|D_n\|_p &\le \left\|\left(n^{-1}\sum_{i=1}^{n}\varepsilon_i^2X_{i-1}^2\right)^{1/2}\right\|_p \\ &\le \left\|\left(n^{-1}\sum_{i=1}^{n}|X_{i-1}\varepsilon_i|^p\right)^{1/p}\right\|_p \\ &= O(1). \end{aligned} \tag{2.15}$$

Also, by the Minkowski and the Schwarz inequalities, the independence of X_n and ε_{n+1} and the stationarity of X_n,

$$\|X_0\varepsilon_1 + n^{-1}\varepsilon_{n+1}X_n - n^{-1}X_nX_{n+1}\|_p = O(n^{-1/2}). \tag{2.16}$$

Therefore, the required result follows from (2.13)–(2.16), the Minkowski and the Schwarz inequalities and (2.4). □

LEMMA 3. *If* $E|\varepsilon_1|^{2p} < \infty$, *for* $p \ge 2$, *then for every* $\varepsilon > 0$,

$$P\left[\left|\hat{\sigma}_n^2/\left(1-\hat{\beta}_n\right)^2 - \sigma^2/(1-\beta)^2\right| > \varepsilon\right] = O(n^{-p/2}). \tag{2.17}$$

PROOF. By Lemma 2 and the Markov inequality, for every $\varepsilon > 0$,

$$P\left[\left|C_{1,n} - \beta\sigma^2/(1-\beta^2)\right| > \varepsilon\right] = O(n^{-p/2}) \tag{2.18}$$

and

$$P\left[\left|C_{0,n} - \sigma^2/(1-\beta^2)\right| > \varepsilon\right] = O(n^{-p/2}). \tag{2.19}$$

From this, Lemma 1 and (2.3)

$$P\left[|\hat{\beta}_n - \beta| > \varepsilon\right] = O(n^{-p/2}). \tag{2.20}$$

This, in turn, yields

$$P\left[|\hat{\beta}_n^2 - \beta^2| > \varepsilon\right] = O(n^{-p/2}). \tag{2.21}$$

Hence, by (2.3)

$$P\left[|\hat{\sigma}_n^2 - \sigma^2| > \varepsilon\right] = O(n^{-p/2}). \tag{2.22}$$

The required result now follows from Lemma 1. □

The next lemma gives a rate on the tail behavior of the stopping rule N, which is crucial for the proof of Theorem 1.

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

$$b = (a/c)^{1/2}, \qquad n_1 = [b^{1/(1+h)}], \qquad n_2 = [n_0(1-\varepsilon)] \tag{2.23}$$

and

$$n_3 = [n_0(1+\varepsilon)], \quad \text{for } 0 < \varepsilon < 1.$$

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

LEMMA 4. *Assume the moment conditions in Theorem* 1. *Then for every* $\varepsilon > 0$, *and for* $s = p/2$

$$P[N \le n_2] = O(c^{(s-1)/2(1+h)}) \tag{2.24}$$

and

$$\sum_{n \ge n_3} P[N > n] = O(c^{(s-1)/2)}). \tag{2.25}$$

PROOF. From (1.7), $N \ge n_1$. Also

$$\begin{aligned} P[N \le n_2] &\le P\left[\hat{\sigma}_n/|1-\hat{\beta}_n| \le b^{-1}n \text{ for some } n_1 \le n \le n_2\right] \\ &\le P\left[\hat{\sigma}_n^2/(1-\hat{\beta}_n)^2 \le (1-\varepsilon)^2\sigma^2/(1-\beta)^2 \text{ for some } n_1 \le n \le n_2\right] \\ &\le \sum_{n=n_1}^{\infty} P\left[\left|\hat{\sigma}_n^2/(1-\hat{\beta}_n)^2 - \sigma^2/(1-\beta)^2\right| \ge \varepsilon(2-\varepsilon)\sigma^2/(1-\beta)^2\right]. \end{aligned} \tag{2.26}$$

Now use Lemma 3 and (2.23) to get (2.24). As for (2.25), it follows from (1.7) that, for $n \ge n_3$,

$$\begin{aligned} P[N > n] &\le P\left[\hat{\sigma}_n/|1-\hat{\beta}_n| > b^{-1}n - n^{-h}\right] \\ &\le P\left[\hat{\sigma}_n/|1-\hat{\beta}_n| - \sigma/|1-\beta| > b^{-1}(n_3 - n_0) - n_3^{-h}\right]. \end{aligned} \tag{2.27}$$

Choose c small enough so that

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

Then

$$\begin{aligned} P[N > n] &\le P\left[\hat{\sigma}_n/|1-\hat{\beta}_n| - \sigma/|1-\beta| > \sigma\varepsilon/2(1-\beta)\right] \\ &\le P\left[\left|\hat{\sigma}_n^2/|1-\hat{\beta}_n|^2 - \sigma^2/(1-\beta)^2\right| > \sigma^2\varepsilon^2/4(1-\beta)^2\right]. \end{aligned}$$

Now argue as for (2.24) to conclude (2.25). □

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

PROOF OF THEOREM 1. From (2.18), (2.19) and the Borel–Cantelli lemma

$$C_{1,n} \to \beta\sigma^2/(1-\beta^2) \quad \text{a.s.} \quad \text{and} \quad C_{0,n} \to \sigma^2/(1-\beta^2) \quad \text{a.s., as } n \to \infty.$$

From this and (2.3)

$$\hat{\beta}_n \to \beta \quad \text{a.s.} \quad \text{and} \quad \hat{\sigma}_n^2 \to \sigma^2 \quad \text{a.s., as } n \to \infty. \tag{2.28}$$

Therefore, it follows that $N < \infty$ a.s. Also, since $N \ge n_1$, $N \to \infty$. Hence, from (2.28)

$$\hat{\beta}_N \to \beta \quad \text{a.s.} \quad \text{and} \quad \hat{\sigma}_N^2 \to \sigma^2 \quad \text{a.s.} \tag{2.29}$$

But

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

and

$$N \le c^{-1/2}a^{1/2}\left[\hat{\sigma}_{N-1}/|1-\hat{\beta}_{N-1}| + (N-1)^{-h}\right] + n_1.$$

Therefore, from (2.29)

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

As for the L_1-convergence (1.15), write

$$N/n_0 - 1 = (N/n_0)I_B + (N/n_0 - 1)I_A + (N/n_0)I_D - I_{B\cup D},$$

so that

$$E|N/n_0 - 1| \le (2-\varepsilon)P[N \le n_2] + \varepsilon + n_0^{-1}\sum_{n \ge n_3} P[N > n] + P(N > n_3).$$

The required result follows from Lemma 4, since ε is arbitrary.

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

$$R_N/R_{n_0} \simeq \{aE\bar{X}_N^2 + cEN\}/2cn_0.$$

From (1.15), it suffices to show that

$$aE\bar{X}_N^2/cn_0 \to 1. \tag{2.30}$$

To establish (2.30) it suffices to show that

$$aE\bar{X}_N^2 I_{\bar{A}}/cn_0 \to 0 \tag{2.31}$$

and

$$aE\left(\bar{X}_N - \bar{X}_{n_0}\right)^2 I_A/cn_0 \to 0. \tag{2.32}$$

Consider (2.31). From (1.1)

$$
\begin{aligned}
E\overline{X}_N^2 I_B &\le E \max_{n_1 \le n \le n_2} \overline{X}_n^2 I_B \\
&\le E\left| \sum_{u=0}^{\infty} |\beta|^u \max_{n_1 \le n \le n_2} n^{-1} \sum_{i=1}^{n} \varepsilon_{i-u} \right|^2 I_B \qquad (2.33)\\
&\le \sum_{u=0}^{\infty} |\beta|^{2u} E M_{n_1, u}^2 I_B + 2 \sum_{u<v} |\beta|^{u+v} E M_{n_1, u} M_{n_1, v} I_B,
\end{aligned}
$$

where $M_{n_1, u} = \max_{n_1 \le n \le n_2} |n^{-1} \sum_{i=1}^{n} \varepsilon_{i-u}|$.

Now, observe that for each fixed $u \ge 0$, $\{n^{-1}\sum_{i=1}^{n}\varepsilon_{i-u},\ n \ge 1\}$ is a reverse martingale. Therefore, by this fact, the Schwarz inequality, the maximal inequality for reverse submartingales [see Sen (1982), (2.2.6), page 13], (2.8) and (2.24)

$$
\begin{aligned}
EM_{n_1, u}^2 I_B &\le E^{1/2} M_{n_1, u}^4 \{P(B)\}^{1/2} \\
&\le \left(\tfrac{16}{9}\right) E^{1/2} \left| n_1^{-1} \sum_{i=1}^{n_1} \varepsilon_{i-u} \right|^4 \{P(B)\}^{1/2} \qquad (2.34)\\
&= O\left(n_1^{-1}\right)\left\{O\left(c^{(p-2)/8(1+h)}\right)\right\}.
\end{aligned}
$$

Since $h < [(p-2)/4]$, we have

$$EM_{n_1, u}^2 I_B / c n_0 \to 0.$$

Therefore, $|\beta| < 1$ and the Schwarz inequality yield

$$E\overline{X}_N^2 I_B / c n_0 \to 0. \qquad (2.35)$$

Similar arguments yield

$$E\overline{X}_N^2 I_D / c n_0 \to 0. \qquad (2.36)$$

This proves (2.31). As for (2.32), it follows from (1.1) that

$$
\begin{aligned}
E\left(\overline{X}_N - \overline{X}_{n_0}\right)^2 I_A &\le E \max_{n_2 \le n \le n_3} |\overline{X}_n - \overline{X}_{n_0}|^2 I_A \\
&\le E\left\{ \sum_{u=0}^{\infty} |\beta|^u \max_{n_2 \le n \le n_3} \left| n^{-1} \sum_{i=1}^{n} \varepsilon_{i-u} - n_0^{-1} \sum_{i=1}^{n_0} \varepsilon_{i-u} \right| \right\}^2 \qquad (2.37)\\
&= \sum_{u=0}^{\infty} |\beta|^{2u} E \max_{n_2 \le n \le n_3} W_{n, n_0, u}^2 \\
&\quad + 2 \sum_{u<v} |\beta|^{u+v} E \max_{n_2 \le n \le n_3} W_{n, n_0, u} \max_{n_2 \le n \le n_3} W_{n, n_0, v},
\end{aligned}
$$

where $W_{n, n_0, u} = n^{-1}\sum_{i=1}^{n}\varepsilon_{i-u} - n_0^{-1}\sum_{i=1}^{n_0}\varepsilon_{i-u}$.

Observe that for each fixed $u \geq 0$, $\{W_{n, n_3, u};\ n_0 \leq n \leq n_3\}$ is a reserve martingale. By the maximal inequality for reverse submartingales

$$E \max_{n_0 \leq n \leq n_3} W^2_{n, n_0, u} \leq 2E \max_{n_0 \leq n \leq n_3} W^2_{n, n_3, u}$$

$$(2.38) \qquad + 2E\left| n_3^{-1} \sum_{i=1}^{n_3} \varepsilon_{i-u} - n_0^{-1} \sum_{i=0}^{n_0} \varepsilon_{i-u} \right|^2$$

$$\leq 10E\left| n_3^{-1} \sum_{i=1}^{n_3} \varepsilon_{i-u} - n_0^{-1} \sum_{i=1}^{n_0} \varepsilon_{i-u} \right|^2$$

$$= 10\varepsilon/n_3.$$

Since ε is arbitrary,

$$(2.39) \qquad E \max_{n_0 \leq n \leq n_3} W^2_{n, n_0, u}/cn_0 \to 0.$$

Similar arguments yield

$$(2.40) \qquad E \max_{n_2 \leq n \leq n_0} W^2_{n, n_0, u}/cn_0 \to 0.$$

The required result now follows from the Schwarz inequality. Hence the theorem. □

PROOF OF THEOREM 2. Write

$$\sqrt{N}\left(\overline{X}_N - \mu\right) = \sqrt{N/n_0}\sqrt{n_0}\left(\overline{X}_n - \overline{X}_{n_0}\right) + \sqrt{N/n_0}\sqrt{n_0}\left(\overline{X}_{n_0} - \mu\right).$$

By (2.31), (2.32) and since $cn_0^2 \simeq a\sigma^2/(1-\beta)^2$

$$(2.41) \qquad n_0 E\left(\overline{X}_N - \overline{X}_{n_0}\right)^2 \to 0.$$

Therefore,

$$(2.42) \qquad \sqrt{n_0}|\overline{X}_N - \overline{X}_{n_0}| \to_P 0.$$

Hence, Theorem 2 follows from (1.10), (1.14) and the Slutsky theorem. □

Acknowledgments. This paper is a part of my Ph.D. dissertation which was written at Michigan State University under the guidance of Professor H. L. Koul. I wish to thank Professor Koul for suggesting the problem, for his careful reading and constructive remarks on the paper which led to this improved version. I also wish to thank the referee and the Associate Editor for their valuable comments.

REFERENCES

ANDERSON, T. W. (1959). On asymptotic distributions 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.

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 sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*. Springer, 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.

GHOSH, M. and MUKHOPADHYAY, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Comm. Statist. A.—Theory Methods* **8** 637–651.

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. *Z. Wahrsch. verw. Gebiete* **70** 411–416.

ROBBINS, H. E. (1959). Sequential estimation of the mean of a normal population. In *Probability and Statistics* (U. Grenander, ed.) 235–245. Wiley, New York.

SEN, P. K. (1982). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. Wiley, New York.

SEN, P. K. and GHOSH, M. (1981). Sequential point estimation of estimable parameters based on *U*-statistics. *Sankhyā Ser. A* **43** 331–344.

SUBRAMANYAM, A. (1984). Contributions to the asymptotic theory of sequential procedures. Ph.D. dissertation, Univ. of Poona.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*. SIAM, Philadelphia.

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602