Reprinted from THEORETICAL POPULATION BIOLOGY
All Rights Reserved by Academic Press, New York and London

Vol. 27, No. 2, April 1985
Printed in Belgium

122

Confidence Intervals for Demographic Projections Based on Products of Random Matrices

C. C. HEYDE

CSIRO Division of Mathematics and Statistics, Canberra, Australia

AND

JOEL E. COHEN

Rockefeller University, 1230 York Avenue, New York 10021

Received September 13, 1983

This work is concerned with the growth of age-structured populations whose vital rates vary stochastically in time and with the provision of confidence intervals. In this paper a model

$$Y_{t+1}(\omega) = X_{t+1}(\omega)\, Y_t(\omega)$$

is considered, where Y_t is the (column) vector of the numbers of individuals in each age class at time t, X is a matrix of vital rates, and ω refers to a particular realization of the process that produces the vital rates. It is assumed that $\{X_i\}$ is a stationary sequence of random matrices with nonnegative elements and that there is an integer n_0 such that any product $X_{j+n_0} \cdots X_{j+1} X_j$ has all its elements positive with probability one. Then, under mild additional conditions, strong laws of large numbers and central limit results are obtained for the logarithms of the components of Y_t. Large-sample estimators of the parameters in these limit results are derived. From these, confidence intervals on population growth and growth rates can be constructed. Various finite-sample estimators are studied numerically. The estimators are then used to study the growth of the striped bass population breeding in the Potomac River of the eastern United States. © 1985 Academic Press, Inc.

1. INTRODUCTION

To describe an age-structured population whose vital rates vary stochastically in time, let Y_t be the (column) vector consisting of the numbers of individuals in each of K age classes at time t. Assume that

$$Y_{t+1}(\omega) = X_{t+1}(\omega)\, Y_t(\omega), \tag{1}$$

0040-5809/85 $3.00

where X is a matrix of vital rates and ω refers to a particular realization of the process that produces the vital rates.

The model (1) emphasizes the effects of variations in the vital rates themselves. The formulation suppresses the variability of births and deaths conditional on given vital rates.

A simple example of the use of (1) would involve confining attention to the females in the population and then taking the Xs as Leslie matrices. These have the age specific fertility rates in the first row, the age specific survival rates in the $(i+1, i)$ elements, $i = 1, 2, \ldots, K-1$, and all other elements zero.

The model (1), under various conditions on the Xs, has been much discussed, particularly with a view to establishing asymptotic results on the long-term population growth rate. Tuljapurkar and Orzack (1980) give a comprehensive listing of earlier work. We shall treat (1) in greater generality and, for the first time, provide confidence intervals and a hypothesis testing framework for population growth rates and total population.

A convenient approach to studying the long-term behaviour of (1) is via the asymptotic theory of subadditive processes. Let $X_1, X_2, \ldots$, be a stationary ergodic random sequence of $K \times K$ matrices with nonnegative elements and write M_{ij} for the i, j element of a matrix M. Suppose that $E|\log \max(X_1)_{ij}| < \infty$ and also that there exists an integer n_0 such that any product $X_{j+n_0} \cdots X_{j+1} X_j$ has all its elements positive with probability one. Then $(X_{n_0 t} \cdots X_{n_0 s+1})_{11} > 0$ and

$$\{x_{st} = -\log(X_{n_0 t} \cdots X_{n_0 s+1})_{11}, s < t\}$$

is a subadditive process and the Kingman ergodic theorem for subadditive processes (e.g., Hall and Heyde, 1980, Theorem 7.5, p. 215) leads to

$$t^{-1} \log(X_t \cdots X_1)_{11} \xrightarrow{\text{a.s.}} \log \lambda$$

(say), which is a finite constant, as $t \to \infty$. This results continues to hold for all matrix entries i, j and, under mild additional conditions, can be augmented by a central limit theorem

$$\sigma^{-1} t^{-1/2} \{\log(X_t \cdots X_1)_{ij} - t \log \lambda\} \xrightarrow{d} N(0, 1)$$

($\xrightarrow{d}$ denoting convergence in distribution and $N(0, 1)$ the unit normal law) for some $\sigma > 0$ and all $1 \leqslant i, j \leqslant K$. The results extend those of Furstenberg and Kesten (1960), Ishitani (1977) and Tuljapurkar and Orzack (1980) and can readily be used on the model (1). In particular, it can be shown that the ergodic theorem gives for the total population size $1'Y_t$,

$$t^{-1} \log 1'Y_t \xrightarrow{\text{a.s.}} \log \lambda \; (= \lim_{t \to \infty} t^{-1} E \log 1'Y_t) \tag{2}$$

and the central limit theorem gives

$$\sigma^{-1} t^{-1/2} \{\log 1'Y_t - t \log \lambda\} \xrightarrow{d} N(0, 1). \tag{3}$$

This last result (in mixing form) provides a basis for the construction of confidence intervals for population projections.

In Parts 2 and 3 of the paper we shall establish the above results and indicate how $\log \lambda$ and σ can be estimated in practice. An application to the striped bass population of the Potomac River is given in Part 5 to illustrate the usefulness of the methodology.

The model (1) is closely related to the one which can be obtained from a formulation of the population process as a multitype Galton–Watson branching process with random environments. In this case the different types would correspond to the different age classes. The offspring probability generating functions

$$\{\phi_{\zeta_n}(s) = (\phi_{1,\zeta_n}(s), \ldots, \phi_{K,\zeta_n}(s)), n = 0, 1, 2, \ldots\}$$

are chosen from a collection Φ of probability generating functions according to a sequence of environmental variables ζ_n, $n = 0, 1, 2, \ldots$, which are assumed to stationary and ergodic (or perhaps even independent and identically distributed). This process evolves in such a way that the conditional generating function of $Y'_{t+1} = (Y_1(t+1), \ldots, Y_K(t+1))$ given Y_t is

$$E\left[\prod_{j=1}^{K} \phi_{j,\zeta_t}^{Y_j(t)}(s) \mid Y_t\right].$$

Tanny (1981) has studied the growth of this process and obtained various results analogous to (2). Results like (3), on which confidence intervals and hypothesis tests must be based, are not available except in the 1-type case or the case of constant environment.

2. Theoretical Results

The evolution of the system (1) depends crucially on properties of products of random matrices since iteration of (1) gives

$$Y_t = X_t X_{t-1} \cdots X_1 Y_0. \tag{4}$$

We begin by specifying the kinds of matrices with which we shall deal. Let $X_1, X_2, \ldots$, be a stationary ergodic random sequence of $K \times K$ matrices with nonnegative elements. For a matrix M denote by M_{ij} the element in the ith row and jth column. We shall suppose that the matrices $\{X_i\}$ satisfy two assumptions:

(A1) There exists an integer n_0 such that any product $X_{j+n_0} \cdots X_{j+1}$ of n_0 of the matrices has all its elements positive with probability one.

(A2) For some constant C, $1 < C < \infty$ and each matrix X_i,

$$1 \leqslant M(X_i)/m(X_i) \leqslant C$$

with probability one, where $M(X)$ and $m(X)$ are, respectively, the maximum and minimum *positive* elements of X.

Assumption (A1) is standard within the theory of nonnegative matrices. It ensures that the effects of the initial composition of the population disappear in the limit (demographic weak ergodicity). Both assumptions define an "ergodic set" in the sense of Hajnal (1976). For products of $K \times K$ Leslie matrices with no zero vital rates $n_0 = K$ and, in general, if n_0 exists then $n_0 \leqslant 2^K - 2$ (Cohen and Sellers, 1982), the bound being sharp.

Assumption (A2) is perhaps less familiar than the more restrictive version in which

$$0 < \beta \leqslant m(X_i) \leqslant M(X_i) \leqslant \gamma < \infty$$

for each matrix X_i and some β, γ used by Tuljapurkar and Orzack (1980); (A2) allows, for example, cases such as $X_i = \alpha_i X$, where the $\alpha_i > 0$ are stationary (and unrestricted) while X is fixed.

Next suppose that the process $X_1, X_2, \ldots$, is defined on a probability space $(\Omega, \mathscr{F}, P)$. Write $\mathscr{M}_a^b$ for the σ-field generated by $X_a, \ldots, X_b$ and let

$$\phi(n) = \sup_{k>0} \{|P(B \mid A) - P(B)|; A \in \mathscr{M}_0^k, B \in \mathscr{M}_{k+n}^{\infty}, P(A) > 0\}.$$

The condition $\phi(n) \to 0$ as $n \to \infty$ (called uniform mixing) is one of a number of standard conditions of asymptotic independence (see, e.g., Hall and Heyde, 1980, Chap. 5).

The basis for the inferential results of the paper is provided by the following theorem.

THEOREM 1. *Suppose that the stationary ergodic sequence of matrices* $\{X_i\}$ *is such that assumptions* (A1) *and* (A2) *are satisfied and* $E|\log M(X_1)| < \infty$. *Then, for all* $1 \leqslant i, j \leqslant K$,

$$t^{-1} \log(X_t \cdots X_1)_{ij} \xrightarrow{\text{a.s.}} \log \lambda \tag{5}$$

(say), which is a finite constant, as $t \to \infty$. *Furthermore, if*

$$E|\log M(X_1)|^2 < \infty \tag{6}$$

and

$$\sum_{n=1}^{\infty} |\phi(n)|^{1/2} < \infty \tag{7}$$

then

$$\lim_{t\to\infty} t^{-1/2} E|\log(X_t \cdots X_1)_{ij} - t \log \lambda| = \sigma(2/\pi)^{1/2}$$

exists for $0 \leqslant \sigma < \infty$ *and if* $\sigma > 0$, *then*

$$(t\sigma^2)^{-1/2}\{\log(X_t \cdots X_1)_{ij} - t \log \lambda\} \xrightarrow{d} N(0, 1) \text{ (mixing)}, \tag{8}$$

as $t \to \infty$, *the convergence being mixing (in the sense of Rényi).*

The conditions imposed to obtain the central limit part of this theorem are just indicative of the possibilities. Many variants on uniform mixing and (7) could equally be proposed. The present choice is merely a convenient one for dealing with the envisaged applications. Condition (7) easily covers the practically important cases: (i) the matrices $\{X_i\}$ are m-dependent, meaning that for $i < k < k+n < j$ the sets $(X_i, \ldots, X_k)$ and $(X_{k+n}, \ldots, X_j)$ are independent whenever $n > m$ (in this terminology an independent sequence is 0-dependent), and (ii) the matrices $\{X_i\}$ are from a finite set and are determined by an irreducible aperiodic Markov chain of fixed finite order.

The mixing convergence in the sense of Rènyi means that

$$P((t\sigma^2)^{-1/2}\{\log(X_t \cdots X_1)_{ij} - t \log \lambda\} \leqslant x \mid E) \to \Phi(x)$$

as $t \to \infty$ for any $-\infty < x < \infty$, where $\Phi(x)$ is the distribution function of the unit normal law, and for any $E \in \mathscr{F}$ with $P(E) > 0$.

Theorem 1 extends the corresponding result of Tuljapurkar and Orzack (1980) which deals with the case where $\{X_i\}$ form a Markov chain and $\phi(n)$ decreases geometrically to zero. That the condition (A3) of Tuljapurkar and Orzack (1980) ensures uniform mixing with a geometric rate follows from Rosenblatt (1971, pp. 209–213). Furthermore, the mixing convergence in Theorem 1 allows the probability measure based on the stationary initial distribution to be replaced by any probability measure which is absolutely continuous with respect to it without perturbing the limit distribution.

Now we apply the results of Theorem 1 to the system (1).

THEOREM 2. *Let* $Y_{t+1} = X_{t+1} Y_t$, $t \geqslant 0$, *and let* $Z_t = a' Y_t$ *where* a *is a nonzero vector of nonnegative elements. Then, under the same conditions as* (5) *of Theorem* 1,

$$\lim_{t\to\infty} t^{-1} \log Z_t = \log \lambda \text{ a.s.}$$

and, if the additional conditions (6) *and* (7) *of Theorem* 1 *hold,*

$$\lim_{t\to\infty} t^{-1/2} E\,|\log Z_t - t\log\lambda| = \sigma(2/\pi)^{1/2}$$

exists for $0 \leqslant \sigma < \infty$ *and if* $\sigma > 0$,

$$(t\sigma^2)^{-1/2}\{\log Z_t - t\log\lambda\} \xrightarrow{d} N(0, 1) \text{ (mixing)}$$

as $t\to\infty$, *the convergence being mixing in the sense of Rényi.*

A wide variety of other combinations of the Z_j, $1 \leqslant j \leqslant t$, also provide strongly consistent estimators of $\log\lambda$. In particular, let $\{a_j^{(t)}, 1 \leqslant j \leqslant t, t \geqslant 1\}$ be a set of nonnegative weights such that for all t, $a_t = \sum_{j=1}^t j a_j^{(t)} > 0$, and define the statistics

$$\log\lambda_t = a_t^{-1} \sum_{j=1}^t a_j^{(t)} \log Z_j.$$

If $a_t^{-1} \max_{1<j<t} a_j^{(t)} \to 0$ then under the same assumptions as (5) it is easily checked that

$$\log\lambda_t \xrightarrow{\text{a.s.}} \log\lambda$$

as $t\to\infty$. However, the estimator $t^{-1}\log Z_t$ has minimum asymptotic variance within a broad class of these combination estimators.

Indeed, if in addition (6) holds and

$$\left(\sum_{j=1}^t j^{1/2} a_j^{(t)}\right)^{-1} A_1^{(t)} \to 0,\ \left(\sum_{j=1}^t j^{1/2} a_j^{(t)}\right)^2 = O\left(\sum_{j=1}^t (A_j^{(t)})^2\right)$$

as $t\to\infty$, where $A_j^{(t)} = \sum_{i=j}^t a_i^{(t)}$, $1 \leqslant j \leqslant t$, then if $\sigma > 0$ in Theorem 2,

$$\operatorname{var}(\log\lambda_t - \log\lambda) \sim \sigma_t^2$$

as $t\to\infty$, where

$$\sigma_t^2 = (\sigma^2/a_t^2) \sum_{j=1}^t (A_j^{(t)})^2 \geqslant \sigma^2/t.$$

The minimum asymptotic variance is achieved when $a_t^{(t)} = 1$ and $a_j^{(t)} = 0$, $j < t$; i.e., for the estimator $t^{-1}\log Z_t$. We shall omit the (rather involved) proof since the result is of a supplementary nature. A result of similar character holds for the estimation of the criticality parameter of a supercritical branching process with random environments (Pakes and Heyde, 1982).

The finite-sample variance of consistent estimators $\log\lambda_t$ with vario of weights will be investigated numerically below (in Sect. 4) in an e provided by a real biological use of these methods.

Theorem 2 gives a means by which approximate confidence interv $\log\lambda$ can be constructed if $\sigma(>0)$ can be estimated. We shall $t^{-1}\log Z_t$ to estimate $\log\lambda$ while a consistent estimator for σ is giver next theorem.

THEOREM 3. *Under the conditions* (5), (6), *and* (7) *we have*

$$(\log t)^{-1} \sum_{i=1}^t |\log Z_i - i\log\lambda|\, i^{-3/2} \xrightarrow{p} \sigma(2/\pi)^{1/2}$$

as $t\to\infty$ ($\xrightarrow{p}$ *denoting convergence in probability*) *and*

$$(\log t)^{-1} \sum_{i=1}^t |\log Z_i - i\log\hat\lambda|\, i^{-3/2} \xrightarrow{p} \sigma(2/\pi)^{1/2}$$

as $t\to\infty$, *where* $\log\hat\lambda = t^{-1}\log Z_t$.

For inferential purposes we would ordinarily use the case where Z total population size $1'Y_t$. Then if $\hat\sigma$ is the estimate obtained via The for σ from a realization $Y_{t_0},\dots, Y_t$, where t is large compared with *approximate* $100(1-\alpha)\%$ *confidence interval for the growth rate* log

$$t^{-1}\log 1'Y_t \pm z_{\alpha/2}\hat\sigma t^{-1/2},$$

where

$$\beta = (2\pi)^{-1/2}\int_{z_\beta}^\infty e^{-(1/2)u^2}\,du = 1 - \Phi(z_\beta).$$

Furthermore, the mixing convergence in Theorem 2 can be used to approximate confidence intervals for the logarithm of the population s later time $\tau > t$. We have, writing

$$W_u = (u\sigma^2)^{-1/2}\{\log 1'Y_u - u\log\lambda\},$$

that

$$P(|W_\tau| < z_{p/2} \,|\, |W_t| < z_{q/2}) \to 1 - p$$

as $\tau\to\infty$ for fixed t and hence for large t,

$$P(|W_\tau| < z_{p/2}, |W_t| < z_{q/2}) \approx (1-p)(1-q).$$

But

$$P(\tau t^{-1}(\log 1'Y_t - (t\sigma^2)^{1/2} z_{q/2}) - (\tau\sigma^2)^{1/2} z_{p/2} < \log 1'Y_\tau$$
$$< \tau t^{-1}(\log 1'Y_t + (t\sigma^2)^{1/2} z_{q/2}) + (\tau\sigma^2)^{1/2} z_{p/2})$$
$$\geqslant P(|W_\tau| < z_{p/2}, |W_t| < z_{q/2}).$$

Thus, *an approximate* $100(1-\alpha)\%$ *confidence interval for the logarithm of the population size* $\log 1'Y_\tau$ *is*

$$\tau t^{-1} \log 1'Y_t \pm \hat{\sigma} \min_{\alpha > q > 0} (\tau t^{-1/2} z_{q/2} + \tau^{1/2} z_{(\alpha - q)/2(1-q)}). \qquad (10)$$

The minimization in (10) is not straightforward in general but approximate results can easily be obtained numerically in particular cases.

In practice the actual generation numbers $t_0, \ldots, t$ will frequently be unknown so we must work with differences from t_0. Using Theorem 2,

$$(t - t_0)^{-1}(\log 1'Y_t - \log 1'Y_{t_0}) = \log \hat{\lambda} \qquad (11)$$

is a strongly consistent estimator of $\log \lambda$ while to estimate σ we can employ

$$(\pi/2)^{+1/2}(\log(t-t_0))^{-1} \sum_{i=1}^{t-t_0} i^{-3/2} |\log 1'Y_{t_0+i} - \log 1'Y_{t_0} - i \log \hat{\lambda}| \qquad (12)$$

which is consistent for σ in view of Theorem 3. Then in (9) and all but the first term of (10) we replace t by $t - t_0$, $\log 1'Y_t$ by $\log 1'Y_t - \log 1'Y_{t_0}$, and (in (10)) τ by $\tau - t$ to obtain approximate $100(1-\alpha)\%$ confidence intervals for $\log \lambda$ and $\log 1'Y_\tau$, respectively. The theory, of course, has the great advantage that no specific information is required about the matrices X_i of vital rates. However, such information *is* required to confirm the relevance of the model (1).

3. Proofs

Proof of Theorem 1. In the case of random matrices X_i all of whose entries are positive with probability one, the result of Theorem 1 is known save for the mixing property of the convergence to normality (8). For the strong law (2) see the corollary of Furstenberg and Kesten (1960) while the remaining results involve a minor modification of Theorem 2 of Ishitani (1977) (see Theorem 7.7 and the discussion in pp. 225, 226 of Hall and Heyde, 1980). We shall indicate the changes that are necessary in the proofs of these results to cope with the present context and to establish mixing convergence in (8).

Fundamental to the proof of the theorem are replacements dealing with matrices which may have zero entries for Lemmas 2 and 3 of Furstenberg and Kesten (1960) and these we shall denote below by Lemmas 2′ and 3′. We write

$${}^tY^s = X_t \cdots X_s$$

while M_{ij} denotes the element in the ith row and jth column of M, and

$$M_{i.} = \sum_{j=1}^K M_{ij}, \qquad M_{.j} = \sum_{i=1}^K M_{ij}, \qquad M_{..} = \sum_{i=1}^K \sum_{j=1}^K M_{ij},$$
$$\|M\| = \max_i M_{i.}.$$

Lemma 2′. *If* (A1) *and* (A2) *are satisfied,*

$$({}^{n+m}Y^m)_{ij} > 0 \qquad (13)$$

for $n \geqslant n_0$ *and all* $1 \leqslant i, j \leqslant K$ *and*

$$1 \leqslant M({}^{n+m}Y^m)/m({}^{n+m}Y^m) \leqslant (KC)^{2n_0} \qquad (14)$$

for $n \geqslant n_0$. *Also, for all* n,

$$1 \leqslant M({}^{n+m}Y^m)/m({}^{n+m}Y^m) \leqslant (KC)^n. \qquad (15)$$

Proof. The result (13) follows by assumption. To prove (15) we first note that for any i, j for which it is positive, $({}^{n+m}Y^m)_{ij}$ is a sum of $(n+1)$-tuples numbering between 1 and K^n. Then, for $({}^{n+m}Y^m)_{ij} > 0$,

$$K^n M(X_{n+m}) \cdots M(X_m) \geqslant ({}^{n+m}Y^m)_{ij}$$
$$\geqslant m(X_{n+m}) \cdots m(X_m)$$
$$\geqslant C^{-n} M(X_{n+m}) \cdots M(X_m),$$

using (A2) and (15) follows. This also gives (14) for $n \leqslant 2n_0$.

To obtain (14) for $n \geqslant 2n_0 + 2$ we note that

$$({}^{n+m}Y^m)_{ij} = \sum_{r,s} ({}^{n+m}Y^{n-n_0+m})_{ir} ({}^{n-n_0-1+m}Y^{n_0+1+m})_{rs} ({}^{n_0+m}Y^m)_{sj}$$
$$\leqslant (KC)^{2n_0} \sum_{r,s} m({}^{n+m}Y^{n-n_0+m}) ({}^{n-n_0-1+m}Y^{n_0+1+m})_{rs} m({}^{n_0+m}Y^m)$$
$$\leqslant (KC)^{2n_0} m({}^{n+m}Y^m)$$

from two applications of the special case $n = n_0$ of (15).

Finally, for $n = 2n_0 + 1$ we have

$$\frac{({}^{2n_0+1+m}Y^m)_{i_1 j_1}}{({}^{2n_0+1+m}Y^m)_{i_2 j_2}} = \frac{\sum_r ({}^{2n_0+1+m}Y^{n_0+1+m})_{i_1 r}({}^{n_0+m}Y^m)_{rj_1}}{\sum_r ({}^{2n_0+1+m}Y^{n_0+1+m})_{i_2 r}({}^{n_0+m}Y^m)_{rj_2}}$$

$$\leqslant (KC)^{n_0}\frac{({}^{2n_0+1+m}Y^{n_0+1+m})_{i_1 \cdot}}{({}^{2n_0+1+m}Y^{n_0+1+m})_{i_2 \cdot}}$$

$$\leqslant (KC)^{n_0}\frac{M({}^{2n_0+1+m}Y^{n_0+1+m})}{m({}^{2n_0+1+m}Y^{n_0+1+m})}$$

$$\leqslant (KC)^{2n_0},$$

and similarly the lower bound is $(KC)^{-2n_0}$. This completes the proof of Lemma 2′.

LEMMA 3′. *For $n \geqslant 2n_0$,*

$$\left| \frac{({}^{n+m}Y^m)_{i_1 j}}{({}^{n+m}Y^m)_{i_1 \cdot}} - \frac{({}^{n+m}Y^m)_{i_2 j}}{({}^{n+m}Y^m)_{i_2 \cdot}} \right| \leqslant R_n,$$

where $R_n = (1-(KC)^{-4n_0})^{[n/n_0]-1}$, $[x]$ denoting the integer part of x.

Proof. For $m_1 \geqslant m$, $r \geqslant 1$, we have upon writing

$$A_{ij} = ({}^{(r+1)n_0+m_1}Y^{rn_0+1+m_1})_{ij}, \qquad B_{ij}(r) = ({}^{rn_0+m_1}Y^m)_{ij}$$

to simplify the expressions,

$$\frac{({}^{(r+1)n_0+m_1}Y^m)_{i_1 j}}{({}^{(r+1)n_0+m_1}Y^m)_{i_1 \cdot}} = \frac{B_{i_1 j}(r+1)}{B_{i_1 \cdot}(r+1)}$$

$$= \sum_s \frac{A_{i_1 s}B_{s\cdot}(r)}{B_{i_1 \cdot}(r+1)}\frac{B_{sj}(r)}{B_{s\cdot}(r)}. \tag{16}$$

But, for all i,

$$\sum_s \frac{A_{is}B_{s\cdot}(r)}{B_{i\cdot}(r+1)} = 1 \tag{17}$$

and all the summands in (17) are positive so that, using Lemma 2′,

$$\frac{A_{i_2 s}B_{s\cdot}(r)}{B_{i_2\cdot}(r+1)} \geqslant (KC)^{-4n_0}\frac{A_{i_1 s}B_{s\cdot}(r)}{B_{i_1\cdot}(r+1)}. \tag{18}$$

We have using (16) that

$$\frac{B_{i_1 j}(r+1)}{B_{i_1\cdot}(r+1)} - \frac{B_{i_2 j}(r+1)}{B_{i_2\cdot}(r+1)}$$

$$= \sum_s \left\{ \frac{A_{i_1 s}B_{s\cdot}(r)}{B_{i_1\cdot}(r+1)} - \frac{A_{i_2 s}B_{s\cdot}(r)}{B_{i_2\cdot}(r+1)} \right\} \frac{B_{sj}(r)}{B_{s\cdot}(r)}, \tag{19}$$

and if S' and S'' are, respectively, the sets of indices S for which the summand on the right-hand side of (19) is positive or negative,

$$0 \leqslant \sum_{s\in S'} \left\{ \frac{A_{i_1 s}}{B_{i_1\cdot}(r+1)} - \frac{A_{i_2 s}}{B_{i_2\cdot}(r+1)} \right\} B_{s\cdot}(r)$$

$$= -\sum_{s\in S''} \left\{ \frac{A_{i_1 s}}{B_{i_1\cdot}(r+1)} - \frac{A_{i_2 s}}{B_{i_2\cdot}(r+1)} \right\} B_{s\cdot}(r) \leqslant 1-(KC)^{-4n_0} \tag{20}$$

using (17) and (18). Consequently, using (20) in (19) and iterating,

$$\max_i \frac{B_{ij}(r+1)}{B_{i\cdot}(r+1)} - \min_i \frac{B_{ij}(r+1)}{B_{i\cdot}(r+1)}$$

$$\leqslant (1-(KC)^{-4n_0}) \left\{ \max_i \frac{B_{ij}(r)}{B_{i\cdot}(r)} - \min_i \frac{B_{ij}(r)}{B_{i\cdot}(r)} \right\}$$

$$\leqslant (1-(KC)^{-4n_0})^r \left\{ \max_i \frac{B_{ij}(1)}{B_{i\cdot}(1)} - \min_i \frac{B_{ij}(1)}{B_{i\cdot}(1)} \right\}$$

$$\leqslant (1-(KC)^{-4n_0})^r$$

from which the result of the lemma follows.

Now we resume the proof of Theorem 1. In order to establish (5) we begin by noting that

$$\{x_{st} = -\log({}^{n_0 t}Y^{n_0 s+1})_{11}, s < t\}$$

is a subadditive process. Subadditivity follows since for $s < u < t$,

$$({}^{n_0 t}Y^{n_0 s+1})_{11} = \sum_{j=1}^{K} ({}^{n_0 t}Y^{n_0 u+1})_{1j}({}^{n_0 u}Y^{n_0 s+1})_{j1}$$

$$\geqslant ({}^{n_0 t}Y^{n_0 u+1})_{11}({}^{n_0 u}Y^{n_0 s+1})_{11}$$

and, upon taking logarithms,

$$x_{st} \leqslant x_{su} + x_{ut}.$$

Furthermore,

$$-Ex_{0t} = E\log({}^{n_0t}Y^1)_{11} \leqslant E\log\|{}^{n_0t}Y^1\| \leqslant n_0 tE\log\|X_1\|$$

so that

$$\inf_{n>0} n^{-1}Ex_{0n} \geqslant -n_0 E\log\|X_1\| > -\infty,$$

since $E|\log M(X_1)| < \infty$. Consequently, the Kingman ergodic theorem for subadditive processes (e.g., Theorem 7.5, p. 215 of Hall and Heyde, 1980) gives

$$\lim_{t\to\infty} t^{-1}x_{0t} = -n_0\log\lambda \qquad \text{a.s.} \tag{21}$$

(say) which is finite and a constant since the stationary process $\{X_i\}$ is ergodic. Then, to obtain (5) for $i=j=1$ we first note that for $0 \leqslant k \leqslant n_0$,

$$\begin{aligned}(KC)^{2n_0}({}^{n_0t}Y^1)_{11}\prod_{j=n_0t+1}^{n_0t+k} M(X_j) &\geqslant ({}^{n_0t+k}Y^1)_{11} \\ &= \sum_{j=1}^{K} ({}^{n_0t+k}Y^{n_0t+1})_{1j}({}^{n_0t}Y^1)_{j1} \\ &\geqslant ({}^{n_0t}Y^1)_{11}\prod_{j=n_0t+1}^{n_0t+k} m(X_j) \qquad (22) \\ &\geqslant C^{-k}({}^{n_0t}Y^1)_{11}\prod_{j=n_0t+1}^{n_0t+k} M(X_j),\end{aligned}$$

using Lemma 2′ and (A2). Then taking logarithms and recalling that $E|\log M(X_1)| < \infty$, we find that

$$\lim_{n\to\infty} n^{-1}\log({}^nY^1)_{11} = \lim_{t\to\infty}(n_0t)^{-1}\log({}^{n_0t}Y^1)_{11} = \log\lambda \qquad \text{a.s.}$$

To deal with other values of i, j note that

$$({}^{n+2n_0}Y^1)_{ij} \geqslant ({}^{n+2n_0}Y^{n+n_0+1})_{i1}({}^{n+n_0}Y^{n_0+1})_{11}({}^{n_0}Y^1)_{1j}, \tag{23}$$

each term on the right-hand side being positive for $n > n_0$ in view of (A1) while (5) for $i=j=1$ and stationarity give

$$\lim_{n\to\infty} n^{-1}\log({}^{n+n_0}Y^{n_0+1})_{11} = \log\lambda \qquad \text{a.s.}$$

Furthermore,

$$({}^{n+2n_0}Y^{n+n_0+1})_{i1} \leqslant K^{n_0}\prod_{j=n+n_0+1}^{n+2n_0} M(X_j)$$

$$({}^{n_0}Y^1)_{ij} \leqslant K^{n_0}\prod_{j=1}^{n_0} M(X_j)$$

and, since $E|\log M(X_1)| < \infty$, we have from (23) that

$$\liminf_{n\to\infty} n^{-1}\log({}^{n+2n_0}Y^1)_{ij} \geqslant \log\lambda \qquad \text{a.s.}$$

Similarly, the inequality

$$({}^{n+2n_0}Y^1)_{11} \geqslant ({}^{n+2n_0}Y^{n+n_0+1})_{1i}({}^{n+n_0}Y^{n_0+1})_{ij}({}^{n_0}Y^1)_{j1}$$

with $n > n_0$ together with (5) for $i=j=1$ and stationarity leads to

$$\log\lambda \geqslant \limsup_{n\to\infty} n^{-1}\log({}^nY^1)_{ij} \qquad \text{a.s.}$$

and hence (5) follows.

Now for the proof of the central limit result (8) it is convenient to use Theorem 7.7, p. 223 of Hall and Heyde (1980). Consequently, we shall begin by showing that their result on convergence to normality can be strengthened to mixing convergence.

Following the notation of Theorem 7.7 and its proof we have that $\{x_{st}, s < t\}$ is a subadditive process and there is a stationary uniform mixing process $\{y_k, -\infty < k < \infty\}$ such that $x_{st} = y_{st} + z_{st}$, where $y_{st} = \sum_{k=s}^{t-1} y_k$ and z_{st} is a nonnegative subadditive process with $z_{0t}/t^{1/2} \to^p 0$ as $t \to \infty$. Furthermore, $Ey_0 = \gamma$, γ being the time constant of the process $\{x_{st}\}$ and $t^{-1}E(y_{0t} - t\gamma)^2 \to \sigma^2$ as $t \to \infty$.

For any $E \in \mathscr{F}$ with $P(E) > 0$ and any $\varepsilon > 0$ we have

$$P(t^{-1/2}|x_{0t} - y_{0t}| > \varepsilon \mid E) \leqslant (P(E))^{-1}P(t^{-1/2}|x_{0t} - y_{0t}| > \varepsilon) \to 0$$

as $t \to \infty$ so that in order to show that

$$(t\sigma^2)^{-1/2}(x_{0t} - t\gamma) \xrightarrow{d} N(0, 1) \text{ (mixing)},$$

it suffices to show that

$$(t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \xrightarrow{d} N(0, 1) \text{ (mixing)}.$$

Now suppose that E_k belongs to the σ-field generated by $y_0, \dots, y_k$ and $P(E_k) > 0$. Then, with $\varepsilon > 0$ and writing

$$F_t = \left\{ (t\sigma^2)^{-1/2} \sum_{i=0}^{[t^{1/4}]} y_i \leqslant \varepsilon \right\},$$

where $[x]$ is the integer part of x, while $\bar{F}_t$ denotes the complement of F_t, we have that

$$P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x \mid E_k)$$

has the same limit behaviour as

$$P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x, F_t \mid E_k),$$

since

$$P(\bar{F}_t \mid E_k) \leqslant E\left(\sum_{i=0}^{[t^{1/4}]} y_i\right)^2 \Big/ t\sigma^2\varepsilon^2 P(E_k) \to 0$$

as $t \to \infty$. But,

$$\begin{aligned} P\left((t\sigma^2)^{-1/2}\left(\sum_{i=[t^{1/4}]+1}^{t} y_i - t\gamma\right) \leqslant x - \varepsilon \mid E_k\right) \\ \leqslant P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x, F_t \mid E_k) \\ \leqslant P\left((t\sigma^2)^{-1/2}\left(\sum_{i=[t^{1/4}]+1}^{t} y_i - t\gamma\right) \leqslant x + \varepsilon \mid E_k\right) \end{aligned}$$

and, since $\{y_i\}$ is uniform mixing and k is fixed, $\sum_{i=[t^{1/4}]+1}^{t} y_i$ and E_k are asymptotically independent as $t \to \infty$ and

$$\begin{aligned} \Phi(x - \varepsilon) &\leqslant \liminf_{t\to\infty} P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x, F_t \mid E_k) \\ &\leqslant \limsup_{t\to\infty} P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x, F_t \mid E_k) \leqslant \Phi(x + \varepsilon), \end{aligned}$$

where Φ is the distribution function of the unit normal law. Since $\Phi(x)$ is continuous in x and ε can be chosen arbitrarily small we have that

$$\lim_{t\to\infty} P((t\sigma^2)^{-1/2}(y_{0t} - t\gamma) \leqslant x \mid E_k) = \Phi(x)$$

and this establishes the required mixing property using the general criterion of Rényi and Révész (1958).

Next we proceed to check the conditions of Theorem 7.7 of Hall and Heyde (1980) in relation to the subadditive process

$$x_{st} = \{-\log({}^{n_0 t}Y^{n_0 s+1})_{11}, s < t\}.$$

We have

$$\begin{aligned} \frac{({}^{tn_0}Y^1)_{11}}{({}^{tn_0}Y^{n_0+1})_{11}} &= \frac{\sum_{i=1}^{K} ({}^{tn_0}Y^{n_0+1})_{1i}({}^{n_0}Y^1)_{i1}}{({}^{tn_0}Y^{n_0+1})_{11}} \\ &\leqslant (KC)^{2n_0} KM({}^{n_0}Y^1), \end{aligned}$$

using Lemma 2′ and (A2). Similarly, for a lower bound we have

$$\begin{aligned} \frac{({}^{tn_0}Y^1)_{11}}{({}^{tn_0}Y^{n_0+1})_{11}} &\geqslant (KC)^{-2n_0} \sum_{i=1}^{K} ({}^{n_0}Y^1)_{i1} \\ &\geqslant (KC)^{-2n_0} m({}^{n_0}Y^1) \\ &\geqslant (KC)^{-3n_0} M({}^{n_0}Y^1) \end{aligned}$$

and thus

$$\begin{aligned} |\log({}^{tn_0}Y^1)_{11}/({}^{tn_0}Y^{n_0+1})_{11}| &= |x_{0t} - x_{1t}| \\ &\leqslant |\log M({}^{n_0}Y^1)| + 3n_0 \log KC \qquad (24) \end{aligned}$$

which provides the boundedness condition (7.55) of Hall and Heyde, since $E|\log M(X_1)|^2 < \infty$ ensures that

$$\begin{aligned} E|\log M({}^{n_0}Y^1)|^2 &\leqslant 2(n_0 \log K)^2 + 2E|\log M(X_{n_0}) + \cdots + \log M(X_1)|^2 \\ &\leqslant 2(n_0 \log K)^2 + n_0 2^{n_0} E|\log M(X_1)|^2 < \infty. \end{aligned}$$

To deal with condition (7.56) we first observe that

$$({}^{(n+m)n_0}Y^1)_{11}/({}^{(n+m)n_0}Y^{n_0+1})_{11} = \alpha_n + \beta_n,$$

where

$$\alpha_n = \sum_{i=1}^{K} ({}^{nn_0}Y^{n_0+1})_{.i}({}^{n_0}Y^1)_{i1}/({}^{nn_0}Y^{n_0+1})_{.1}$$

and

$$\beta_n = \frac{\sum_{l=1}^{K}\sum_{j=1}^{K} ({}^{(n+m)n_0}Y^{nn_0+1})_{1j}\left[\frac{({}^{nn_0}Y^{n_0+1})_{jl}}{({}^{nn_0}Y^{n_0+1})_{.l}} - \frac{({}^{nn_0}Y^{n_0+1})_{j1}}{({}^{nn_0}Y^{n_0+1})_{.1}}\right]({}^{nn_0}Y^{n_0+1})_{.i}({}^{n_0}Y^1)_{i1}}{\sum_{j=1}^{K} ({}^{(n+m)n_0}Y^{nn_0+1})_{1j}({}^{nn_0}Y^{n_0+1})_{j1}}.$$

Now, using Lemma 2′ we have

$$(KC)^{-2n_0}({}^{n_0}Y^1)_{.1} \leqslant \alpha_n \leqslant (KC)^{2n_0}({}^{n_0}Y^1)_{.1}. \tag{25}$$

Also, using Lemmas 2′ and 3′ we have for $n \geqslant 4$,

$$|\beta_n| \leqslant R_{(n-2)n_0}\frac{\sum_{l=1}^{K}\sum_{j=1}^{K} ({}^{(n+m)n_0}Y^{nn_0+1})_{1j}({}^{nn_0}Y^{n_0+1})_{.i}({}^{n_0}Y^1)_{i1}}{\sum_{j=1}^{K} ({}^{(n+m)n_0}Y^{nn_0+1})_{1j}({}^{nn_0}Y^{n_0+1})_{j1}}$$
$$\leqslant R_{(n-2)n_0}K(KC)^{2n_0}({}^{n_0}Y^1)_{.1}. \tag{26}$$

Then, from (25) and (26),

$$|\beta_n|/\alpha_n \leqslant (KC)^{4n_0+1}R_{(n-2)n_0}$$

and, since

$$\log(\alpha_n + \beta_n) = \log \alpha_n + \log(1 + \beta_n/\alpha_n) = \log \alpha_n + O(|\beta_n|/\alpha_n)$$

as $n \to \infty$ we have

$$x_{0,m+n} - x_{1,m+n} = -\log\{({}^{(n+m)n_0}Y^1)_{11}/({}^{(n+m)n_0}Y^{n_0+1})_{11}\}$$
$$= -\log \alpha_n + O(R_{nn_0}) \tag{27}$$

uniformly in m and points ω of the underlying probability space.

Since $\log \alpha_n$ is $\mathscr{M}_{-\infty}^{nn_0}$-measurable, we have from (27) that

$$(x_{0t} - x_{1t}) - E(x_{0t} - x_{1t} | \mathscr{M}_{-\infty}^{nn_0}) = 0, \qquad t \leqslant nn_0$$
$$= O(R_{nn_0}), \qquad t > nn_0,$$

uniformly in t and ω and condition (7.56) is satisfied.

Finally, upon taking expectations in (27) and writing $g_t = Ex_{0t}$, we have

$$g_{m+n} - g_{m+n-1} = -E \log \alpha_n + O(R_{nn_0}) \tag{28}$$

uniformly in m. Now the subadditive function g_t necessarily satisfies

$t^{-1}g_t \to \gamma = -n_0 \log \lambda$ as $t \to \infty$ (with γ finite in view of (24)) and hence from (27),

$$(2n)^{-1}g_{2n} = (2n)^{-1}\{(g_{2n} - g_{2n-1})$$
$$+ (g_{2n-1} - g_{2n-2}) + \cdots + (g_{n+1} - g_n) + g_n\}$$
$$= (2n)^{-1}\{-nE \log \alpha_n + g_n\} + nO(R_{nn_0}) \to \gamma$$

as $n \to \infty$ which assures that $-E \log \alpha_n \to \gamma$ as $n \to \infty$. Furthermore, (28) gives in particular

$$-E \log \alpha_n + O(R_{nn_0}) = -E \log \alpha_{n+m} + O(R_{(n+m)n_0})$$

and letting $m \to \infty$,

$$|-E \log \alpha_n - \gamma| = O(R_{nn_0}) \tag{29}$$

as $n \to \infty$. From (28) and (29) we then readily conclude that $n^{-1/2}(g_n - n\gamma) \to 0$ as $n \to \infty$. The conditions of Theorem 7.7 of Hall and Heyde (1980) are then satisfied and the proof of (8) for $i = j = 1$ is completed by using (22). Other values of i, j may be dealt with by noting that for $n \geqslant n_0$,

$$|\log({}^nY^1)_{ij} - \log({}^nY^1)_{11}| \leqslant 2n_0 \log KC$$

using Lemma 2′. This completes the proof of Theorem 1.

Proof of Theorem 2. From Lemma 2′ we have for $t \geqslant n_0$,

$$(KC)^{-2n_0} \leqslant ({}^tY^1)_{ij}/({}^tY^1)_{11} \leqslant (KC)^{2n_0} \tag{30}$$

and hence

$$a'Y_t = a'({}^tY^1)\,Y_0$$

satisfies

$$(KC)^{-2n_0}({}^tY^1)_{11}(1'a)(a'Y_0) \leqslant a'Y_t \leqslant K(KC)^{2n_0}({}^tY^1)_{11}(1'a)(a'Y_0). \tag{31}$$

Using (31), the result of Theorem 2 is then immediate from Theorem 1.

The result of Theorem 3 is based on the following proposition which is of independent interest.

PROPOSITION 1. *Let* $\{Z_i\}$ *be a stationary uniform mixing process with* $EZ_1 = 0$, $EZ_1^2 < \infty$, *and write* $S_n = \sum_{i=1}^{n} Z_i$. *Suppose that* $n^{-1}ES_n^2 \to \sigma^2$ *and* $n^{-1/2}E|S_n| \to \sigma(2/\pi)^{1/2}$, $0 \leqslant \sigma < \infty$, *as* $n \to \infty$. *Then, for a sequence* $\{a_i\}$ *of* *positive constants,* $b_n^{-1}\sum_{i=1}^n i^{-1/2}a_i|S_i|$, *where* $b_n=\sum_{i=1}^n a_i$ *is a consistent estimator of* $\sigma(2/\pi)^{1/2}$ *if the following three conditions are satisfied:*

(i) $\alpha(n)=b_n^{-1}\max_{1<i<n}a_i\to 0$ *as* $n\to\infty$.

(ii) $\beta(n)=b_n^{-2}\sum_{i=2}^n i^{-1/2}a_i\sum_{j=1}^{i-1} j^{1/2}a_j\to 0$ *as* $n\to\infty$.

(iii) *There exists* $p=p(n)<n$ *with* $p(n)\uparrow\infty$ *as* $n\to\infty$ *such that*

$$\gamma(n)=b_n^{-2}p^{1/2}\sum_{i=p+1}^n i^{-1/2}a_ib_i\to 0,\qquad \delta(n)=b_n^{-1}b_p\to 0$$

as $n\to\infty$. *These conditions are satisfied, in particular, if* $a_i=i^{-1}$.

Proof. Since $n^{-1/2}E|S_n|\to\sigma(2/\pi)^{1/2}$ as $n\to\infty$ we have from Kronecker's lemma (e.g., Hall and Heyde, 1980, p. 31) that

$$b_n^{-1}\sum_{i=1}^n i^{-1/2}a_iE|S_i|\to\sigma(2/\pi)^{1/2}$$

and hence we need to establish that

$$b_n^{-1}\sum_{i=1}^n i^{-1/2}a_i(|S_i|-E|S_i|)\xrightarrow{p}0$$

which is accomplished by showing that

$$b_n^{-2}E\left(\sum_{i=1}^n i^{-1/2}a_i(|S_i|-E|S_i|)\right)^2\to 0 \tag{32}$$

as $n\to\infty$ when (i), (ii), and (iii) hold. Now

$$\begin{aligned}E&\left(\sum_{i=1}^n i^{-1/2}a_i(|S_i|-E|S_i|)\right)^2\\&=\sum_{i=1}^n i^{-1}a_i^2E(|S_i|-E|S_i|)^2\\&\quad+2\sum_{i=2}^n i^{-1/2}a_i\sum_{j=1}^{i-1} j^{-1/2}a_j(E|S_iS_j|-E|S_i|E|S_j|)\end{aligned}\tag{33}$$

and since $ES_n^2=O(n)$ as $n\to\infty$,

$$b_n^{-2}\sum_{i=1}^n i^{-1}a_i^2E(|S_i|-E|S_i|)^2\to 0$$

as $n\to\infty$ using (i), so that (32) follows from (33) if

$$b_n^{-2}\sum_{i=2}^n i^{-1/2}a_i\sum_{j=1}^{i-1} j^{-1/2}a_j(E|S_iS_j|-E|S_i|E|S_j|)\to 0 \tag{34}$$

as $n\to\infty$.

Next we note that using stationarity we may write for $i>j$,

$$S_i=S_j+Z_{j+1}+\cdots+Z_i=S_j+S^*_{i-j},$$

say, where S^*_{i-j} has the same distribution as S_{i-j}. Then note that

$$\begin{aligned}|E|S_iS_j|&-E|S_i|E|S_j|-E|S_iS^*_{i-j}|+E|S^*_{i-j}|E|S_j||\\&\leqslant E|S_j||\,|S_i|-|S^*_{i-j}|\,|+E|S_j|E|\,|S_i|-|S^*_{i-j}|\,|\\&\leqslant ES_j^2+(E|S_j|)^2\leqslant 2ES_j^2=O(j)\end{aligned}$$

as $j\to\infty$ and hence, using (ii), (34) may be replaced by establishing

$$b_n^{-2}\sum_{i=2}^n i^{-1/2}a_i\sum_{j=1}^{i-1} j^{-1/2}a_j(E|S^*_{i-j}S_j|-E|S^*_{i-j}|E|S_j|)\to 0 \tag{35}$$

as $n\to\infty$.

To deal with (35) we let $p=p(n)<n$ be an integer with $p(n)\uparrow\infty$ as $n\to\infty$ and for $p>2$ decompose (35) into the components summed over $\sum_{i=2}^p\sum_{j=1}^{i-1}$, $\sum_{i=p+1}^n\sum_{j=i-p}^{i-1}$, and $\sum_{i=p+1}^n\sum_{j=1}^{i-p-1}$ which will be denoted by $I_1(n)$, $I_2(n)$, and $I_3(n)$, respectively. Using Schwarz' inequality,

$$\begin{aligned}I_1(n)&\leqslant b_n^{-2}\sum_{i=2}^p i^{-1/2}a_i\sum_{j=1}^{i-1} j^{-1/2}a_j(ES_{i-j}^2ES_j^2)^{1/2}\\&=O\left(b_n^{-2}\sum_{i=2}^p a_i\sum_{j=1}^{i-1}a_j\right)=O(b_n^{-1}b_p)^2=o(1)\end{aligned}$$

as $n\to\infty$ using (iii). Furthermore, since in the sum $I_2(n)$ $i\leqslant j+p$,

$$\begin{aligned}I_2(n)&=b_n^{-2}\sum_{i=p+1}^n i^{-1/2}a_i\sum_{j=i-p}^{i-1} j^{-1/2}a_j(ES_{i-j}^2ES_j^2)^{1/2}\\&=O\left(b_n^{-2}p^{1/2}\sum_{i=p+1}^n i^{-1/2}a_i\sum_{j=i-p}^{i-1}a_j\right)=o(1)\end{aligned}$$

as $n\to\infty$ also using (iii).

Finally, for $i > j + p$ we have

$$S^*_{i-j} = Z_{j+1} + \cdots + Z_{j+p} + Z_{j+p+1} + \cdots + Z_i$$
$$= S'_p + S''_{i-j-p},$$

say, where S'_p and S''_{i-j-p} have, respectively, the distributions of S_p and S_{i-j-p}. Then,

$$\begin{aligned}
&|E|S_jS^*_{i-j}| - E|S_j|E|S^*_{i-j}||\\
&\quad \leqslant |E|S_jS^*_{i-j}| - E|S_j|E|S^*_{i-j}| - E|S_jS''_{i-j-p}| + E|S_j|E|S''_{i-j-p}||\\
&\qquad + |E|S_jS''_{i-j-p}| - E|S_j|E|S''_{i-j-p}||\\
&\quad \leqslant E|S_j||S^*_{i-j}| - |S''_{i-j-p}|| + E|S_j|E||S^*_{i-j}| - |S''_{i-j-p}||\\
&\qquad + |E|S_jS''_{i-j-p}| - E|S_j|E|S''_{i-j-p}||\\
&\quad \leqslant E|S_jS'_p| + E|S_j|E|S'_p| + |E|S_jS''_{i-j-p}| - E|S_j|E|S''_{i-j-p}||\\
&\quad \leqslant 2(ES_j^2ES_p^2)^{1/2} + |E|S_jS''_{i-j-p}| - E|S_j|E|S''_{i-j-p}||. \qquad (36)
\end{aligned}$$

Furthermore, the stationary sequence $\{Z_k\}$ is uniform mixing so that, using a well-known mixing inequality (e.g., Theorem (A.6), p. 278 of Hall and Heyde, 1980),

$$|E|S_jS''_{i-j-p}| - E|S_j|E|S''_{i-j-p}|| \leqslant 2\phi^{1/2}(p)(ES_j^2ES_{i-j-p}^2)^{1/2}, \quad (37)$$

where $\{\phi(k)\}$ with $\phi(k) \to 0$ as $k \to \infty$ are the mixing coefficients. Thus, from (36) and (37) we have for $i > j + p$,

$$|E|S_jS^*_{i-j}| - E|S_j|E|S^*_{i-j}|| = O((jp)^{1/2} + (\phi(p)j(i-j-p))^{1/2})$$

so that

$$\begin{aligned}
I_3(n) &= b_n^{-2} \sum_{i=p+1}^{n} i^{-1/2}a_i \sum_{j=1}^{i-p-1} j^{-1/2}a_j |E|S_jS^*_{i-j}| - E|S_j|E|S^*_{i-j}||\\
&= O\left(b_n^{-2}p^{1/2} \sum_{i=p+1}^{n} i^{-1/2}a_i \sum_{j=1}^{i-p-1} a_j\right)\\
&\quad + O\left(b_n^{-2}\phi^{1/2}(p) \sum_{i=p+1}^{n} i^{-1/2}a_i \sum_{j=1}^{i-p-1} a_j(i-j-p)^{1/2}\right)\\
&= O\left(b_n^{-2}p^{1/2} \sum_{i=p+1}^{n} i^{-1/2}a_ib_i\right) + O(\phi^{1/2}(p)) = o(1)
\end{aligned}$$

as $n \to \infty$ in view of (iii) and the uniform mixing property. This completes the proof of (35) and hence that of Proposition 1. Note that (ii) does not hold for $a_i = i^\alpha$, $\alpha > 1$, while if $a_i = i^{-1}$,

$$\alpha(n) = O(\log n)^{-1}, \qquad \beta(n) = O(\log n)^{-1}$$
$$\gamma(n) = O(\log p(\log n)^{-2}) \leqslant O(\log n)^{-1}, \qquad \delta(n) = O((\log p)^2(\log n)^{-2})$$

as $n \to \infty$.

Remarks (i) A more obvious estimator of σ^2 of the form $b_n^{-1}\sum_{i=1}^n i^{-1}a_iS_i^2$ is not adequate for the purposes of this paper. Consistency will be destroyed in the transition from additive processes to subadditive processes necessary for Theorem 3.

(ii) The estimator $n^{-1}\sum_{i=1}^n i^{-1}S_i^2$ obtained by setting $a_i = 1$, $i \geqslant 1$, in the expression discussed in (i) is not consistent for σ^2. Indeed, in the case where the Z_i are independent and normally distributed with mean 0 and variance σ^2, $\text{var}(n^{-1}\sum_{i=1}^n i^{-1}S_i^2) \to \sigma^4$ as $n \to \infty$.

Proof of Theorem 3. As in the proof of Theorem 1 we let

$$x_{st} = -\log({}^{n_0t}Y^{n_0s+1})_{11}, \qquad s < t,$$

and then using the decomposition theorem for subadditive processes (e.g., Theorem 7.6, p. 216 of Hall and Heyde, 1980) decompose x_{st} into

$$x_{st} = y_{st} + z_{st}, \tag{38}$$

where $\{y_{st}, s < t\}$ is an additive process with

$$Ey_{0t} = -n_0t\log\lambda = t\lim_{u\to\infty} u^{-1}Ex_{0u}$$

and $\{z_{st}, s < t\}$ is a nonnegative subadditive process with

$$\lim_{t\to\infty} t^{-1}Ex_{0t} = 0.$$

As indicated in the proof of Theorem 7.7 of Hall and Heyde (1980) we may, under the conditions (5), (6), (7) choose $y_{st} = \sum_{k=s}^{t-1} y_k$, where $\{y_k\}$ is stationary, uniform mixing, and satisfies

$$t^{-1/2}E|y_{0t} + n_0t\log\lambda| \to (2n_0\sigma^2/\pi)^{1/2}, \; t^{-1}E(y_{0t} + n_0t\log\lambda)^2 \to n_0\sigma^2$$

as $t \to \infty$.

Using Proposition 1 we have

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} |y_{0i} + n_0 i \log \lambda| \xrightarrow{p} (2n_0\sigma^2/\pi)^{1/2} \tag{39}$$

as $t \to \infty$. Also, from the proof of Theorem 1 we have

$$t^{-1/2} E z_{0t} = t^{-1/2}(E x_{0t} + n_0 t \log \lambda) \to 0$$

as $t \to \infty$ and hence

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} E z_{0i} \to 0$$

which implies

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} z_{0i} \xrightarrow{p} 0 \tag{40}$$

as $t \to \infty$. Combining (39) and (40) then gives, in view of (38),

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} |x_{0i} + n_0 i \log \lambda| \xrightarrow{p} (2n_0\sigma^2/\pi)^{1/2} \tag{41}$$

as $t \to \infty$.

Furthermore, from (22) we deduce the existence of a finite constant A such that for $0 \leqslant k < n_0$

$$E |\log({}^{in_0+k}Y^1)_{11} - \log({}^{in_0}Y^1)_{11} - k \log \lambda| < A < \infty$$

and then

$$(\log t)^{-1} \sum_{i=1}^{t} \sum_{k=0}^{n_0-1} i^{-3/2} |\log({}^{in_0+k}Y^1)_{11} - \log({}^{in_0}Y^1)_{11} - k \log \lambda| \xrightarrow{p} 0 \tag{42}$$

as $t \to \infty$, since

$$(\log t)^{-1} \sum_{i=1}^{t} \sum_{k=0}^{n_0-1} i^{-3/2} E |\log({}^{in_0+k}Y^1)_{11} - \log({}^{in_0}Y^1)_{11} - k \log \lambda| \to 0$$

as $t \to \infty$. Consequently, from (41) and (42),

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} \sum_{k=0}^{n_0-1} |\log({}^{in_0+k}Y^1)_{11} - (in_0 + k) \log \lambda|$$
$$\xrightarrow{p} n_0^{3/2}(2\sigma^2/\pi)^{1/2}$$

and hence

$$(\log t)^{-1} \sum_{i=1}^{t} \sum_{k=0}^{n_0-1} (in_0 + k)^{-3/2} |\log({}^{in_0+k}Y^1)_{11} - (in_0 + k) \log \lambda|$$
$$\xrightarrow{p} \sigma(2/\pi)^{1/2}$$

or, equivalently,

$$(\log t)^{-1} \sum_{i=n_0}^{n_0 t} i^{-3/2} |\log({}^{i}Y^1)_{11} - i \log \lambda| \xrightarrow{p} \sigma(2/\pi)^{1/2}$$

as $t \to \infty$ and

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} |\log Z_i - i \log \lambda| \xrightarrow{p} \sigma(2/\pi)^{1/2} \tag{43}$$

as $t \to \infty$ follows using (31).

Finally, we have from Theorem 2 that

$$(t^{1/2} \log t)^{-1}(\log Z_t - t \log \lambda) \xrightarrow{p} 0$$

as $t \to \infty$ and hence

$$(\log t)^{-1} \sum_{i=1}^{t} i^{-3/2} |\log Z_i - i t^{-1} \log Z_t| \xrightarrow{p} \sigma(2/\pi)^{1/2}$$

as $t \to \infty$ follows from (43). This completes the proof of Theorem 3.

4. Finite-Sample Estimators

This section considers the problems of estimating $\log \lambda$ and σ given a finite set of data. Recall that $\{Y_t, t = 0, 1, 2,...\}$ is a time-series or sample path of K-dimensional vectors generated by (1). Each vector Y_t corresponds to a population census by age. $Z_t = a'Y_t$ is a nonzero nonnegative linear functional of Y_t, e.g., the total population size, or the number of individuals in any age class or set of age classes, or in an economic context, if a is the vector of labor force participation rates, the size of the labor force at time t. Suppose one is given T consecutive data points $Z_{t_0+1},..., Z_{t_0+T}$, and t_0 is unknown.

The estimator (11) for $\log\lambda$ apparently fails to use most of the data. Both estimators (11) and (12) give a privileged role to Z_{t_0+1}, and (11) also gives a privileged role to Z_{t_0+T}. It is clear from Theorem 2 that for large T,

$$(T-1)^{-1}(\log Z_{T+t_0} - \log Z_{t_0+1})$$

is a consistent estimator of $\log\lambda$, and similarly from Theorem 3 that for large T

$$(\pi/2)^{+1/2}(\log(T-1))^{-1}\sum_{i=1}^{T-1} i^{-3/2}\,|\log Z_{t_0+1+i} - \log Z_{t_0+1} - i\log\hat{\lambda}|$$

is a consistent estimator of σ. Intuitively, one might hope that some averaging of such consistent estimators might produce estimators with lower finite-sample variance than those of the estimators (11) and (12). The numerical results to be presented in this section show that, at least for the example considered, this intuition fails for (11) but appears to be correct for (12).

For numerical investigations of possible estimators of $\log\lambda$ and σ, a time series of 1000 values $W_t = \log Z_t$, $t = 1,..., 1000$ was constructed as follows. For all t, X was taken to be a 15×15 matrix with all elements equal to 0 except

$$x_{1,4} = 17{,}500,\quad x_{1,5} = 85{,}000,\quad x_{1,6} = 175{,}000$$
$$x_{1,7} = 265{,}000,\quad x_{1,8} = 340{,}000,\quad x_{1,9} = 450{,}000$$
$$x_{1,10} = 500{,}000,\quad x_{1,11} = 600{,}000,\quad x_{1,12} = 750{,}000$$
$$x_{1,13} = 900{,}000,\quad x_{1,14} = 1{,}050{,}000,\quad x_{1,15} = 1{,}200{,}000$$
$$x_{i+1,i} = 0.5,\qquad i = 2,..., 14,$$

and the values of $x_{2,1}$ were chosen independently over time with probability $\frac{1}{13}$ from the following list of 13 numbers:

$$2.98\times 10^{-7},\ 3.23\times 10^{-5},\ 1.59\times 10^{-5},\ 4.5\times 10^{-6},\ 7.39\times 10^{-6},$$
$$5.74\times 10^{-6},\ 2.13\times 10^{-5},\ 8.32\times 10^{-6},\ 5.84\times 10^{-6},\ 3.21\times 10^{-5},$$
$$9.82\times 10^{-6},\ 1.08\times 10^{-5},\ 7.05\times 10^{-6}.$$

These numerical values define one version of the striped bass population model of Cohen, Christensen, and Goodyear (1983), which is described

briefly in the next section. Computations were carried out in BASIC on a Tektronix 4052 desk-top computer.

An initial vector Y_{-100} was defined to have all 15 elements equal to 1. One hundred random matrices X_t, $t = -99, -98,..., 0$ were generated as just described and vector Y_0 was computed according to (1). The purpose of the procedure so far was to eliminate from Y_0 any effects of the inevitably arbitrary choice of Y_{-100}. Then another 1000 random matrices X_t, $t = 1, 2,..., 1000$ were generated as above, Y_t was again computed from (1), and $W_t = \log 1'Y_t$, $t = 1,..., 1000$ were recorded for future analysis. In terms of the population model, W_t is the logarithm of total population size, but the scale has no meaning.

Let $G = T - 1$ be the "gap" between the epoch $t_0 + 1$ of the first data point in the sample and the epoch $t_0 + T$ of the last data point. Values of G were chosen to cover the range likely to be observed in biological applications: $G = 3, 9, 27, 81$. For each G, and for all $t_0 = 0, 1, 2,..., 1000 - T$, nine estimators L_i, $i = 1,..., 9$ of $\log\lambda$ and four estimators D_i, $i = 1,..., 4$ of σ were computed:

$$L_1(G, t_0) = (W_{t_0+G+1} - W_{t_0+1})/G,$$
$$L_2(G, t_0) = (\tfrac{1}{2})[L_1(G, t_0) + (W_{t_0+G+1} - W_{t_0+2})/(G-1)],$$
$$L_3(G, t_0) = (\tfrac{1}{3})[2L_2(G, t_0) + (W_{t_0+G} - W_{t_0+1})/(G-1)],$$
$$L_4(G, t_0) = (\tfrac{1}{4})[3L_3(G, t_0) + (W_{t_0+G} - W_{t_0+2})/(G-2)],$$
$$L_5(G, t_0) = (\tfrac{1}{6})[4L_4(G, t_0) + (W_{t_0+G+1} + W_{t_0+G-1} - W_{t_0+3} - W_{t_0+1})/(G-2)],$$
$$L_6(G, t_0) = (\tfrac{1}{2})[L_1(G, t_0) + (W_{t_0+G+1} + W_{t_0+G} - W_{t_0+1} - W_{t_0+2})/(G-1)],$$
$$L_7(G, t_0) = (2W_{t_0+G+1} + W_{t_0+G} - W_{t_0+2} - 2W_{t_0+1})/(3G-2),$$
$$L_8(G, t_0) = (\tfrac{1}{3})[2L_6(G, t_0) + (W_{t_0+G} - W_{t_0+2})/(G-2)],$$
$$L_9(G, t_0) = (\tfrac{1}{2})(W_{t_0+G} + W_{t_0+G+1} - W_{t_0+1} - W_{t_0+2})/(G-1),$$
$$D_1(G, t_0) = (\pi/2)^{1/2}[\log G]^{-1}\sum_{j=1}^{G} j^{-3/2}\,|W_{t_0+j+1} - W_{t_0+1} - jL_1(G, t_0)|,$$
$$D_2(G, t_0) = (\pi/2)^{1/2}[\log G]^{-1}\sum_{j=1}^{G} j^{-3/2}\,|W_{t_0+j+1} - W_{t_0+1} - jL_5(G, t_0)|,$$
$$D_3(G, t_0) = (\tfrac{1}{2})[D_1(G, t_0) + D_1(G-1, t_0+1)],$$
$$D_4(G, t_0) = (\tfrac{1}{2})[D_2(G, t_0) + D_2(G-1, t_0+1)].$$

L_1 is identical to the estimator of $\log\lambda$ in (11). L_2, L_3, L_4, and L_5 average increasing numbers of pairs of points. L_6 and L_7 are two other plausible ways of weighting the same pairs of data points that are used in

L_3; similarly L_8 and L_9 are two other ways of weighting the pairs of data points used in L_4.

D_1 and D_2 are identical to the estimator of σ in (12) but use, respectively, L_1 and L_5 as the estimator of $\log \lambda$. D_3 and D_4 use, loosely speaking, two initial epochs t_0 and $t_0 + 1$, with D_3 based on L_1 and D_4 based on L_5.

The means and standard deviations for each estimator and each value of G were computed, for $i = 1, 2, \ldots, 9$, according to

$$M(L_i(G)) = \sum_{t_0=0}^{1000-G-1} L_i(G, t_0)/(1000 - G),$$

$$S(L_i(G)) = \left[\sum_{t_0=0}^{1000-G-1} (L_i(G, t_0) - M(L_i(G)))^2/(1000 - G) \right]^{1/2},$$

and similarly for D_i, $i = 1, 2, 3, 4$. Table I gives the results.

According to Table I, for every value of G, except $G = 81$, the estimator of $\log \lambda$ with the smallest standard deviation is L_1. For $G = 81$, the standard deviation of L_5 is slightly smaller than that of L_1. That L_1 has the smallest deviation is consistent with the asymptotic result in the remark following Theorem 2, but still surprisingly because L_1 is based on only one pair of data points. We recommend L_1 as the best among the estimators of $\log \lambda$ that we have considered.

For values of G greater than 3, the estimator of σ with the smallest standard deviation is D_3. For every G considered, the mean value of D_3 is never the largest or the smallest of the means of the four estimators of σ. We recommend D_3 as the best among the estimators of σ that we have considered.

It is possible to supply some heuristics to suggest why $S(D_3) < S(D_1)$ is not uncommon. Write

$$D_1(G - 1, t_0 + 1) = D_1^*, \qquad D_i(G, t_0) = D_i, \; i = 1, 3.$$

Then $S(D_3) \leqslant S(D_1)$ means

$$\tfrac{1}{4} | S^2(D_1) + S^2(D_1^*) + 2\rho S(D_1) \, S(D_1^*) | \leqslant S^2(D_1), \tag{44}$$

where ρ denotes the correlation between D_1 and D_1^*. Clearly ρ will be positive and close to one. Now the inequality (44) holds if and only if

$$S(D_1^*) \leqslant \{(3 + \rho^2)^{1/2} - \rho\} \, S(D_1). \tag{45}$$

As G increases, $\rho \to 1$. For G within the range to be expected in biological problems we have $\rho < 1$ and $(3 + \rho^2)^{1/2} - \rho > 1$ while $S(D_1^*)$ and $S(D_1)$ are likely to be around the same value. This suggests that (45) should hold more often than not, implying $S(D_3) \leqslant S(D_1)$ is more likely than $S(D_3) > S(D_1)$.

TABLE I

Means and Standard Deviations of Estimators

	Mean	Standard deviation	Mean	Standard deviation
	$T = 1000$	$G = 3$	$T = 1000$	$G = 27$
L_1	−0.083052	0.096232	−0.082667	0.021324
L_2	−0.083051	0.102794	−0.082653	0.021412
L_3	−0.083054	0.101017	−0.082658	0.021375
L_4	−0.083058	0.107645	−0.082653	0.021502
L_5	−0.083053	0.098562	−0.082650	0.021377
L_6	−0.083054	0.099774	−0.082660	0.021339
L_7	−0.083054	0.100303	−0.082658	0.021373
L_8	−0.083058	0.109046	−0.082653	0.021505
L_9	−0.083056	0.103588	−0.082654	0.021493
D_1	0.124968	0.091676	0.141653	0.063425
D_2	0.127451	0.092913	0.144234	0.067470
D_3	0.123544	0.072000	0.142212	0.057170
D_4	0.121388	0.071485	0.142823	0.059069
	$T = 1000$	$G = 9$	$T = 1000$	$G = 81$
L_1	−0.082858	0.040897	−0.082730	0.011729
L_2	−0.082853	0.041403	−0.082719	0.011729
L_3	−0.082846	0.041256	−0.082725	0.011716
L_4	−0.082839	0.042093	−0.082723	0.011730
L_5	−0.082835	0.041809	−0.082721	0.011698
L_6	−0.082849	0.041045	−0.082726	0.011715
L_7	−0.082847	0.041217	−0.082725	0.011716
L_8	−0.082838	0.042156	−0.082723	0.011730
L_9	−0.082840	0.041912	−0.082723	0.011729
D_1	0.153916	0.075775	0.132777	0.052168
D_2	0.164354	0.087292	0.133701	0.053555
D_3	0.157286	0.068402	0.133117	0.047462
D_4	0.160211	0.073828	0.133374	0.048048

Note. Nine estimators $(L_1, \ldots, L_9)$ of $\log \lambda$ and four estimators $(D_1, \ldots, D_4)$ of σ in a simulation of length $T = 1{,}000$ for gaps $G = 3, 9, 27, 81$.

The last part of Theorem 2, which describes a convergence in distribution to a normal law, suggests that the relationship between $S(L_1(G))$, the standard deviation of $L_1(G)$, and σ, which is estimated by $M(D_3(G))$, would be, for large G,

$$S(L_1(G)) \approx \sigma G^{-1/2} \approx M(D_3(G)) \, G^{-1/2}.$$

Consequently, for large G, the ratio $M(D_3(G)) \, G^{-1/2}/S(L_1(G))$ should approximate to 1, a result which can be rigorously established via the apparatus of the proof of Theorem 1. For $G = 3, 9, 27$, and 81, this ratio is 0.75, 1.25, 1.30, and 1.23 according to the values of $M(D_3(G))$ and $S(L_1(G))$ given in Table I. None of these ratios is very far from 1, but even for $G = 81$, the ratio 1.23 differs from 1 enough to suggest that $L_1(81, t_0)$ is still some distance from normality.

We have carried out the same numerical analysis of another example, not to be reported here, and obtained qualitatively similar results. Unfortunately, theory which could determine how general these conclusions are does not seem to be presently available.

In summary, for $\log \lambda$, the simplest estimator (11) has the smallest variance in this numerical example. For σ, one can reduce the variance of the estimator (12) by averaging over two initial epochs $t_0 + 1$ and $t_0 + 2$, using (11) for $\log \hat{\lambda}$, as in D_3.

5. Striped Bass in the Potomac River

The Potomac River, which flows into the Chesapeake Bay on the east coast of the United States, has along its shores breeding sites of the striped bass (*Morone saxatilis*), an important sporting and commercial fish. The Maryland Department of Natural Resources has conducted annually a standardized seining procedure which has yielded the time series shown in Table II. The data for Chesapeake sites other than along the Potomac River are analysed by Goodyear, Cohen, and Christensen (in press). The numbers give the average number of fingerling striped bass caught per beach seine haul. These numbers are interpreted as the second element of the vector Y_t. The first element corresponds to the striped bass eggs, which were not measured directly.

Before employing the estimators presented above, it is important to verify that the model for which these estimators were developed is appropriate to the data. Otherwise the results of using the estimators have little meaning.

Cohen, Christensen, and Goodyear (1983) derived estimates of the fecundity (female eggs laid per female) as a function of years of age of striped bass that are given as the elements of the first row of the matrix X in the previous section. Since the first age class corresponds to eggs, $x_{1,4} = 17{,}500$ is the effective fecundity of a 3-year-old female striped bass, for example.

For the annual probability of survival of a striped bass of age one year or more, Cohen *et al.* considered a range of values, one of which was 0.5. Goodyear *et al.* considered a wider range of post-egg survival probabilities and a refinement in the assumption of constant post-egg survival.

Using the fixed matrix elements (all but $x_{2,1}$) given in the previous section

TABLE II

Average catch (number of individuals of both sexes) per beach seine haul of fingerling striped bass in the Potomac River breeding site of Chesapeake Bay

Year	Young of year[a]
1954	5.2
1955	5.7
1956	6.2
1957	2.6
1958	8.4
1959	1.6
1960	4.3
1961	25.7
1962	19.7
1963	1.1
1964	29.2
1965	3.4
1966	10.5
1967	1.9
1968	0.7
1969	0.2
1970	20.1
1971	8.5
1972	1.8
1973	2.1
1974	1.5
1975	7.7
1976	3.2
1977	1.9
1978	7.9
1979	2.1
1980	2.3
1981	1.4

Note: Based on annual surveys conducted by the Maryland Department of Natural Resources (from Cohen *et al.*, 1983).

[a] Young of year t is taken as an index of the number of 1-year-old female fish at the time of the spawning season in calendar year $t + 1$. Thus the average catch of 2.3 individuals in 1980 indexes the 1-year-old female population in 1981. Data courtesy of Maryland Department of Natural Resources (B. Florence).

and the time series of 1-year-olds given in Table II, Cohen *et al.* obtained 13 estimates for the annual probability of survival of eggs, which were given in the previous section as the values for $x_{2,1}$. Their detailed analysis of these values of $x_{2,1}$ failed to reject the hypothesis that they were independently and identically distributed. Other assumptions of the model (1), such as the absence of density dependence, were also discussed and justified. The neglect of demographic fluctuations in the numbers of births and deaths, conditional on the rates of fecundity and survival, is appropriate in view of the very large numbers of fish in the striped bass populations.

Cohen *et al.* adopted the model used in the previous section before the estimators and supporting theory presented here were available. Their estimator of $\log \lambda$ (their Eq. (35)) is identical to our (11). However, they were not able to estimate σ.

Table III presents the results of using all the proposed estimators, taking W_t as the logarithms of the data in Table II and $G = 27$, since there are 28 data points. According to L_1, the asymptotic growth rate of population size is $\log \hat{\lambda} = -0.049$. According to D_3, $\hat{\sigma} = 0.42$. Thus a 95% confidence interval for $\log \lambda$ is $-0.049 \pm (1.96) \times (0.42)/5.3 = -0.049 \pm 0.155$. The data are thus consistent with the possibility that the population is increasing,

TABLE III

Estimates of $\log \lambda(L_i)$ and $\sigma(D_i)$ for the Striped Bass Data in Table II and for the Same Model Simulation on Which Table I Is Based

	Striped bass data ($G = 27$)	Model simulation ($G = 999$)
	$T = 28$	$T = 1000$
L_1	−0.048599	−0.083135
L_2	−0.051300	−0.083126
L_3	−0.044658	−0.083113
L_4	−0.042569	−0.083101
L_5	−0.044345	−0.083082
L_6	−0.045643	−0.083118
L_7	−0.044708	−0.083113
L_8	−0.042530	−0.083101
L_9	−0.042687	−0.083101
D_1	0.387223	0.108699
D_2	0.382306	0.109229
D_3	0.421421	0.116829
D_4	0.418728	0.117357

Note: Since $G = T - 1$ in these calculations, only one value was obtained for each estimator.

decreasing, or stationary in the long run, a conclusion also reached by Cohen *et al.*

Suppose that the data continued to be collected in the future and that the values of $\log \hat{\lambda}$ and $\hat{\sigma}$ remained at the values presently observed. The number of years of observation that would be required to demonstrate at the 5% significance level that the population is declining is $(z_{\alpha/2}\hat{\sigma}/\log \hat{\lambda})^2 = 284$. By that time, assuming the stock of striped bass continued to decline at nearly 5% per year, there would hardly be any striped bass left to count.

The estimate of $\log \hat{\lambda} = -0.083$ obtained from the simulation of 1,000 years (Table III) falls well within the 95% confidence interval estimated from the data. The 95% confidence interval estimated from the 1,000 simulation points is -0.083 ± 0.007, which includes the estimate $\log \hat{\lambda} = -0.086$ obtained by Cohen *et al.* from a simulation of 1,000 points independently programmed on another computer in another language. Thus the two simulations are consistent with each other and with the data.

Finally, supposing that no further information about the young of year becomes available after 1981, what is a 95% confidence interval for the young of year (average catch of fingerling striped bass per beach seine haul) in 1991?

Given T data points $z_{t_0+1}, \ldots, z_{t_0+T}$ without knowledge of the time origin t_0, (10) can be modified to yield the following approximate $100(1-\alpha)$% confidence interval for the $W_\tau = \log Z_\tau$ at epoch $\tau > t_0 + T$:

$$W_{t_0+T} + (\tau - |t_0+T|)(T-1)^{-1}(W_{t_0+T} - W_{t_0+1})$$
$$\pm\, \hat{\sigma} \min_{\alpha > q > 0} \left((\tau - |t_0+T|)(T-1)^{-1/2} z_{q/2}\right.$$
$$\left. + (\tau - |t_0+T|)^{1/2} z_{(\alpha-q)/[2(1-q)]}\right).$$

In the present example, $t_0 + 1 = 1954$, $t_0 + T = 1981$, $\tau = 1991$, $W_{1954} = \log 5.2$, $W_{1981} = \log 1.4$, and using D_3 we have $\hat{\sigma} = 0.421421$. The factor to be minimized above is

$$\min_{0.05 > q > 0} \left|10(27)^{-1/2} z_{q/2} + 10^{1/2} z_{(0.05-q)/[2(1-q)]}\right| = 11.316.$$

This minimization is easily carried out numerically by scanning a grid of values of q and using a standard approximation for z_β (Abramowitz and Stegun, 1964, 26.2.22, p. 933). The resulting 95% confidence interval for W_τ is (−4.92, +4.62) which implies a 95% confidence interval for average young of year per beach seine haul from 0.007 to 101.42.

Figure 1 shows the natural logarithms of the observed mean number of fingerlings per beach seine haul from 1954 to 1981 and the projected 95% confidence intervals (on the logarithmic scale) for the years from 1982 to

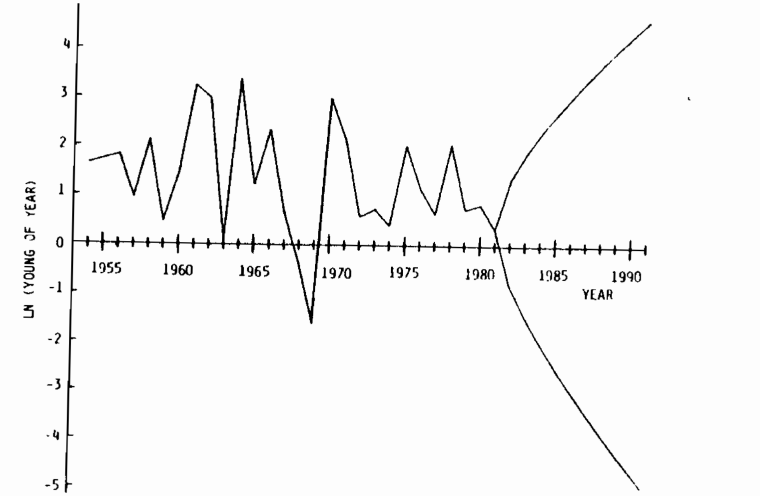

FIG. 1. Natural logarithm of young of year of Potomac striped bass observed from 1954 to 1981, and estimated 95% confidence intervals from 1982 to 1991. Observed values are logarithms of data in Table II.

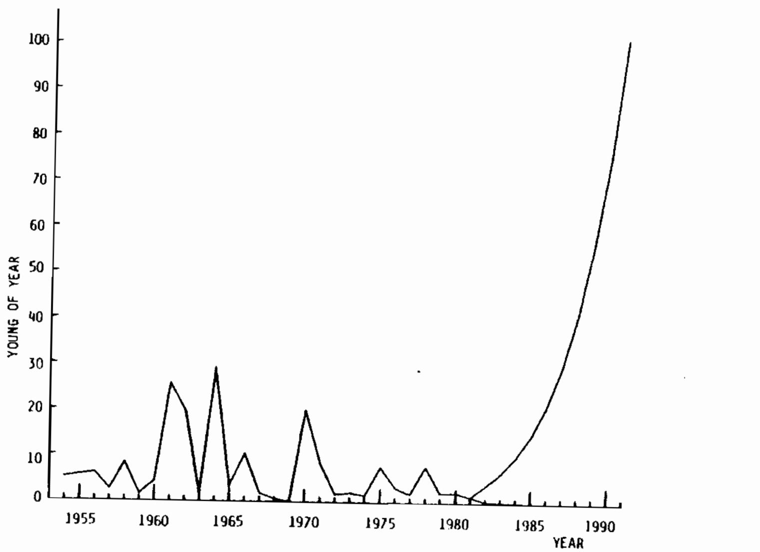

FIG. 2. Young of year (average number of individuals per beach seine haul) of Potomac striped bass observed from 1954 to 1981, and estimated 95% confidence intervals from 1982 to 1991. Observed values are the data in Table II.

1991. Figure 2 gives the observed mean number of fingerlings per beach seine haul and projected confidence intervals.

While the observed annual means fluctuate between 0 and 30 fingerlings per beach seine haul, the projected confidence intervals expand rapidly. Ten years in the future, the confidence intervals include, with low probability, fingerling numbers that have not been seen since 1954 at least.

These wide confidence intervals result from the conjunction of the data, our model of the population dynamics (based on stationary ergodic products of random nonnegative matrices), and the statistical estimators of $\log \lambda$ and σ derived here. If one accepts the data and the model as defensible, one can interpret the width of the projected confidence intervals in two ways.

First, the confidence intervals may be so wide because our estimator of σ gives numerical values that are too large. Research in progress indicates that estimators, based on time-series methods, that are quadratic functions of the data may estimate σ more efficiently than the estimators proposed here. If improved estimators lower our estimates of σ, then the projected confidence intervals will also become narrower.

Second, if improved estimators do not substantially change our numerical estimates of σ, then the existing data and our model of population dynamics together may imply a very wide range of uncertainty about the numbers of fingerlings that can be anticipated even a few years in the future. In this case, both commercial planners and governmental managers of resources would do better to face this great uncertainty squarely than to be surprised constantly by unanticipated fluctuations.

ACKNOWLEDGMENTS

J. E. C. thanks the U. S. National Science Foundation for grant DEB 80-11026, the John D. and Catherine T. MacArthur Foundation for a Fellowship, and the National Bureau of Economic Research (Stanford, through the courtesy of Victor R. Fuchs), and Mr. and Mrs. William T. Golden for hospitality during this work. This collaboration was initiated as a result of an invitation and travel grant to J. E. C. from the Australian Mathematical Society arranged by Robert Gibberd.

REFERENCES

ABRAMOWITZ, M., AND STEGUN, I. A. 1964. "Handbook of Mathematical Functions." U. S. Nat. Bur. of Standards Appl. Math. Ser. 55. Government Printing Office, Washington, D.C.

COHEN, J. E., CHRISTENSEN, S. W., AND GOODYEAR, C. P. 1983. A stochastic age-structured population model of striped bass (*Morone saxatilis*) in the Potomac River, *Canad. J. Fish. Aquatic Sci.* **40**, 2170–2183.

COHEN, J. E., AND SELLERS, P. H. 1982. Sets of nonnegative matrices with positive inhomogeneous products, *Linear Algebra Appl.* **47**, 185–192.

FURSTENBERG, H., AND KESTEN, H. 1960. Products of random matrices, *Ann. Math. Stat.* **31**, 457–469.

GOODYEAR, C. P., COHEN, J. E., AND CHRISTENSEN, S. W. 1985. Maryland striped bass: recruitment declining below replacement. *Trans. Amer. Fish. Soc.*

HAJNAL, J. 1976. On products of non-negative matrices, *Math. Proc. Cambridge Philos. Soc.* **79**, 521–530.

HALL, P., AND HEYDE, C. C. 1980. "Martingale Limit Theory and its Application," Academic Press, New York.

ISHITANI, H. 1977. A central limit theorem for the subadditive process and its application to products of random matrices, *Publ. Res. Inst. Math. Sci. Kyoto Univ.* **12**, 565–575.

PAKES, A. G., AND HEYDE, C. C. 1982. Optimal estimation of the criticality parameter of a supercritical branching process having random environments, *J. Appl. Probab.* **19**, 415–420.

RÉNYI, A., AND RÉVÉSZ, P. 1958. On mixing sequences of random variables, *Acta. Math. Acad. Sci. Hungar.* **9**, 389–393.

ROSENBLATT, M. 1971. "Markov Processes. Structure and Asymptotic Behaviour," Springer-Verlag, New York.

TANNY, D. 1981. On multitype branching processes in a random environment, *Adv. Appl. Probab.* **13**, 464–497.

TULJAPURKAR, S. D., AND ORZACK, S. H. 1980. Population dynamics in variable environments I. Long-run growth rates and extinction, *Theor. Popul. Biol.* **18**, 314–342.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium