ERGODICITY OF AGE STRUCTURE IN POPULATIONS WITH MARKOVIAN VITAL RATES, III: FINITE-STATE MOMENTS AND GROWTH RATE; AN ILLUSTRATION

JOEL E. COHEN, *The Rockefeller University, New York*

Abstract

Leslie (1945) models the evolution in discrete time of a closed, single-sex population with discrete age groups by multiplying a vector describing the age structure by a matrix containing the birth and death rates. We suppose that successive matrices are chosen according to a Markov chain from a finite set of matrices. We find exactly the long-run rate of growth and expected age structure. We give two approximations to the variance in age structure and total population size. A numerical example illustrates the ergodic features of the model using Monte Carlo simulation, finds the invariant distribution of age structure from a linear integral equation, and calculates the moments derived here.

ERGODIC SETS; LESLIE MATRICES; NON-NEGATIVE MATRICES; POPULATION DYNAMICS; PRODUCTS OF RANDOM MATRICES; STOCHASTIC POPULATION MODELS

1. Introduction

We suppose (Cohen (1976), (1977); subsequently referred to as Parts I and II respectively) that the successive Leslie (1945) matrices affecting a closed, single-sex population show a Markovian sequential dependence. Assuming a finite state space of bounded Leslie matrices and an ergodic Markov chain governing transitions in the state space, we give here a simple direct technique for obtaining the exact mean and approximate variance of the number of individuals of each age; the exact mean and approximate variance of the total population size; and the exact long-run rate of population growth. We illustrate the results numerically.

2. The model

An array (vector, matrix, or column or row of a matrix) all of whose elements are non-negative and not infinite is called non-negative; an array all of whose elements are positive and not infinite is called positive; an array all of whose elements are zero is a zero array.

Definition (Hajnal (1976)).   An ergodic set $H(s, g, r)$ is a non-void set of $s \times s$ non-negative matrices with no zero columns and no zero rows such that (1) any product of g factors which are members of $H(s, g, r)$ is positive; and (2) for each h in $H(s, g, r)$, $\min^+(h)/\max(h) > r > 0$, where $\min^+(h)$ is the smallest of the positive elements of h, and r is some positive number independent of h.

Every primitive Leslie matrix (that is, any Leslie matrix some power of which is positive (Sykes (1969), Theorem 6)) is by itself an ergodic set. The incidence matrix of a non-negative matrix is defined as the matrix obtained by replacing each positive element of the matrix by 1. Any finite set of primitive Leslie matrices sharing an identical incidence matrix is an ergodic set.

In demographic applications of the following model, X is the state space of Leslie matrices, S is the set of transition probability matrices for the Markov chain on X, and Y is the set of possible age structures. When the Markov chain is homogeneous, S contains only a single transition matrix t.

If v is a vector with elements $v(i)$, $\| v \| = \Sigma_i |v(i)|$.

The model: Let $X = H_1(k, g_1, r_1)$ be an ergodic set containing s distinct members (s finite) labelled $x_1, \cdots, x_s$, and let $S = H_2(s, g_2, r_2)$ be an ergodic set each of whose members is stochastic. (We define a non-negative matrix to be stochastic if each *column* sums to 1.) Let $\{t_n\}_{n=1}^{\infty}$ be an infinite sequence, with repetitions possible, of members of S, and let $\{X_n\}_{n=1}^{\infty}$ be a Markov chain with state space X and arbitrary initial distribution such that $P[X_{n+1} = x_i \mid X_n = x_j] = t_n(i, j)$. (This is the transpose of the convention adopted in Part II.)

Let Y be the set of all positive column k-vectors with elements which sum to 1. Let $y_0 \in Y$. Define $\{Y_n\}_{n=0}^{\infty}$ to be the family of random vectors with sample space Y such that $Y_0 = y_0$ with probability 1, and for $n > 0$, $Y_n = X_n \cdot Y_{n-1}/\| X_n \cdot Y_{n-1} \|$. Let $\{Z_n\}_{n=1}^{\infty} = \{(X_n, Y_n)\}_{n=1}^{\infty}$.

3. First and second moments; asymptotic growth rate

Definition.   Let s, s', k and k' be positive integers. Let X be a set of s matrices $x_1, \cdots, x_s$ each of size $k \times k'$. Let t be an $s \times s'$ matrix, with elements $t(i, j)$, $i = 1, \cdots, s$; $j = 1, \cdots, s'$. Define $t \otimes X$ to be the matrix of size $ks \times k's'$ containing ss' block submatrices each of size $k \times k'$ in which the i, jth *block* equals $t(i, j)x_i$, $i = 1, \cdots, s$; $j = 1, \cdots, s'$. More explicitly, for $u = 1, \cdots, ks$; $v = 1, \cdots, k's'$, if $u = ik + a$ and $v = jk' + b$ where $a = 1, \cdots, k$; $b = 1, \cdots, k'$, then $(t \otimes X)(u, v) = t(i+1, j+1)x_{i+1}(a, b)$ is the element in the uth

row and vth column of $t \otimes X$. If T is a (not necessarily finite) set of $s \times s'$ matrices, define $T \otimes X = \{t \otimes X : t \in T\}$. (Thus $\otimes$ is defined only when the left argument is one or a set of $s \times s'$ matrices and the right argument is a set containing s labelled matrices.)

When the elements of X are distinctly labelled copies of a single matrix x, then $t \otimes X$ is the Kronecker or tensor product of t and x.

Lemma. If $X = H_1(k, g_1, r_1)$ is a finite ergodic set with s members and $T = H_2(s, g_2, r_2)$ is an ergodic set, then $T \otimes X$ is an ergodic set.

Proof. Let A be a member of $T \otimes X$. A has no zero columns and no zero rows. Also $\min^+(A)/\max(A) > r_1 r_2 > 0$. Let $g = g_1 g_2$; let $A_1, \cdots, A_g$ be any g members of $T \otimes X$; and let $A = A_1 \cdots A_g$. Each block submatrix of A is the sum of a product of a scalar contributed by members of T and a $k \times k$ matrix contributed by X. The scalar is an element of the product of g_1 positive matrices each of which is the product of g_2 members of T (hence positive). The $k \times k$ matrix is a product of g_2 positive matrices, each of which is the product of g_1 members of X (hence positive). Since every block of A is positive, so is A. The lemma is proved.

Under the assumptions and notation of the model in Section 2, let U, an $s \times 1$ matrix, specify the initial distribution of the Markov chain $\{X_n\}$, i.e. $U(i) = P[X_1 = x_i]$, $i = 1, \cdots, s$. Let $B = U \otimes X$ and $A_n = t_n \otimes X$. (B is $sk \times k$; each A_n is $sk \times sk$.) Let J be the $k \times sk$ matrix constructed by juxtaposing horizontally s copies of the $k \times k$ identity matrix I_k. If 1_k is a column of k 1's, $J = (1_k \otimes I_k)^T$. Let $M[i, j]$ refer to the i, jth block submatrix (of size $k \times k$) of any $s'k \times s''k$ matrix M, whatever s' and s'', and $V[i]$ to the ith block k-vector in a column sk-vector V. Round parentheses are reserved for individual elements of a vector $(\cdot)$ or matrix $(.,.)$.

In our notation, a standard result for Markov chains is that $P[X_n = x_i] = (t_{n-1} \cdots t_1 U)(i)$, for $n > 1$. For $n \geq g_2$, since S is an ergodic set, $P[X_n = x_i] > 0$.

Theorem 1 (means).   (i) $E(X_1 y_0) = JBy_0$.

(ii) For all $n \geq 1$, let L_n denote the random matrix $X_n \cdots X_1$. Then for $n \geq g_2$, $E(L_{n+1} y_0 \mid X_{n+1} = x_i) = (A_n \cdots A_1 By_0)[i]/P[X_{n+1} = x_i]$ and

$$E(L_{n+1} y_0 \mid X_{n+1} = x_i)/E(\| L_{n+1} y_0 \| \mid X_{n+1} = x_i)$$

$$= (A_n \cdots A_1 By_0)[i]/\|(A_n \cdots A_1 By_0)[i]\|.$$

(iii) For $n \geq 1$, $E(L_{n+1} y_0) = JA_n \cdots A_1 By_0$.

(iv) If f is a homogenous linear function such that the expectation exists, $E[f(L_{n+1} y_0)] = f(JA_n \cdots A_1 By_0)$.

(v) Let d be the Hilbert projective metric between two positive vectors; that

is, if a and b are positive vectors, $d(a, b) = \ln[\max_i(a(i)/b(i))/\min_i(a(i)/b(i))]$. Let y_0, y_0' be two elements of Y; U, U' any two initial distributions for X_1 $(U, U' \geqq 0$ and $\|U\| = \|U'\| = 1)$.   Then $\lim_n d(E\{L_{n+1}Y_0 \mid Y_0 = y_0; U\}$, $E\{L_{n+1}Y_0 \mid Y_0 = y_0'; U'\}) = 0$. The limit is approached exponentially fast.

Proof. (i) $(By_0)[i] = U(i)x_iy_0 = P[X_1 = x_i]E[X_1y_0 \mid X_1 = x_i]$ so $JBy_0 = \Sigma_i U(i)x_iy_0 = E(X_1y_0)$.

(ii) and (iii). We proceed by induction.

$$(A_1By_0)[i] = \sum_{j=1}^{s} A_1[i, j](By_0)[j] = \sum_{j=1}^{s} t_1(i, j)x_iU(j)x_jy_0$$

$$= \sum_{j=1}^{s} P[X_1 = x_j, X_2 = x_i] x_ix_jy_0$$

$$= P[X_2 = x_i] \sum_{j=1}^{s} P[X_1 = x_j, X_2 = x_i \mid X_2 = x_i] x_ix_jy_0$$

$$= P[X_2 = x_i] E(X_2X_1y_0 \mid X_2 = x_i).$$

Thus $JA_1By_0 = E(X_2X_1y_0)$.

Assuming inductively that

$$(A_{n-1} \cdots A_1By_0)[i]$$

$$= \sum_{j_1=1}^{s} \cdots \sum_{j_{n-1}=1}^{s} P[X_1 = x_{j_1}, \cdots, X_{n-1} = x_{j_{n-1}}, X_n = x_i] x_ix_{j_{n-1}} \cdots x_{j_1}y_0$$

which we have established for $n = 2$, we calculate, using the Markovian structure of $\{X_n\}$,

$$(A_nA_{n-1} \cdots A_1By_0)[i] = \sum_{j=1}^{s} t_n(i, j)x_i(A_{n-1} \cdots A_1By_0)[j]$$

$$= \sum_{j_1=1}^{s} \cdots \sum_{j_{n-1}=1}^{s} \sum_{j=1}^{s} P[X_1 = x_{j_1}, \cdots, X_{n-1} = x_{j_{n-1}}, X_n = x_j, X_{n+1} = x_i]$$

$$\cdot x_ix_jx_{j_{n-1}} \cdots x_{j_1}y_0$$

and, repeating the argument for $n = 2$, find

$$(A_nA_{n-1} \cdots A_1By_0)[i] = P[X_{n+1} = x_i] E(X_{n+1} \cdots X_1y_0 \mid X_{n+1} = x_i),$$

and

$$JA_n \cdots A_1By_0 = \sum_{i=1}^{s} (A_n \cdots A_1By_0)[i] = E(L_{n+1}y_0).$$

(iv) If f is a homogenous linear function of a k-vector such that the expectation exists,

$$E[f(L_{n+1}y_0)] = \sum_{i=1}^{s} E[f(L_{n+1}y_0) \mid X_{n+1} = x_i] P[X_{n+1} = x_i]$$

$$= \sum_{i=1}^{s} f((A_n \cdots A_1 By_0)[i]) = f\left(\left(\sum_{i=1}^{s} (A_n \cdots A_1 B)[i,1]\right) y_0\right)$$

$$= f(JA_n \cdots A_1 By_0).$$

(v) By the lemma, $T \otimes X$ is an ergodic set. Apply (iii) in combination with Theorem 3 of Hajnal (1976). This proves Theorem 1.

For example, if c_k is a row k-vector with elements $c_k(i)$, then $E(c_k L_{n+1}y_0) = c_k JA_n \cdots A_1 By_0$. When X contains only primitive Leslie matrices and $c_k(i) = 1$ for $i = 1, \cdots, k$, $E(c_k L_{n+1}y_0)$ given by Theorem 1(iv) is the expected total population at time $n + 1$ starting from a population of size 1 with age structure y_0 at time 0. If $c_k(i_0) = 1$, $c_k(i) = 0$ for $i \neq i_0$, then Theorem 1(iv) gives the expected number of individuals in age category i_0.

Corollary 1. For all $n \geq 1$, let $t_n = t$. Let $A = t \otimes X$. Then t has positive maximal eigenvalue 1 and A has positive maximal eigenvalue λ, each with corresponding eigenmanifold of dimension 1. Let π denote the unique positive right eigenvector of t associated with eigenvalue 1 such that $\|\pi\| = 1$. Then for any initial distribution U, $\lim_n P[X_n = x_i] = \pi(i)$. Let α be the unique positive right (column) eigenvector of A and β the unique positive left (row) eigenvector of A associated with the eigenvalue λ and satisfying $\|\alpha\| = \Sigma_i \alpha(i)\beta(i) = 1$.
Then (i)

$$\lim_n E(L_{n+1}y_0 \mid X_{n+1} = x_i)/\lambda^n = (\beta By_0)\alpha[i]/\pi(i)$$

and

$$\lim_n E(L_{n+1}y_0)/\lambda^n = J\alpha(\beta By_0) = (\beta By_0)\Sigma_i \alpha[i].$$

(ii) $\lim_n E(L_{n+1}y_0 \mid X_{n+1} = x_i)/E(\|L_{n+1}y_0\| \mid X_{n+1} = x_i) = \alpha[i]/\|\alpha[i]\|$ and $\lim_n E(L_{n+1}y_0)/\|E(L_{n+1}y_0)\| = \Sigma_i \alpha[i]/\Sigma_u \alpha(u)$. The sum in the last numerator is the blockwise sum of s k-vectors, while the sum in the denominator is the scalar sum of all elements of α.

(iii) $\lim_n E[(L_{n+1}y_0)(j) \mid X_{n+1} = x_i]/E[(L_n y_0)(j) \mid X_n = x_i] = \lambda$

and

$$\lim_n E[(L_{n+1}y_0)(j)]/E[(L_n y_0)(j)] = \lambda,$$

for $j = 1, \cdots, k$; $i = 1, \cdots, s$.

(iv) For all i and j, with probability 1, $\lim_n n^{-1} \log[L_n(i,j)] = \lambda$.

Thus λ is the limit proved to exist and denoted as E by Furstenberg and Kesten ((1960), corollary on p. 462). Our Corollary 1 gives an easy explicit means of calculating that limit.

Corrections to "Ergodicity of age structure in populations with Markovian vital rates, III: Finite-state moments and growth rate; an illustration," by Joel E. Cohen. Adv. Appl. Prob. 9, 462-475 (1977).

P. 466. Replace the last four lines by:
(iv) For all i and j, with probability 1, $\lim_n n^{-1} \log [L_n(i,j)] \leq \log \lambda$, with strict inequality in general.

The almost sure limit in (iv) is proved to exist and denoted as E by Furstenberg and Kesten (1960, corollary on p. 462). Our Corollary 1 gives an easy explicit means of calculating an upper bound on that limit. The inequality follows from the upward concavity of log.

P. 470. In Eq. (3), $t^{32} = t^{33} = \pi \cdot 1_3^T$, not $1_3 \cdot \pi^T$, where π is the given 3-vector.

Proof. Let $Q = \alpha\beta$, that is, $Q(i, j) = \alpha(i)\beta(j)$. Q is of rank 1 and $\lim_n A^n/\lambda^n = Q$ (Karlin (1968), pp. 478–479). This matrix limit is to be understood elementwise. Hence $A^n By_0/\lambda^n = QBy_0 = \alpha(\beta By_0)$, where βBy_0 is a scalar. Theorem 1 (ii), (iii) imply (i). Taking norms and dividing gives (ii). (iii) follows from (i). Finally the s^{s_1} matrices consisting of all possible products of g_1 factors drawn from the set X satisfy assumption AI of Furstenberg and Kesten ((1960), p. 462). Since $\{X_n\}$ satisfy the conditions of Theorem 2 of Furstenberg and Kesten ((1960), p. 460), their corollary (p. 462) and Part (i) above imply (iv).

Corollary 2.   Under the assumptions of Theorem 1 and Corollary 1, suppose that $\{X_n\}$ are identically and independently distributed, that is, $t(i, j) = U(i)$ where U describes the distribution of X_1 and $t(i, j) = P[X_{n+1} = x_i \mid X_n = x_j]$. Let $x^* = \Sigma_{i=1}^k U(i)x_i$. Then x^* is primitive and has a positive maximal eigenvalue of multiplicity 1, denoted λ^*. Every eigenvalue of x^* is an eigenvalue of $A = t \otimes X$, and if λ is the positive maximal eigenvalue of multiplicity 1 of the matrix A, then $\lambda = \lambda^*$. Moreover, if α_0 is the positive right eigenvector of A corresponding to λ and α^* is the positive right eigenvector of x^* corresponding to λ^*, normalized so that $\|\alpha_0\| = \|\alpha^*\| = 1$, then $\alpha^* = J\alpha_0$ or $\Sigma_i \alpha_0[i]$.

In demographic terms, the long-run growth rate of a population subject to identically and independently distributed Leslie matrices is the same as the long-run growth rate of a population subject to a single average Leslie matrix.

Proof.   Since x^* is a weighted average of members of an ergodic set, it must be primitive, and therefore possess a spectral radius λ^* as described. Let c be any eigenvalue of x^* with corresponding eigenvector α, so $x^*\alpha = c\alpha$. Define the sk-vector β blockwise by $\beta[j] = U(j)x_j\alpha/c$, $j = 1, \cdots, s$. Then $(A\beta)[i] = U(i)x_i(\Sigma_j\beta[j]) = U(i)x_i(\Sigma_j U(j)x_j\alpha/c) = U(i)x_i\alpha = c\beta[i]$. Thus c is an eigenvalue of A with corresponding eigenvector β. In particular λ^* is an eigenvalue of A, so $\lambda^* \leqq \lambda$. Now let c be any eigenvalue of A with corresponding eigenvector β. Writing $A\beta = c\beta$ blockwise gives $U(i)x_i(\Sigma_j\beta[j]) = c\beta[i]$. Summing over i gives $x^*(\Sigma_j\beta[j]) = c(\Sigma_i\beta[i])$. So λ is an eigenvalue of x^* and $\lambda \leqq \lambda^*$.

Theorem 2 (variances).   Resume the assumptions of Theorem 1.

(i) If f is a homogeneous linear function such that the variance exists, then a first approximation to its variance is the variance of the conditional means, which is

$$\mathrm{Var}\,[E\{f(L_{n+1}y_0) \mid X_{n+1}\}]$$
$$= \sum_{i=1}^{s} (f\{(A_n \cdots A_1 By_0)[i]\})^2/P\,[X_{n+1} = x_i] - (f[JA_n \cdots A_1 By_0])^2.$$

(ii) A second approximation to the variance is obtained by approximating the distribution of $L_{n+1}By_0$ by the s^2 discrete points

$$x_i E(L_n y_0 | X_n = x_j) \qquad \text{with probability} \qquad t_n(i,j) P[X_n = x_i].$$

The resulting approximation is

$$\text{Var}\{f(L_{n+1} B y_0)\} \doteq \sum_{i=1}^{s} \sum_{j=1}^{s} t_n(i,j)[f(x_i(A_{n-1} \cdots A_1 B y_0)[j])]^2 / P[X_n = x_j]$$

$$- (f[J A_n \cdots A_1 B y_0])^2.$$

Proof. (i) For any measurable function g of a vector, the variance when it exists is given by

$$\text{Var}[g(L_{n+1} y_0)] = \text{Var}[E\{g(L_{n+1} y_0) | X_{n+1}\}]$$
$$+ E\{E[g(L_{n+1} y_0) - E\{g(L_{n+1} y_0) | X_{n+1}\}]^2 | X_{n+1}\}$$
$$= \text{the variance of the conditional means}$$
$$+ \text{ the mean of the conditional variances}$$

(Parzen (1962), p. 55). Hence a first approximation to $\text{Var}\{f(\cdot)\}$ for any *linear* f is to ignore the mean of the conditional variances and calculate the variance of the s quantities

$$E\{f(L_{n+1} y_0) | X_{n+1} = x_i\} = f\{(A_n \cdots A_1 B y_0)[i] / P[X_{n+1} = x_1]\}, \quad i = 1, \cdots, s,$$

where $(A_n \cdots A_1 B y_0)[i] / P[X_{n+1} = x_i]$ should be interpreted as 0 if $P[X_{n+1} = x_i] = 0$. Thus as a first approximation

$$\text{Var}[f(L_{n+1} y_0)] \doteq \text{Var}[E\{f(L_{n+1} y_0) | X_{n+1}\}]$$

$$= \sum_{i=1}^{s} (t_n \cdots t_1 U)(i)(f\{(A_n \cdots A_1 B y_0)[i] / (t_n \cdots t_1 U)(i)\})^2$$

$$- \left(\sum_{i=1}^{s} (t_n \cdots t_1 U)(i) f\{(A_n \cdots A_1 B y_0)[i] / (t_n \cdots t_1 U)(i)\} \right)^2$$

which simplifies slightly when f is linear and homogeneous to $\sum_{i=1}^{s} (f\{(A_n \cdots A_1 B y_0)[i]\})^2 / P[X_{n+1} = x_i] - (f[J A_n \cdots A_1 B y_0])^2$.

(ii) As a second approximation to $\text{Var}\{f(L_{n+1} y_0)\}$, we add to the first an approximation to the conditional variances. This approximation is obtained by replacing each exact conditional distribution of $f(L_{n+1} y_0)$ given $X_{n+1} = x_i$, $i = 1, \cdots, s$, by a distribution of s point probability masses. In the approximate distribution of $f(L_{n+1} y_0)$ given $X_{n+1} = x_i$, the jth point is constructed by assuming that $X_n = x_j$ and that $L_n y_0$ exactly equals its conditional expectation $E(L_n y_0 | X_n = x_j)$.

Explicitly, the ith conditional variance, given $X_{n+1} = x_i$, is

$$E\{[f(L_{n+1} y_0)]^2 | X_{n+1} = x_i\} - [E\{f(L_{n+1} y_0) | X_{n+1} = x_i\}]^2.$$

The second term is known exactly. We approximate the first by

$$E\{[f(L_{n+1} y_0)]^2 | X_{n+1} = x_i\} \doteq \sum_{j=1}^{s} P[X_n = x_j | X_{n+1} = x_i] \cdot [f(x_i E\{L_n y_0 | X_n = x_j\})]^2.$$

Since

$$P[X_n = x_j, X_{n+1} = x_i] = t_n(i,j)P[X_n = x_j] = P[X_n = x_j \mid X_{n+1} = x_i]P[X_{n+1} = x_i],$$

we have the transition probabilities of the so-called 'reversed process' associated with $\{X_n\}$ (Karlin (1968), pp. 136–137) as

$$P[X_n = x_j \mid X_{n+1} = x_i] = t_n(i,j)P[X_n = x_j]/P[X_{n+1} = x_i].$$

Thus

$$E\{[f(L_{n+1}y_0)]^2 \mid X_{n+1} = x_i\} \doteq \sum_{j=1}^{s} t_n(i,j)(P[X_n = x_j]/P[X_{n+1} = x_i])$$

$$\cdot [f(x_i(A_{n-1}\cdots A_1 By_0)[j]/P[X_n = x_j])]^2$$

$$= \sum_{j=1}^{s} t_n(i,j)/(P[X_n = x_j]P[X_{n+1} = x_i])[f(x_i(A_{n-1}\cdots A_1 By_0)[j])]^2.$$

Thus our approximation to the mean of the conditional variances is

$$\sum_{i=1}^{s}\sum_{j=1}^{s} t_n(i,j)[f(x_i(A_{n-1}\cdots A_1 By_0)[j])]^2/P[X_n = x_j]$$

$$- \sum_{i=1}^{s} [f((A_n \cdots A_1 By_0)[i])]^2/P[X_{n+1} = x_i].$$

The negative term in this expression equals the positive term in the exact expression for the variance of the conditional means, so that finally our second approximation is

$$\sum_{i=1}^{s}\sum_{j=1}^{s} t_n(i,j)[f(x_i(A_{n-1}\cdots A_1 By_0)[j])]^2/P[X_n = x_j] - (f[JA_{n+1}\cdots A_1 By_0])^2.$$

This proves Theorem 2.

As in the example following Theorem 1, this theorem can give two approximations to the variance of the total population size and the variance of the number of individuals in any age category.

4. Numerical illustration

This artificial numerical example illustrates by Monte Carlo simulation that the distribution of age structure becomes independent of initial conditions. It shows that under homogeneity the long-run distribution of age structure obtained by Monte Carlo simulation can be calculated by numerical solution of a renewal equation in Part II. It finds the mean and approximations to the variance of age structure using Theorems 1 and 2 and shows that these calculations are consistent with the Monte Carlo results. Finally it compares the long-run rate of growth with the growth rates implied by three approximations to the Markovian

model. These comparisons show that the mean age structure under the Markovian model may differ from the stable age structure of the mean of the vital rates obtained by weighting each Leslie matrix according to its long-run probability.

Let $k = 2$; individuals are young (age group 1) or old (age group 2). Let $s = 3$ and $X = \{x_1, x_2, x_3\}$ where

$$(1) \qquad x_1 = \begin{pmatrix} 0.5 & 0.4 \\ 0.8 & 0 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 0.4 & 0.857143 \\ 0.7 & 0 \end{pmatrix}, \qquad x_3 = \begin{pmatrix} 2 & 1 \\ 0.2 & 0 \end{pmatrix}$$

with corresponding maximal eigenvalues $\lambda_1 = 0.8684658$, $\lambda_2 = 1.0000001$, and $\lambda_3 = 2.0954451$. Let the Markov chain $\{X_n\}_{n \in N}$ be homogeneous where $t(i, j) = P[X_{n+1} = x_i \mid X_n = x_j]$ and

$$(2) \qquad t = \begin{bmatrix} 0.5 & 0 & 0.8 \\ 0.3 & 0.9 & 0.2 \\ 0.2 & 0.1 & 0 \end{bmatrix}.$$

To 7 decimal places, the chain's equilibrium probabilities are given by π in

$$(3) \qquad t^{32} = t^{33} = 1_3 . \pi^T = 1_3 (0.1702128, 0.7234043, 0.1063830),$$

where 1_3 is a 3-vector each element of which is 1.

4.1. *Uniform weak ergodicity.* Two hundred independently simulated sample paths of the joint Markov process $\{Z_n\} = \{(X_n, Y_n)\}$ started from $Z_1 = (x_1, y_1)$. Another 200, referred to as $\{Z'_n\} = \{(X'_n, Y'_n)\}$, started from $Z'_1 = (x_2, y_2)$, where (1) gives x_1 and x_2, and

$$(4) \qquad\qquad y_1 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \qquad y_2 = \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix}.$$

Since there are only two age classes, age structure is determined by the scale-free ratio $W_n = Y_n(2)/Y_n(1)$. W_n is the number of old individuals at time n, conditional on the choice, for each realization, of an initial population size such that the number of young individuals at time n is exactly 1. Y_n is not in general independent of total population size $1_2^T \cdot L_n y_0$ at time n nor, equivalently, is $(L_n y_0)(2)$ independent of $(L_n y_0)(1)$. Hence the expectation of ratios $E(W_n) = E[(L_n y_0)(2)/(L_n y_0)(1)]$ is not in general equal to the ratio of expectations $E[(L_n y_0)(2)]/E[(L_n y_0)(1)]$.

To test the equality of distribution of the two random variables W_{201} and $W'_{201} = Y'_{201}(2)/Y'_{201}(1)$, the corresponding sample cumulative distribution functions, each based on 200 simulated values, were compared by a two-sample Kolmogorov–Smirnov test. The difference was not significant at the 0.2 level. The sample mean and sample variance of W_{201} were, respectively, 0.68873 and 0.09367. The sample mean and sample variance of W'_{201} were 0.73658 and 0.12055. These sample variances do not differ significantly, according to the

F-test. The sample means do not differ significantly by the t-test. The sample density functions, with the grouping in Table 1, do not differ significantly, according to a χ^2-test of homogeneity. Thus the 400 simulations offer no evidence for a difference in the distributions of W_{201} and W'_{201}, which had a combined sample mean of 0.71266 and variance of 0.10742.

4.2. *Renewal equation for invariant distribution* (Part II). If $W_n = w_n$ and if the vital rates at the next time are $X_{n+1} = x_i$, then

(5)
$$w_{n+1} = x_i(2, 1)/(x_i(1, 1) + w_n x_i(1, 2)),$$
$$\equiv f_i(w_n).$$

Here f_i is not the linear homogeneous function of Theorem 1. Let B be any Borel set of real numbers (for example, a closed interval). Then in terms of (X_n, W_n) rather than of $(X_n, Y_n) = Z_n$, the limiting or invariant probability F is the solution of

(6)
$$F(\{x_j\} \times B) = \int_w \sum_{i=1}^{3} F(\{x_i\} \times dw)t(j, i)I_B(f_j(w)),$$

for $j = 1, 2, 3$ and every B. Here $I_B(b) = 1$ if $b \in B$, $I_B(b) = 0$ if $b \notin B$.

To find F numerically, we approximated (6) by a family of linear equations and iteratively substituted each trial solution for F into the right side of the approximation to (6).

TABLE 1

Monte Carlo simulation and renewal equation solution for equilibrium distribution of age structure

Intervals B of age ratio $w = y(2)/y(1)$	Monte Carlo simulation frequencies			$400F(X \times B)$ from Equation (6)
	$W_{201} = Y_{201}(2)/Y_{201}(1)$	$W'_{201} = Y'_{201}(2)/Y'_{201}(1)$	combined	
[0, 0.1)	21	20	41	42.552
[0.1, 0.2)	0	0	0	0
[0.2, 0.3)	0	0	0	0
[0.3, 0.4)	0	0	0	0
[0.4, 0.5)	8	6	14	17.872
[0.5, 0.6)	12	7	19	18.306
[0.6, 0.7)	51	49	100	111.693
[0.7, 0.8)	82	79	161	131.617
[0.8, 0.9)	1	3	4	6.553
[0.9, 1.0)	10	6	16	20.340
[1.0, 1.1)	2	8	10	8.510
[1.1, 1.2)	0	0	0	0
[1.2, 1.3)	0	0	0	0
[1.3, 1.4)	0	0	0	0
[1.4, 1.5)	0	0	0	0.408
[1.5, 1.6)	13	22	35	42.144
[1.6, ∞)	0	0	0	0
Total	200	200	400	399.995

Table 1 shows the aggregated frequency distribution of the 400 simulated values of W_{201} and the expected frequencies $400\,F(X \times B)$ over w-intervals B of length 0.1. The goodness of fit of the Monte Carlo frequencies to those predicted from F was assayed after pooling the predicted frequencies over $[1.4, 1.6)$ by assigning 8 degrees of freedom to the X^2 value of 12.227, which is not significant at the 0.1 level.

In Figure 1, the density of F is approximated by a rectangle of appropriate height over each of the 256 intervals of length $w = 0.006250$ in $[0,1.6]$, for each x_j in X. Sometimes the age structure gives full information about the vital rates and (nearly) vice versa. For example, $W_n \in [0,0.1)$ if and only if $X_n = x_3$.

4.3. *Mean and variance of age ratios.* With X given by (1) and t given by (2), $A = t \otimes X$ is

$$(7)\quad A = \begin{bmatrix} 0.25 & 0.2 & 0 & 0 & 0.4 & 0.32 \\ 0.4 & 0 & 0 & 0 & 0.64 & 0 \\ 0.12 & 0.2571429 & 0.36 & 0.7714287 & 0.08 & 0.1714286 \\ 0.21 & 0 & 0.63 & 0 & 0.14 & 0 \\ 0.4 & 0.2 & 0.2 & 0.1 & 0 & 0 \\ 0.04 & 0 & 0.02 & 0 & 0 & 0 \end{bmatrix}$$

with maximal eigenvalue $\lambda = 1.0771776$ and corresponding right (column) and left (row) eigenvectors α and β given by
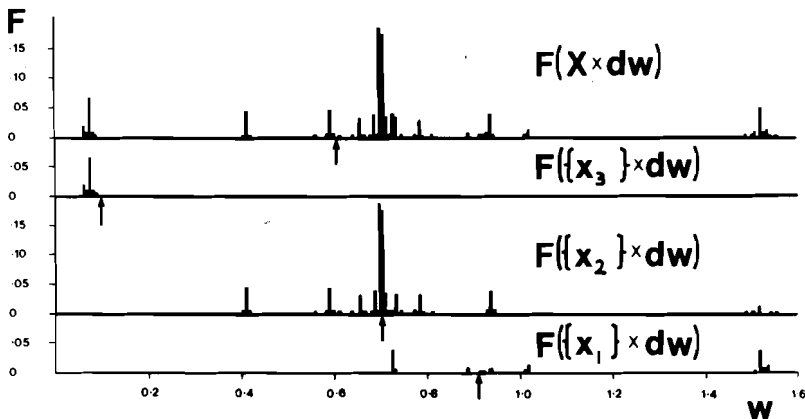

Figure 1

The invariant joint distribution F of age structure and vital rates, assuming Equations (1) and (2). The abscissa w is the ratio $y(2)/y(1)$ of old to young individuals. The top panel, labelled $F(X \times dw)$, approximates the marginal density of w regardless of concurrent rates. Each panel below approximates the joint density of w and the most recently applied Leslie matrix x_j. For each w, the density in the top panel is the sum of the densities in the three below. For each x_j, the sum of the densities over all w is $\pi(j)$ in Equation (3). The arrow under each of the three lower panels indicates the stable age ratio when only the corresponding vital rates are repeatedly applied. The arrow under the top panel marks $\alpha^*(2)/\alpha^*(1)$.

$$\alpha^T = (0.1098627, 0.1316486, 0.3493792, 0.2456306, 0.1529123, 0.0105666)$$

(8)

$$\beta = (1.2682025, 0.6966162, 1.0913724, 0.8819029, 1.0805011, 0.5504359).$$

Thus the long-run expected value of W_n, given that $X_n = x_i$, is $\alpha(2i)/\alpha(2i-1)$, $i = 1, 2, 3$, so that $E(W_n) = \Sigma_i \pi(i)\alpha(2i)/\alpha(2i-1) = 0.7198994$. The ratio of expectations is $[(J\alpha)(2)]/[(J\alpha)(1)] = \Sigma_i \alpha(2i)/\Sigma_i \alpha(2i-1) = 0.6335754$. According to a two-tailed t-test using the Monte Carlo sample variance, the theoretical mean $E(W_n)$ does not differ from the Monte Carlo sample mean of 0.71266 at the 10 percent level. The ratio of expectations was not recorded for the Monte Carlo simulation.

As a first approximation to the variance of W_n, the variance $\mathrm{Var}[E(W_n | X_n)]$ of the conditional mean of W_n, again based on α, is 0.0842221. This value is significantly smaller than the Monte Carlo sample variance of 0.10742 since $X^2 = (399)(0.10742)/(0.084222) = 508.9$ with 399 degrees of freedom.

As a second approximation, we calculate the variance of the random variable W^* which takes values $W(i, j)$ with probabilities $F^*(i, j)$, $i, j = 1, 2, 3$, where $W(i, j)$ is the ratio of the second to the first element of the 2-vector $x_i\alpha[j]$ and $F^*(i, j) = t(i, j)\pi(j)$ from (2) and (3). Explicitly,

(9)
$$W = \begin{bmatrix} 0.8168929 & 1.0240400 & 1.5161827 \\ 0.4904999 & 0.6981755 & 1.5242890 \\ 0.0625332 & 0.0739905 & 0.0966603 \end{bmatrix}.$$

The variance of W^* is then 0.1137813, which is acceptably close to the Monte Carlo sample variance since now $X^2_{399} = 376.7$.

Robert M. May (personal communication, 11 April 1974) and Ronald Lee (personal communication, 9 July 1974) independently suggested other approximations to the long-run distribution of W based on x^* defined in Corollary 2.

4.4. *Comparison of models.* Table 2 compares three long-run parameters of the Markovian model (line 1) with three approximations to that model. The

TABLE 2
Comparison of long-run growth rate and age structure

Model	Growth rate (I)	Age structure, old ÷ young	
		Expectation of ratios (II)	Ratio of expectations (III)
(1) X_n Markovian	1.07718	0.71990	0.63358
(2) X_n i.i.d.	1.07697	0.73600	0.61639
(3) X_n constant	1.07697	0.61639	0.61639
(4) Linear averages	1.09414	0.67333	0.63601

parameters are: the rate of growth (column I), the expected ratio $E(W_n)$ of old to young individuals (column II), and the ratio of the expected number of old to the expected number of young (column III).

Line 2 describes the case of Corollary 2. In spite of being subject to the same matrices of vital rates with the same long-run frequencies as the population in line 1, the population in line 2 has (in this example, but not in general) a lower long-run growth rate. A population which alters (by migration or habitat selection, for example) the sequential dependence among its vital rates, all else remaining constant, can also alter its long-run growth rate. This control of sequential dependence may be an arena for the action of natural selection.

The expected ratio of old to young (column II) differs from the ratio of expected old to expected young individuals (column III), in both lines 1 and 2. In the analysis of variable populations, it is quantitatively important to distinguish between averaging before and averaging after taking proportions.

The two indices of age structure in columns II and III change in qualitatively opposite directions (in this example) as $\{X_n\}$ changes from Markovian (line 1) to independent and identically distributed (line 2). It is not safe to use one index of age structure as a qualitative indicator of the behavior of the other index under a perturbation of the sequential dependence in the vital rates.

The model in line 3 replaces the random X_n by a fixed $x^* = \Sigma_i \pi(i)x_i$. The growth rates and ratios of expectations in lines 2 and 3 are identical (Corollary 2). Within line 3, the two indices of age structure are identical because there is no long-run variation in a population's age structure.

Line 4 demonstrates that one naive approach to estimating the vital rates in lines 1 or 2 does not give exact results. Recall that λ_i is the dominant eigenvalue of x_i. Let α_i be the corresponding right eigenvector, $\|\alpha_i\| = 1$, $i = 1, 2, 3$. Then the 'growth rate' in line 4 is $\Sigma_i \pi(i)\lambda_i$, the 'expectation of ratios' is $\Sigma_i \pi(i)\alpha_i(2)/\alpha_i(1)$ and the 'ratio of expectations' is $(\Sigma_i \pi(i)\alpha_i(2))/(\Sigma_i \pi(i)\alpha_i(1))$. None of these naive estimates matches lines 1, 2 or 3 because none of the parameters is a linear function of the Leslie matrix to which it refers. Another naive approximation to the growth rate, $\Pi\lambda_i^{\pi(i)} = 1.056217$, fails equally.

References

COHEN, J. E. (1976) Ergodicity of age structure in populations with Markovian vital rates, I: Countable states. *J. Amer. Statist. Assoc.* **71**, 335–339.

COHEN, J. E. (1977) Ergodicity of age structure in populations with Markovian vital rates, II: General states. *Adv. Appl. Prob.* **9**, 18–37.*

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

HAJNAL, J. (1976) On products of non-negative matrices. *Math. Proc. Camb. Phil. Soc.* **79**, 521–530.

KARLIN, S. (1968) *A First Course in Stochastic Processes.* Academic Press, New York.

LESLIE, P. H. (1945) On the use of matrices in certain population mathematics. *Biometrika* **33**, 183–212.

PARZEN, E. (1962) *Stochastic Processes.* Holden-Day, San Francisco.

SYKES, Z. M. (1969) On discrete stable population theory. *Biometrics* **25**, 285–293.

* Typographical errors on p. 22 of this paper should be corrected. In Theorem 3(iv), the region of integration should be Z, not C. In Theorem 3(v), ll.3 and 4, x should be x' and y should be y'.