

ERGODIC THEOREMS IN DEMOGRAPHY

BY JOEL E. COHEN¹

ABSTRACT. The ergodic theorems of demography describe the properties of a product of certain nonnegative matrices, in the limit as the number of matrix factors in the product becomes large. This paper reviews these theorems and, where possible, their empirical usefulness. The strong ergodic theorem of demography assumes fixed age-specific birth and death rates. An approach to a stable age structure and to an exponentially changing total population size, predicted by the Perron-Frobenius theorem, is observed in at least some human populations. The weak ergodic theorem of demography assumes a deterministic sequence of changing birth and death rates, and predicts that two populations with initially different age structures will have age structures which differ by less and less. Strong and weak stochastic ergodic theorems assume that the birth and death rates are chosen by time-homogeneous or time-inhomogeneous Markov chains and describe the probability distribution of age structure and measures of the growth of total population size. These stochastic models and theorems suggest a scheme for incorporating historical human data into a new method of population projection. The empirical merit of this scheme in competition with existing methods of projection remains to be determined. Most analytical results developed for products of random matrices in demography apply to a variety of other fields where products of random matrices are a useful model.

1. Introduction. According to his autobiography, Ulam [1976, p. 6] once introduced himself as a pure mathematician who had sunk so low that his latest paper contained numbers with decimal points. This paper will sink—if possible—even lower, to pictures of numbers with decimal points. The reasons are that I make no pretense of being a pure mathematician (although some of my best friends are) and that I will describe a young, not a mature, field of science. This field is still very close to its empirical roots. Consequently, even the mathematical parts of this paper will be framed in concrete language. Many of the assumptions made here can be weakened, at the cost of more technicalities.

The ergodic theorems of demography describe the properties of a product of certain nonnegative matrices, in the limit as the number of matrix factors

An invited address delivered at the 84th Annual Meeting of the American Mathematical Society in Atlanta, Georgia, on January 4, 1978; received by the editors May 1, 1978.

AMS (MOS) subject classifications (1970). Primary 15A48, 60J20; Secondary 92A15, 60B15.

Key words and phrases. Ergodic theorems, random ergodic theorems, products of random matrices, uniform mixing, Perron-Frobenius theorem, contractive mappings, nonnegative matrices, ergodic sets, Markov chains, random environments, age structure, demography, stable populations, age census, Leslie matrix, population projection, Hilbert projective pseudometric, periodic environments.

¹Supported in part by U. S. National Science Foundation grant DEB 74-13276.

© 1979 American Mathematical Society
0002-9904/79/0000-0119/\$06.25

in the product becomes large. We review the historical motivation and applications of these theorems. We present some properties, perhaps surprising, of these products when successive factors are chosen from a set of possible matrices by a Markov chain. We assume elementary knowledge of linear algebra and stochastic processes, but no previous exposure to demography. We give some references to extensions and generalizations and indicate some unanswered questions which may require more mathematical power.

An age-structured population is a set, with membership possibly changing in time, of individuals identified by age. These individuals may be people, other animals or plants, cells, or items of equipment such as railroad ties, light bulbs, and aircraft engines. We will restrict our attention to human populations.

UNITED STATES SUMMARY

GENERAL POPULATION CHARACTERISTICS

Population by Age: 1970 and 1960

NUMBER IN MILLIONS

	1970	1960	YEARS
	7.6	5.6	75+
	5.4	4.7	70-74
	7.0	6.3	65-69
	8.6	7.1	60-64
	10.0	8.4	55-59
	11.1	9.6	50-54
	12.1	10.9	45-49
	12.0	11.6	40-44
	11.1	12.5	35-39
	11.4	11.9	30-34
	13.5	10.9	25-29
	16.4	10.8	20-24
	19.1	13.2	15-19
	20.8	16.8	10-14
	20.0	18.7	5-9
	17.2	20.3	0-4

FIGURE 1. United States population in 1970 and 1960; number in millions in five-year age groups. Source: U. S. Bureau of the Census, 1970 U. S. Census of Population. vol. 1, pt. 1, sec. 1, p. 259.

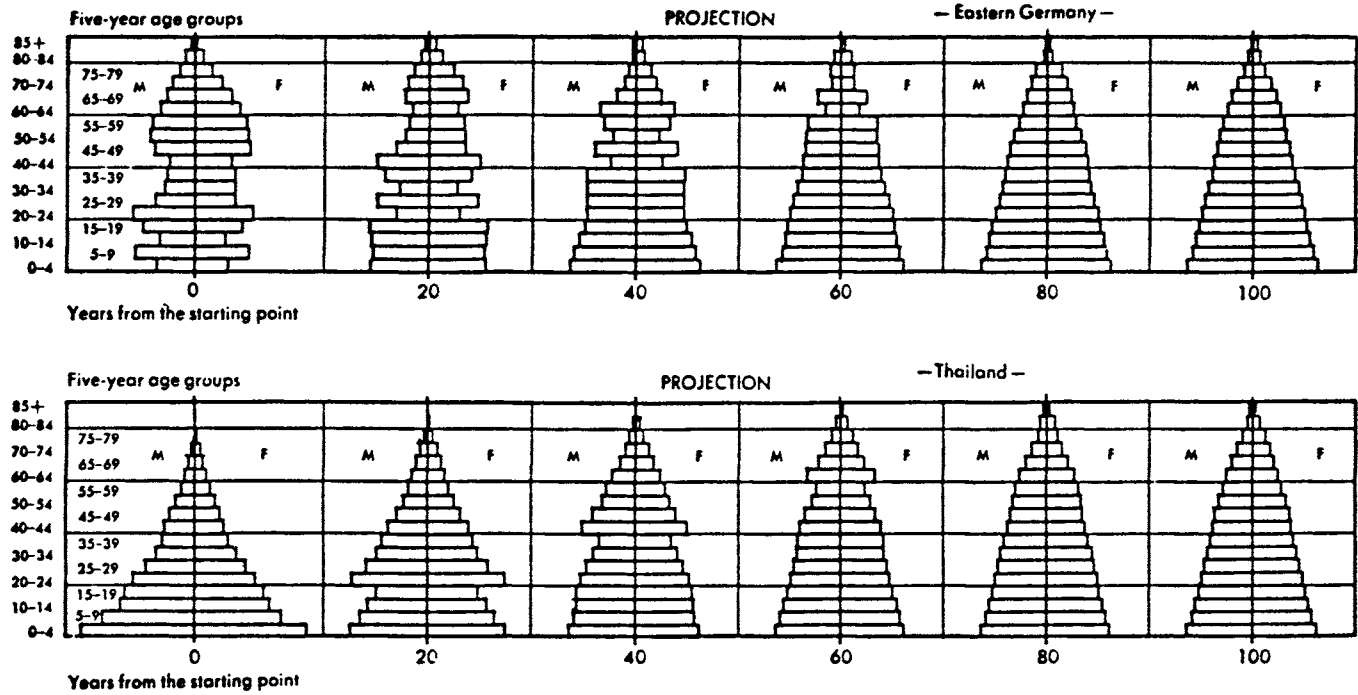


FIGURE 2. Two sets of projections computed on the basis of the population of Eastern Germany in 1957 and of an estimate of the population of Thailand in 1955, respectively; age distribution by five-year age groups. M = male; F = female. Hypothetical vital rates used in both projections assume an expectation of life at birth for both sexes of 60.4 years and a gross reproduction rate of 1.50. Source: Bourgeois-Pichat 1968, p. 6.

The age structure of a population is of interest for both scientific and practical reasons. Censuses show that the proportions of individuals of various ages in national populations vary substantially in time and from place to place. Figure 1 compares the number of individuals of each age in the United States censuses of 1960 and 1970. The leftmost panels of Figure 2 compare the age structure, grossly distorted by war and depression, of East Germany in 1957 (above) with that of the rapidly growing population of Thailand in 1955 (below). These observations raise the scientific question of accounting in quantitative detail for such variation.

From a practical point of view, it is desirable to predict the number of schools which will be needed (as well as the number of teachers and professors in them, of course), the size of the labor force, and the number of people over 65 who may be drawing Social Security benefits. In each of these examples, the quantity of immediate interest, the number of students, working people, or pensioners, depends both on the number of people in the appropriate age class and the proportion of such people who go to school, work, or are retired. So the demography of age-structured populations provides only part of the answers to these practical questions. In other cases, such as a mosquito population divided into larval, pupal, and adult stages, it is safe to assume that every adult female will seek a blood meal. The proportion of adults is of direct interest.

Even if one has no direct interest in the age structure of a population, but would like to improve predictions of total population size, one might plausibly divide a population into homogeneous age classes and apply age-specific birth and death rates to each such class. The overall, or crude, birth and death rates will clearly vary with the proportions of different age classes in the population, because the chance that an individual will have a child or will die in the next year depends on the age of the individual.

By focusing attention on the causes and effects of age structure, we do not intend to ignore the obvious, that birth and death rates depend on many factors besides the age of individuals. Many demographers now believe that one reason for the very limited predictive ability of demography is precisely that it has not paid attention to nondemographic factors which influence demographic variables. Still, it is helpful to start with an understanding of age structure.

To investigate age structure mathematically, we simplify. We treat age and time as discrete. We define age as the number of completed time units since the birth of an individual. We assume, since no one lives forever, a finite number k of age categories. We consider a closed population subject to birth and death only, without immigration or emigration. We consider one sex only. It might appear at first glance that studying populations without sex could hardly be fun, and certainly not useful; but that is not so. Our study of a single sex does not ignore the existence of two sexes in human populations (and of many more than two sexes in, for example, fungal species). We simply assume that there are enough individuals of the other sex (or sexes) not to alter the birth or death rates, as a function of age, of the sex we are studying. In order to avoid repeating the phrase, "birth and death rates," we shall refer to such rates as "vital" rates. We assume that age-specific vital rates apply

uniformly and equally to all individuals in an age class. Finally, we restrict our attention to large populations in which it is sufficient to study expected numbers of births and deaths, conditional on given vital rates. (Schweder [1971] argues for this simplification.) For such large populations, it is reasonable to let the number of individuals in an age class be a continuous nonnegative variable, not restricted to the integers.

Mathematical models of age-structured populations based on a variety of alternatives to these simplifying assumptions have been constructed (Hoppensteadt [1975]; Keyfitz [1977]).

For concreteness, we shall speak in terms of human females. We will use years or multiples of years as our unit of time and age.

2. Censuses, projections, and ergodic theorems. By an age *census* at time t we mean a nonnegative k -vector $Y(t)$, where $k \geq 2$ is the number of age classes and the i th element $Y_i(t) \geq 0$ is the number of females at time t who will be i years old at their next birthday. We adopt the square-block norm $\|Y\| \equiv |Y_1| + \dots + |Y_k|$. By the age *structure* $y(t)$ of a census $Y(t)$ we mean the normalized vector $y(t) = Y(t)/\|Y(t)\|$. Clearly $\|y(t)\| = 1$.

To describe the action of vital rates in transforming an age census at one time into an age census at the next time, we let $x(t)$ be a sequence of operators, $t = 1, 2, \dots$, mapping the nonnegative k -vectors into the nonnegative k -vectors. The basic model we shall consider is given by

$$Y(t+1) = x(t+1)Y(t), \quad t = 0, 1, 2, \dots \quad (1)$$

Particular models of age-structured populations specify the form of $x(t)$ and the choice of the sequence $x(1), x(2), \dots$

Ergodic theorems in demography have the following form: given assumptions about $\{x(t)\}$, describe the long run behavior of population size $\|Y(t)\|$ and of age structure $y(t)$ and show that the behavior of these quantities is independent of initial conditions, over at least some range of initial conditions. "Ergodic" refers here to behavior which is independent of initial conditions, and not, as in statistical mechanics, to the equality of time averages with ensemble averages. For the reader who came this far in the by now disappointed hope of learning about classical ergodic theory, I recommend the lucid introduction, at a high level, by Mackey [1974]. The ergodic theorems which we shall describe are also not to be confused with the development, due to Demetrius ([1974], [1977] and elsewhere), of analogies in population biology to the ergodic theory of statistical mechanics.

We consider three ergodic theorems or classes of theorems. The strong ergodic theorem assumes that $x(t)$ is constant in time t . The weak ergodic theorem assumes that $\{x(t)\}$ is a determinate sequence. Stochastic ergodic theorems assume that $\{x(t, \omega)\}$ is a sample path of a stochastic process which chooses $x(t)$ from a set of possible operators X . As in the deterministic case, strong stochastic ergodic theorems assume that the stochastic process determining $x(t)$ is stationary. Weak stochastic ergodic theorems assume the stochastic process may be nonstationary.

As a further (enormous!) simplification we shall assume that each $x(t)$ is a linear operator, represented by a $k \times k$ projection matrix of the form

$$x(t) = \begin{pmatrix} b_1(t) & b_2(t) & \cdots & b_{k-1}(t) & b_k(t) \\ s_1(t) & 0 & \cdots & 0 & 0 \\ 0 & s_2(t) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{k-1}(t) & 0 \end{pmatrix}. \quad (2)$$

Here $b_i(t) \geq 0$ is the effective fertility per unit time of age class i . The qualification "effective" is necessary because we are assuming that the number of females in age class 1 at $t + 1$ born between t and $t + 1$ to females in age class i at t is $b_i(t + 1)Y_i(t)$. Thus we count only the females born in the interval from t to $t + 1$ who survive to $t + 1$. The newborn females who do not survive to $t + 1$ are not included in the effective fertility rates. The total number of individuals in age class 1 at $t + 1$ is the sum of the contributions from each age class at t :

$$Y_1(t + 1) = \sum_{i=1}^k b_i(t + 1)Y_i(t). \quad (3)$$

In the projection matrix $s_i(t) \geq 0$ is the survival proportion per unit time, $s_i(t) < 1$. Thus the number of females in age class $i + 1$ at $t + 1$ is

$$Y_{i+1}(t + 1) = s_i(t + 1)Y_i(t), \quad i = 1, 2, \dots, k - 1. \quad (4)$$

Equations (2), (3), and (4) specify the details of and are consistent with the basic model (1) if the action of the operator $x(t + 1)$ is now viewed simply as matrix multiplication. Population projections based on specific numerical assumptions for the effective fertility rates and survival proportions were carried out by an English economist Cannan [1895], and by demographers (Bowley [1924]; Whelpton [1936]) long before it was recognized (during World War II by Bernadelli, Lewis, and Leslie; see Keyfitz [1968] for references) that the process could be conveniently formulated in matrix terms.

We shall assume that every projection matrix x of the form (2) in the set X of projection matrices satisfies the further requirements: $s_1 > 0, \dots, s_{k-1} > 0$; $b_{k-1} > 0, b_k > 0$; and the ratio of the smallest positive element of x to the largest element of x is not less than $R > 0$. A consequence of these assumptions is that X is an ergodic set of matrices (Hajnal [1976]). Every element of x^k is positive and every product of any k matrices from X is positive. (We say a matrix is positive if each of its elements is positive; nonnegative, if each element is nonnegative.)

The restrictions we have placed on the elements of each x in X in order to guarantee that the product of any k of them is positive are by far not the weakest sufficient for that conclusion (see Sykes [1969a]; Pollard [1973]). What is important about the restrictions is that they are satisfied by real human populations. If you think I believe that 99- and 100-year-old women are still giving birth (so that $b_{99} > 0$ and $b_{100} > 0$), fear not. For projections, it is possible to truncate the age structure after the last age with positive effective fertility. When the birth sequence has been projected as far as

required, the survivors to all ages can be filled in; females who survive past the last age with positive effective fertility have, according to the assumptions of this model, no effect on future fertility. (Thus there lies hidden in this model the sociological assumption that the availability of grandmothers as babysitters or as competitors for housing has no effect on fertility. Innocuous mathematics may be strong sociology.)

3. The strong ergodic theorem. First we will examine the mathematical consequences of assuming that an age-structured population is repeatedly subject to age-specific vital rates which are constant in time. Then we will briefly review the empirical usefulness of such an assumption.

The strong ergodic theorem is a corollary of the Perron-Frobenius theorem (Seneta [1973]), a beautiful theorem which is worth knowing because of its wide usefulness in economics, ecology, genetics, and the theory of Markov chains, in addition to demography:

Let x be a $k \times k$ nonnegative matrix which is primitive (some power of x is positive). Then

(1) The eigenvalue λ of x which is largest in modulus has algebraic multiplicity 1. (This means that λ is a simple root of the characteristic equation $|\lambda I - x| = 0$.)

(2) λ has geometric multiplicity 1. (This means that for any two column k -vectors V and V' , if $xV = \lambda V$ and $xV' = \lambda V'$, then there exists a nonzero constant c such that $V' = cV$. Similarly for any two row k -vectors W^T and W'^T , if $W^T x = \lambda W^T$ and $W'^T x = \lambda W'^T$, then there exists a nonzero c' such that $W' = c'W$.)

(3) λ is real and positive.

(4) The right and left eigenvectors V and W corresponding to x are positive (elementwise).

(5) $\lim_{t \rightarrow \infty} x^t / \lambda^t = B = VW^T > 0$ where W and V are scaled so that $W^T V = 1$.

λ is called the spectral radius or dominant eigenvalue or Perron-Frobenius root of x , and is written $\lambda = \rho(x)$.

The *strong ergodic theorem* of demography follows from the observation that every matrix x in the set X of projection matrices is primitive:

For all $t = 1, 2, \dots$, let $x(t) = x \in X$. Let $Y(0), Y'(0) \neq 0, Y(0) \neq Y'(0)$ be two nonnegative nonzero and different initial age censuses (k -vectors), and let $Y(t) = x^t Y(0), Y'(t) = x^t Y'(0)$. Then $\lim_{t \rightarrow \infty} Y(t) / \lambda^t = V(W^T Y(0))$. λ is called the stable growth rate per unit of (discrete) time, and $\log \lambda$ is often called the Malthusian parameter or intrinsic rate of natural increase. Moreover, $\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y'(t) = v = V / \|V\|$. v is called the stable age structure.

Thus $Y(t)$ and $Y'(t)$ eventually grow at the same rate λ per unit time and the corresponding age structures eventually approach the same limiting age structure v . $W^T Y(0)$ is called the stable equivalent of $\|Y(0)\|$, which is the initial total population size of the age census $Y(0)$, because if a population with age structure v and total initial population size $W^T Y(0)$ grew geometrically at the rate λ per unit time, that population would eventually come arbitrarily close in total size and age structure to $Y(t)$.

Because of the particularly simple form of a projection matrix (2), it is easy to calculate explicitly the stable age structure in terms of the elements of x and the stable growth rate λ (Pollard [1973, p. 43]).

If $\lambda = 1$, the population is called stationary. Ultimately such a population must cease either to grow or to contract. However, if the initial age structure is not the stable age structure v , then the total population size may very well change as it approaches the stationary limit. A simple expression for the change in population size between the initial age census and the stationary limit has been found for a continuous-time model (Keyfitz [1971b]) and for the discrete model (2) (Lange in press). If λ exceeds or is less than 1 the population will ultimately grow or contract exponentially.

Figure 2 illustrates how two different initial age structures subjected to the same projection matrix converge to the same age structure.

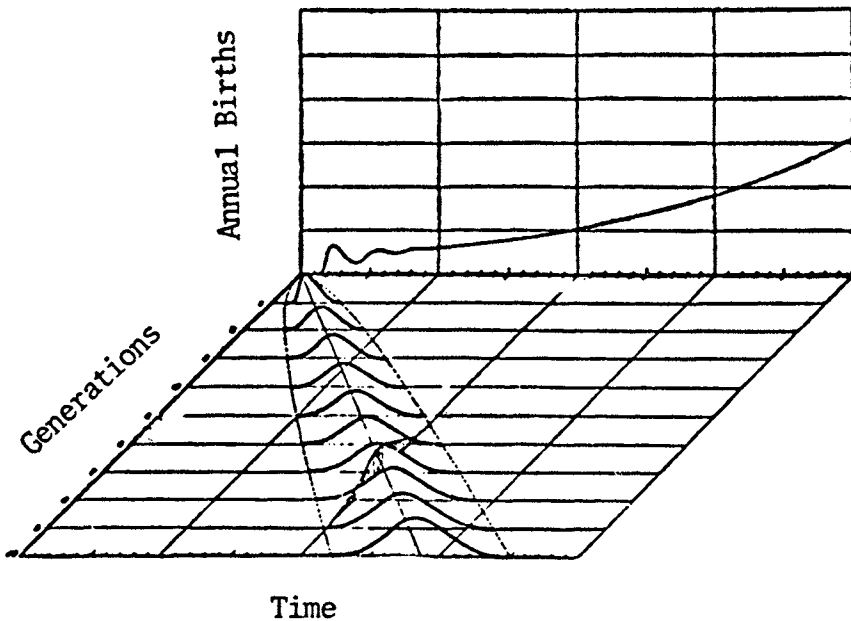


FIGURE 3. Distribution over time of births in a sequence of generations. The curve on the vertical panel at the rear indicates the total number of births to all generations present at a given time. Source: Lotka 1939, p. 80.

Figure 3 shows on the rear panel the number of births per year in a hypothetical population consisting initially only of newborn babies and subject to constant vital rates. At first there are no births. Once the females reach reproductive age there is a wave of births. There is a second but damped wave as the offspring of those births reach reproductive age. The damped waves eventually approach exponential growth. In human populations, the period of these waves is very close to $2\pi/b$ where $\lambda_2 = a + ib$ is the eigenvalue of x next largest in modulus after λ (Keyfitz [1972b]). Invariably $b > 0$ for human populations, since children don't have babies. The plot in the foreground of Figure 3 gives annual births according to the number of generations since the initial birth cohort.

A model of such charming simplicity lends itself to analytical investigations which have occupied (and some would say, preoccupied) mathematical demographers for decades. For example, one can investigate quantitatively and qualitatively the behavior of the stable growth rate λ under perturbations of elements of x due to changes in age-specific vital rates (Demetrius [1969]; Goodman [1971]; Keyfitz [1971a]; Boyce [1977]; Cohen [1978a]; Cohen, submitted). Kato [1976] gives much more general techniques for studying such perturbations. One could investigate the rate of convergence of an age structure to the stable age structure and the convergence of the rate of growth of total population size to exponential growth (Coale [1972]; Keyfitz [1972b]). The rate of convergence depends on the ratio $|\lambda_2|/\lambda$. The convergence of $Y(t)/\lambda^t$ to $BY(0)$ is exponential and complete (Cohen [1979]), in the sense that

$$\lim_{t \rightarrow \infty} \sum_{m=0}^{t-1} (x^m Y(0)/\lambda^m - BY(0)) = (Z - B)Y(0) < \infty,$$

where $Z = (I + B - x/\lambda)^{-1}$ and

$$\lim_{t \rightarrow \infty} \sum_{m=0}^{t-1} |x^m Y(0)/\lambda^m - BY(0)| < \infty.$$

A closed form for the series on the left seems to be unknown.

The history of the strong ergodic theorem illustrates how long it may take for different parts of mathematics and science to become connected in ways that, retrospectively, seem obvious. The Perron-Frobenius theorem was proved in stages between 1907 and 1912. Simultaneously, between 1907 and 1911, Lotka and Sharpe gave the first modern development of the theory of stable populations. They used a model with continuous time and age, in which the characteristic equation for the stable growth rate is an integral equation rather than an algebraic polynomial. (Euler's much earlier discovery of some of the same equations has only recently been recognized. Reprints of the early papers of Euler, Lotka and Sharpe are now readily available; see Weiss and Ballonoff [1975], Smith and Keyfitz [1977].) The relevance of the Perron-Frobenius theorem to the theory of stable populations in discrete age and time did not become apparent until the matrix formulation of population projection during World War II. The full reconciliation of the matrix approach, the integral equation approach of Lotka and Sharpe, and some other equivalent formulations of stable population theory did not come for another score of years after World War II (Keyfitz [1968]).

Aside from its aesthetic virtues, the strong ergodic theorem has retained the interest of demographers for so long because it has considerable practical use. Given a projection matrix x based on current birth and death rates, the long run rate of growth λ and the stable age structure v indicate what would happen if the vital rates in x were maintained indefinitely. A speedometer on a car serves the same function: if it registers 90 kilometers per hour, that is not necessarily a prediction that the car will be 90 kilometers distant after one hour, but is an indicator of the present velocity.

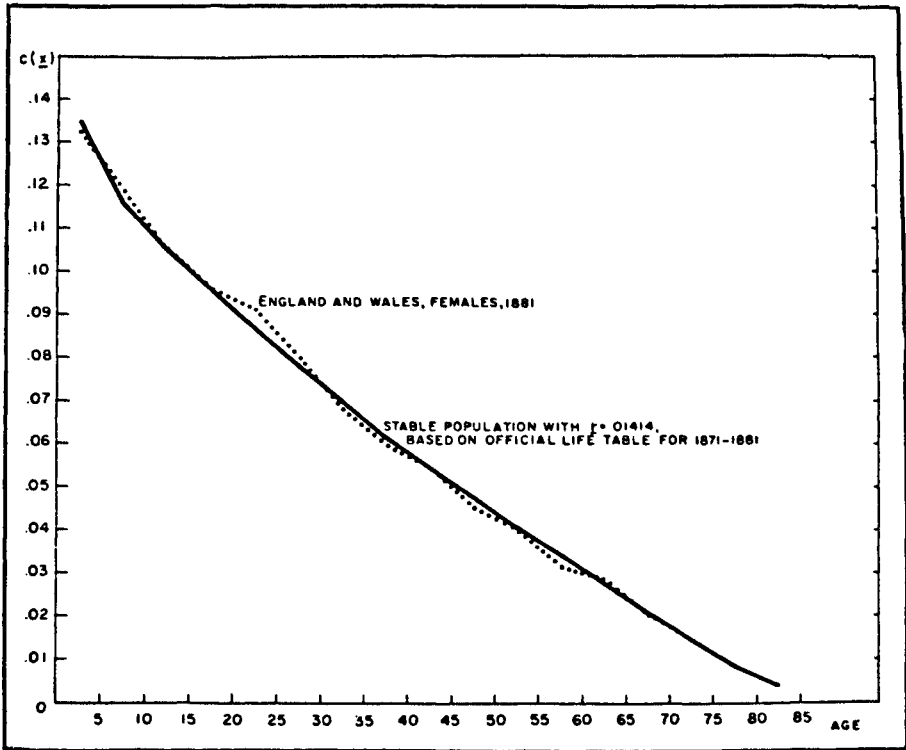


FIGURE 4. Female age distribution in England and Wales, by five-year intervals, as recorded in the census of 1881 (dotted line) and as approximated by a stable population (solid line) constructed on the basis of the intercensal (1871–1881) rate of natural increase and the official English life table for the same period, both for females. Source: Coale and Demeny 1969, p. 13.

The earliest papers of Lotka and Sharpe include a numerical comparison of the calculated stable age structure with an observed age structure in England and Wales. Figure 4 compares the observed proportions of females by age in 1881 with the predicted proportions in a stable population having the death rates and intercensal rate of increase observed between 1871 and 1881 in England and Wales. This population has computed its own dominant eigenvector and acted accordingly. If you are suspicious about how far this example may be generalized, it is only fair to admit that these data were chosen to illustrate agreement between stable and observed age structures, although they are not the only such data. While many current populations have age structures that are not very close to their stable limit, there are enough populations that are nearly stable, particularly among those that are rapidly growing, to make the strong ergodic theorem the basis of very useful procedures for estimating demographic parameters from incomplete data (Coale and Demeny [1969]). For example, if a country has reasonable estimates of an age census, of age-specific death rates, and an overall rate of population growth, the strong ergodic theorem can be used to estimate age-specific fertility rates. There are many other such examples (Bourgeois-Pichat [1968]).

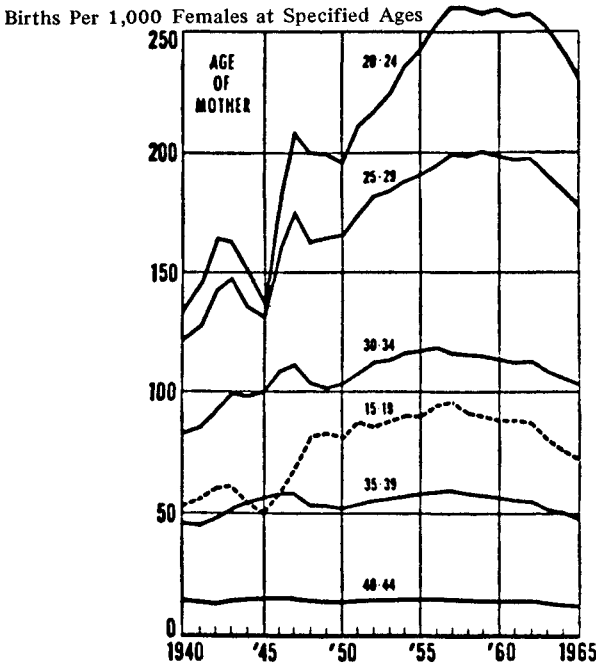


FIGURE 5. United States births per 1,000 females in specified five-year age groups, 1940-1965. Source: Spiegelman 1968, p. 264.

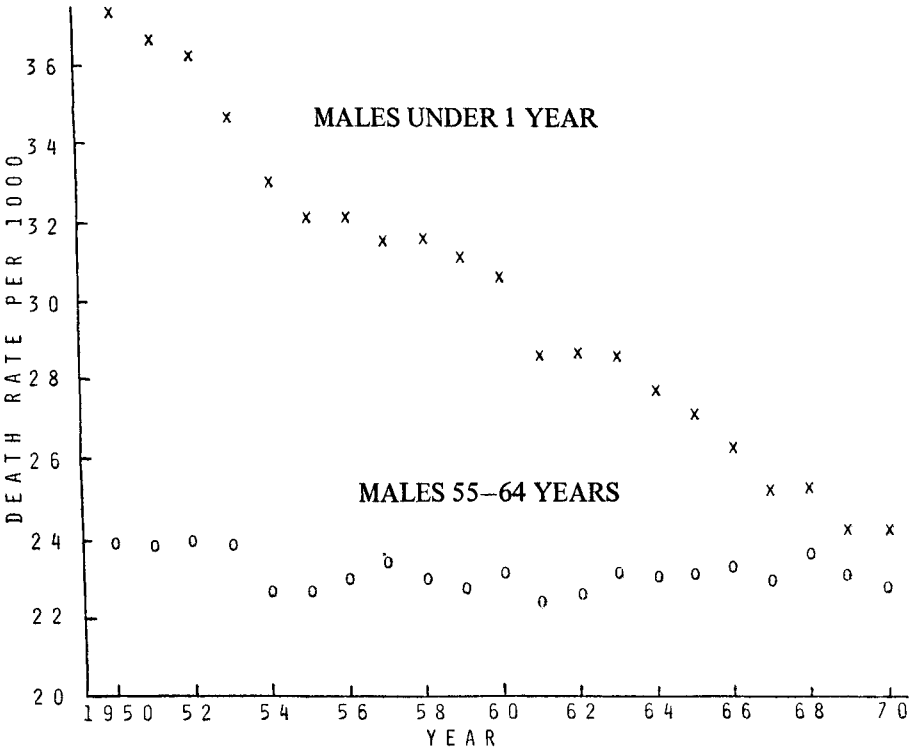


FIGURE 6. United States deaths per 1,000 males under 1 year old (above) and per 1,000 males aged 55 to 64 years (below), 1950-1970. Source of data: U. S. Bureau of the Census, Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition, pt. 1, p. 61.

But populations do not grow exponentially forever. The strong ergodic theorem cannot provide an accurate long term prediction of total population size for the many human populations in which the current stable growth rate λ exceeds 1. Contrary to the assumptions of the strong ergodic theorem, for some populations neither age-specific birth rates (Figure 5) nor age-specific death rates (Figure 6) are constant over time. What can be said about age-structured populations in which vital rates do vary in time?

4. The weak ergodic theorem. In 1957, Coale conjectured that two different initial age censuses subjected to the same sequence of vital rates have age structures that gradually become increasingly like each other, though they may both continue to change in time. In 1961, his student Lopez proved the weak ergodic theorem, using concepts developed by Hajnal for inhomogeneous Markov chains:

If $x(1), x(2), \dots$ are projection matrices (with repetitions possible) from the set X , $Y(0)$ and $Y'(0)$ are two different initial nonzero age censuses, $Y(t) = x(t) \cdots x(1)Y(0)$, $Y'(t) = x(t) \cdots x(1)Y'(0)$, then $\lim_{t \rightarrow \infty} \|y(t) - y'(t)\| = 0$. Thus age structures forget their remote past.

Without going through the details of a proof, one can see why this is so by considering the sequences $\{Y_1(t)\}$ and $\{Y'_1(t)\}$ which approximate the sequences of births in the two populations. In any population, current births are an average of births in previous years, weighted by the proportions surviving and the effective fertility of those who survive. Thus

$$\frac{Y_1(t)}{Y'_1(t)} = \frac{b_1(t)Y_1(t-1) + b_2(t)s_1(t-1)Y_1(t-2) + \dots}{b_1(t)Y'_1(t-1) + b_2(t)s_1(t-1)Y'_1(t-2) + \dots}. \quad (5)$$

Since the same coefficients (which approximate the so-called net maternity function) are used to compute the average in the numerator and denominator of (5), it is not surprising that $Y_1(t)$ and $Y'_1(t)$ eventually become proportional; and then the remaining elements of age censuses $Y(t)$ and $Y'(t)$ must also become proportional.

The weak ergodic theorem makes a science of age structures possible. If in order to explain the current age structure of a population it were necessary to know its prior age structures indefinitely far into the past, the task would be hopeless. The weak ergodic theorem provides assurance that, regardless of the age structure of a population some number of years ago, the vital rates since then completely determine the current age structure. To determine how far into the past it is necessary to know vital rates in order to explain a current age structure is an empirical question. According to numerical experiments with 10×10 projection matrices for women in 5-year age groups, the most recent 15 to 20 matrices (representing 75 to 100 years of vital rates) determine the current age structure for all practical purposes (Kim and Sykes [1976]). These numerical experiments have uncovered empirical regularities which invite theoretical explanation.

Part of the results of Kim and Sykes [1976] may be explained by the recent demonstration (Hajnal [1976], based on earlier results of Birkhoff [1967] and Golubitsky et al. [1975]) that the convergence of age structures is exponential, regardless of the sequence $x(t)$, in the Hilbert projective pseudometric defined

by

$$d(Y(t), Y'(t)) = \ln \frac{\max_i(Y_i(t)/Y'_i(t))}{\min_i(Y_i(t)/Y'_i(t))}$$

for strictly positive vectors $Y(t), Y'(t)$. (Clearly if $Y(t)$ and $Y'(t)$ are proportional then $d(Y(t), Y'(t)) = 0$.) The rate of convergence is given by

$$d(Y(t), Y'(t)) \leq d(Y(0), Y'(0)) \left(\frac{1 - \delta^k}{1 + \delta^k} \right)^{\lfloor t/k \rfloor}$$

Here $\lfloor a \rfloor$ is the greatest integer less than or equal to a , and $\delta > 0$ is the ratio of R , used above to define X , to k .

An immediate consequence of the weak ergodic theorem, which Coale noted in 1970 and many have reproved since then, is that if the sequence of projection matrices is periodic with period T , then so is the sequence of age structures, with period not exceeding T . Some interesting biological parables can be drawn from this simple example (MacArthur [1968]).

Valuable though the weak ergodic theorem be for interpreting the past and the present, it is a weak guide for projections. As Niels Bohr reportedly said (Ulam [1976, p. 286]), "It is very hard to predict, especially the future." Figure 7 shows the official projections of births for the United Kingdom

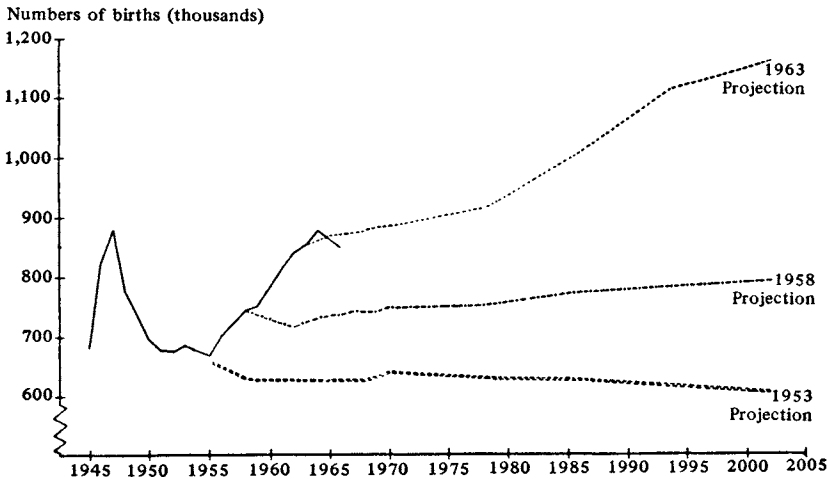


FIGURE 7. England and Wales actual births 1945–1965 (solid line) and births officially projected in 1953, 1958, and 1963 (dashed lines); number in thousands. Source: Cox 1970, p. 438.

prepared in 1953, 1958, and 1963. The startling variation among projections and their deviations from reality suggest that the choice of projection matrices for prediction remains an art. (Dorn [1950]; Hajnal [1955]; Grauman [1967]; Keyfitz [1972a] discuss the problems of population prediction.)

5. Stochastic ergodic theorems. It is worth studying models of age-structured populations with randomly varying vital rates for three reasons: in order to recognize what appears to be random variation in past vital rates (e.g. Figures 5 and 6), in order to improve projections of the future, and in

order to associate with each projection some probability distribution (or confidence interval, in statistical language) to indicate an anticipated range of variation.

The empirical usefulness of the stochastic models we will now describe has not yet been demonstrated. At least these models are formulated so that they are empirically testable.

The exclusion of nondemographic factors from these models is not a denial that such factors are important. It will undoubtedly be essential to incorporate economic, social and technological factors in future models.

The framework of these stochastic ergodic theorems, except for some modifications, is described by Furstenberg and Kesten [1960]. Mark Kac brought Kesten's attention to the model in connection with a physical problem arising at Bell Telephone Laboratories.

As before, $X = \{x^{(i)}\}_{i \in I}$ is a (not necessarily countable) ergodic set (Hajnal [1976]) of projection matrices. $x(t)$ is chosen from X by a random matrix-valued process $x(t, \omega) = x(t)$, $t = 1, 2, \dots$. Here ω is a point in an underlying probability space. $x(1, \omega) = x(1)$, $x(2, \omega) = x(2)$, \dots is one realization or sample path of the process specifying the vital rates. The age censuses at each time t are random vectors specified by $Y(t, \omega) = x(t, \omega) \cdot \dots \cdot x(1, \omega) Y(0, \omega)$. The corresponding age structures are $y(t, \omega) = Y(t, \omega) / \|Y(t, \omega)\|$, $t = 0, 1, 2, \dots$.

Furstenberg and Kesten [1960] assume that the process generating $x(t, \omega)$ is a strictly stationary metrically transitive process. We shall assume that the process is Markovian but not necessarily time-homogeneous. By Markovian, we mean that, if A is a measurable subset of X , then $\mathbf{P}[x(t, \omega) \in A | x(t-1, \omega), x(t-2, \omega), \dots] = \mathbf{P}[x(t, \omega) \in A | x(t-1, \omega)]$. We shall denote this transition probability function by $P_{t-1}(x(t-1), A)$. We shall speak of $Y(t)$, suppressing the ω , as the age census of a population in a Markovian environment. We interpret each matrix $x^{(i)}$ in X as, or as corresponding to, one environment. $x(t, \cdot)$ could be the expected value matrix of a multitype branching process in a Markovian environment (Smith [1968], Smith and Wilkinson [1971], Athreya and Karlin [1971]). Since $y(t+1) = x(t+1)y(t) / \|x(t+1)y(t)\|$ depends on $x(t+1)$ and $y(t)$, and $x(t+1, \omega)$ depends only on $x(t, \omega)$ if $x(t, \cdot)$ is Markovian, it follows that $(x(t+1, \omega), y(t+1, \omega))$ depends only on $(x(t, \omega), y(t, \omega))$. Therefore the bivariate process $z(t, \cdot) = (x(t, \cdot), y(t, \cdot))$ is Markovian if $x(t, \cdot)$ is.

To determine the transition probabilities for $z(t)$, let A , as before, be a measurable set of matrices x in X , and let B be a measurable set of age structures (k -vectors) $y \geq 0$ satisfying $\|y\| = 1$. Then clearly

$$\begin{aligned} \mathbf{P}[x(t+1) \in A, y(t+1) \in B | x(t), y(t)] \\ = \mathbf{P}[x(t+1) \in A \cap \{x: xy(t) \in B\} | x(t)]. \end{aligned}$$

We denote this transition probability function of $z(t)$ by $G_t(x(t), y(t), A, B)$ and observe that the transition probability function on $x(t, \cdot)$ determines G_t in a simple way.

Let $F_t(A, B) = \mathbf{P}[x(t) \in A, y(t) \in B]$. Then

$$F_{t+1}(A, B) = \int_{x \in X} \int_{\substack{y > 0 \\ \|y\|=1}} F_t(dx, dy) G_t(x, y, A, B). \tag{6}$$

In words, the probability that $x(t + 1)$ is in A and $y(t + 1)$ is in B is just the integral over all possible values of x and y at time t of the probability density $F_t(dx, dy)$ of x and y at time t multiplied by the conditional probability G_t of the transition from (x, y) into (A, B) .

If the Markov process on X is suitably ergodic or mixing, so that it forgets its initial distribution as $t \rightarrow \infty$, and if X is an ergodic set, so that long products of operators from X become increasingly close to matrices of rank 1, then the two kinds of forgetting can be spliced together so that as $t \rightarrow \infty$, F_t becomes independent of F_1 .

If one assumes that the Markovian environments are homogeneous, so that $G_t = G$, then $\lim_{t \rightarrow \infty} F_t = F$ where

$$F(A, B) = \int \int F(dx, dy) G(x, y, A, B). \tag{7}$$

This linear integral equation is the fundamental renewal equation for age-structured populations in homogeneous Markovian environments, analogous to the characteristic equation for age-structured populations with fixed vital rates. In cases of practical interest, (7) can be approximated by a large system of linear algebraic equations. A computer can solve these linear equations to give an arbitrarily good approximation to F . A detailed numerical example, with a picture of the resulting F , is given in Cohen [1977b].

When the Markov chain on X is homogeneous, ergodic, and stationary (started at its equilibrium distribution), then the stochastic process governing the vital rates is a special case of the processes studied by Furstenberg and Kesten [1960]. They proved that there exists an almost sure limiting growth rate of total population size $\|Y(t, \omega)\|$ and that the probability distribution of age structure $y(t, \omega)$ approaches a limiting probability distribution. They did not specify how to calculate the almost sure limiting growth rate and the limit probability distribution of y in any concrete cases. We see from (7) that in our special case, the limit law or probability distribution of age structure y is obtained from the bivariate limit law F as the marginal distribution obtained when A is replaced by X .

We now turn to measures of the growth rate of total population size $\|Y(t, \omega)\|$.

Suppose that for each sample path ω , the total population size $\|Y(t, \omega)\|$ ultimately changes exponentially in time with a growth factor $\lambda(\omega)$ which may depend on ω , so that

$$\lim_{t \rightarrow \infty} \|Y(t, \omega)\| / (\lambda(\omega))^t = a(\omega), \quad 0 < a(\omega) < \infty.$$

Furstenberg and Kesten proved that with probability 1 $\lim_{t \rightarrow \infty} t^{-1} \ln \|Y(t, \omega)\|$ exists and is independent of ω ; moreover this limit, which we shall denote by $\ln \lambda$, $\lambda > 0$, almost surely equals $\lim_{t \rightarrow \infty} t^{-1} E_\omega \ln \|Y(t, \omega)\|$. By stationarity of

the process on X ,

$$\ln \lambda = E_{\omega} \ln \frac{\|Y(2, \omega)\|}{\|Y(1, \omega)\|}.$$

In the special case where $x(t, \omega)$ is a Markov chain on X , we have

$$\ln \lambda = \int_{(x,y)} \int_{x'} \ln \left(\frac{\|x'y\|}{\|y\|} \right) \cdot \mathbf{P}[x(t+1) = x' | x(t) = x] F(dx, dy). \quad (8)$$

$\mathbf{P}[x(t+1) = x' | x(t) = x]$ has to be interpreted correctly if X is not countable. Equations (7) and (8) do for age-structured populations in a homogeneous Markovian environment what the Euler-Lotka equation does for age-structured populations with constant vital rates. In cases of practical interest, from a knowledge of the transition function of the Markov chain with state space X , we can compute (with a real computer, not just in principle) F from (7) and then $\ln \lambda$ from (8). When X is finite, $\ln \lambda$ is bounded by

$$-\infty < \sum_{i \in I} \pi_i \ln c_i \leq \ln \lambda \leq \sum_{i \in I} \pi_i \ln c^{(i)} < \infty, \quad (9)$$

where c_i is the smallest of the column sums of $x^{(i)}$, $c^{(i)}$ is the largest of the column sums of $x^{(i)}$, and π_i is the equilibrium probability of $x^{(i)}$ in the regular Markov chain on X (Cohen [1978b]).

An unsolved mathematical problem is to find some nontrivial example in which the almost sure long run growth rate λ can be studied analytically as a function of the members of X and the transition probability function on X .

Those who enjoy historical coincidences may be amused to consider that Furstenberg and Kesten proved their lemmas concerning the contractive properties of positive matrices during 1958–1959 at Princeton University, in the old Fine Hall, former home of the Mathematics Department. At the same time, Alvaro Lopez, working on his doctoral thesis under Ansley Coale, was proving essentially the same lemmas across the street in the University's Office of Population Research. The connection between the work of Furstenberg and Kesten [1960] and that of Lopez [1961] seems not to have been made until 15 years later (Cohen [1976]).

As a further coincidence, I recently learned from Mark Kac of the independent rediscovery by Morgenstern et al. [1978] of special cases of (7) and (8). Their studies of an Ising model in random magnetic fields assume that 2×2 positive matrices $x(t)$ are chosen from X independently and identically distributed.

The almost sure long run growth rate λ is not the only plausible measure of the rate of growth of the population in a Markovian environment (Boyce [1977]; Cohen [1977b], [1978b], submitted). Suppose that the expected total population size at time t , where the expectation is over all sample paths, ultimately changes exponentially with t as t gets large. Then $\lim_{t \rightarrow \infty} \mu^{-t} E_{\omega} \|Y(t, \omega)\| = a$, $0 < a < \infty$, implies

$$\ln \mu = \lim_{t \rightarrow \infty} t^{-1} \ln E_{\omega} \|Y(t, \omega)\|, \quad \mu > 0.$$

Since the logarithm is concave, it is immediate that $\ln \lambda \leq \ln \mu$ with strict inequality in general. In fact when X is finite, the expected total population

size does asymptotically change exponentially, and μ is the spectral radius of a certain nonnegative matrix (Cohen [1977b]).

Suppose again for simplicity that the set X of projection matrices is finite and that the homogeneous Markov chain on X is regular (irreducible and aperiodic). If successive projection matrices are independently and identically distributed, then μ is just the spectral radius of the average of the projection matrices occurring at a given time. Other properties of μ are somewhat less expected.

Suppose one is given the spectral radius λ_i of each $x^{(i)}$, that is, the long run rate of growth λ_i of a population which experiences only the vital rates in $x^{(i)}$. Suppose one is also given the transition matrix of the Markov chain on X . While this information specifies a lower bound on μ , it does not in general specify any upper bound: μ can be arbitrarily large. Thus the average sample path can grow at a rate μ arbitrarily greater than $\max_i \lambda_i$, even though each matrix $x^{(i)}$ of vital rates by itself permits a known rate of growth λ_i .

Now suppose that the elements of the projection matrices $x^{(i)}$ in X are determined but that the transition probability matrix of the Markov chain governing successive projection matrices $x(t)$ is undetermined. Then there exists a transition probability matrix such that the rate of growth μ of the mean population size is arbitrarily close to the largest of the λ_i while the spectral radius of the average of the projection matrices is arbitrarily close to the smallest of the λ_i . The average projection matrix to which the population is subject is, of course, just the sum of the matrices in X weighted by the equilibrium probabilities π_i of the Markov chain on X . Thus sequential dependence of environments can give a growth rate of the mean population size which is near the largest of the growth rates λ_i of any single $x^{(i)}$ even though the average vital rates would suggest a growth rate near $\min_i \lambda_i$, the lowest of the growth rates of any single environment.

Leaving out some of the technical details, we may summarize our major results in a weak stochastic ergodic theorem and a strong stochastic ergodic theorem.

Weak stochastic ergodic theorem: If the sequence of Leslie matrices applied to an age census $Y(0)$ is a sample path of a Markov chain, then the joint process consisting of the current Leslie matrix $x(t)$ and the current age structure vector $y(t)$ is a Markov chain with transition function G_t which we have stated explicitly in terms of the transition function of $x(t)$. If the Leslie matrices are chosen from an ergodic set X of Leslie matrices, and if the Markov chain on X is S -uniformly ergodic in the sense of Griffeath [1975], then the Markov chain $(x(t), y(t))$ is "uniformly weakly ergodic" in the sense that, for every origin of time, for every $\epsilon > 0$, and for every measurable set A of Leslie matrices and every measurable set B of age structures, there exists an integer m_0 such that for all $m \geq m_0$,

$$\begin{aligned} \sup_{(x,y), (x',y')} & |\mathbf{P}[(x(m), y(m)) \in (A, B) | (x(1), y(1)) = (x, y)] \\ & - \mathbf{P}[(x(m), y(m)) \in (A, B) | (x(1), y(1)) = (x', y')]| < \epsilon; \end{aligned}$$

that is, the joint distribution of the current Leslie matrix and current age structure $(x(t), y(t))$ becomes independent of the initial Leslie matrix and

initial age structure after a long time, uniformly with respect to initial conditions.

Strong stochastic ergodic theorem: When the Markov chain on X is homogeneous (when the probabilities of transition from one Leslie matrix to another are constant in time), the joint distribution F_t of the current Leslie matrix and the current age structure $(x(t), y(t))$ approaches a limiting invariant probability distribution F which is the solution of the renewal equation (7). For any Borel function g of $(x(t), y(t))$,

$$\lim_{t \rightarrow \infty} \sum_{k=1}^t g(x(k), y(k)) / t = \int g(x, y) F(dx, dy)$$

almost surely if the integral (over x and y) on the right exists. At last we have an ergodic theorem in the traditional sense! (The details and proofs of the stochastic ergodic theorems up to this point, stated in general operator-theoretic terms without restriction to a matrix representation for members x of X , appear in Cohen [1977a]. The details of the remainder of the strong stochastic ergodic theorem below appear in Cohen [1977b].) In the simplest case, when X contains a finite number of Leslie matrices and the Markov chain on X is homogeneous and regular, the long run rate of growth μ of the expected population size is the dominant eigenvalue of a certain matrix. The long run age structure of the expected population may be calculated from the dominant eigenvector of this matrix.

Lange (in press *b*) reformulates and extends parts of this strong stochastic ergodic theorem.

6. Some applications and extensions. These stochastic models and theorems suggest a scheme for incorporating historical human data into a new method of population projection. Arrange all the age-specific effective fertility and survival coefficients in a projection matrix into a vector. Fit a linear first-order autoregressive scheme to a historically observed sequence of such vectors. Use the estimated parameters and an initial array of vital rates to project a distribution of arrays of future vital rates. Given an initial age structure, this distribution of future vital rates implies a distribution of projected subsequent age structures and population sizes.

The empirical merit of this scheme, or of other possible parametric specifications of the Markovian model, in competition with existing methods of projection, remains to be determined. Similar Markovian and more elaborate autoregressive models are now being applied to age-structured human (Lee [1974], [1975], Saboia [1977]) and even duck populations (Anderson [1975]). These are by no means all the interesting models for age-structured populations which have been proposed (Goodman [1968], Sykes [1969b], Pollard [1973], Ludwig [1974]). The question whether some models are empirically better than others has been neglected, however, as each author tends to promote his own favorite. To evaluate the empirical merit of various population projection techniques, it would be essential to draw on the recent sophistication of some demographers (Henry and Gutierrez [1977]) in using historical data.

On grounds of common sense, it seems likely that populations in stochastic

environments do not grow exponentially forever, either on average or almost surely. It would seem desirable to investigate stochastic age-structured models in which the members of x are nonlinear operators dependent, perhaps, on the most recent age census. Recent writers on deterministic density dependent age-structured models (e.g. Rorres [1976]) are continuing earlier work on the same subject using continuous time and age (Lotka [1939]) or discrete time and age (Leslie [1948]). Almost everything remains to be done in the context of stochastic population models with density dependence.

The stochastic models of age-structured populations described here are identical or similar in form to discrete multiplicative processes in random environments which have applications in the theory of polymer chemistry (Morgenstern et al. [1978]), nuclear reactors, automata, learning, and ecology (Cohen [1978b]). Insight gained into these models is likely to have widespread rewards.

Here is an opportunity to put to work Kingman's [1977] maxim: "... mathematicians should direct their attention to questions to which someone, somewhere, wants to know the answers."

REFERENCES

- David R. Anderson, *Optimal exploitation strategies for an animal population in a Markovian environment: a theory and an example*, Ecology **56** (1975), 1281-1297.
- K. B. Athreya and S. Karlin, *On branching processes with random environments: I. Extinction probabilities: II. Limit theorems*, Ann. Math. Statist. **42** (1971), 1499-1520, 1843-1858.
- Garrett Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloq. Publ., no. 25, Amer. Math. Soc., Providence, R. I., 1967.
- J. Bourgeois-Pichat, *The concept of a stable population: application to the study of populations of countries with incomplete demographic statistics*, United Nations, New York, ST/SOA/SER.A/39, 1968.
- A. L. Bowley, *Births and population of Great Britain*, J. Roy. Econom. Soc. **34** (1924), 188-192.
- Mark S. Boyce, *Population growth with stochastic fluctuations in the life table*, Theoret. Population Biology **12** (1977), 366-373.
- Edwin Cannan, *The probability of a cessation of the growth of population in England and Wales during the next century*, The Economic Journal **5** (1895), 505-515.
- Ansley J. Coale, *How the age distribution of a human population is determined*, Cold Spring Harbor Symposia on Quantitative Biology (ed. K. B. Warren) **22** (1957), 83-89.
- _____, *The use of Fourier analysis to express the relation between time variations in fertility and the time sequence of births in a closed human population*, Demography **7** (1970), 93-120.
- _____, *The growth and structure of human populations*, Princeton Univ. Press, Princeton, N. J., 1972.
- A. J. Coale and Paul Demeny, *Methods of estimating basic demographic measures from incomplete data*, United Nations Manual IV on Methods of Estimating Population, ST/SOA/Ser.A/42, United Nations, New York, 1969.
- Joel E. Cohen, *Ergodicity of age structure in populations with Markovian vital rates. I: Countable states*, J. Amer. Statist. Assoc. **71** (1976), 335-339.
- _____, *Ergodicity of age structure in populations with Markovian vital rates. II: General states*, Advances in Appl. Probability **9** (1977a), 18-37.
- _____, *Ergodicity of age structure in populations with Markovian vital rates, III: Finite-state moments and growth rates; illustration*, Advances in Appl. Probability **9** (1977b), 462-475.
- _____, *Derivatives of the spectral radius as a function of nonnegative matrix elements*, Math. Proc. Cambridge Philos. Soc. **83** (1978a), 183-190.
- _____, *Long-run growth rates of discrete multiplicative processes in Markovian environments*, J. Math. Anal. Appl. (1978b).
- _____, *The cumulative distance from an observed to a stable age structure*, SIAM J. Appl. Math. (to appear).

Joel E. Cohen, *Comparative statics and stochastic dynamics of age-structured populations*, Theoret. Population Biology (submitted).

Peter R. Cox, *Demography*, 4th ed., Cambridge Univ. Press, London and New York, 1970.

Lloyd Demetrius, *The sensitivity of population growth rate to perturbations in the life cycle components*, Math. Biosciences 4 (1969), 129–136.

———, *Demographic parameters and natural selection*, Proc. Nat. Acad. Sci. U.S.A. 71 (1974), 4645–4647.

———, *Adaptedness and fitness*, Amer. Natur. 111 (1977), 1163–1168.

Harold F. Dorn, *Pitfalls in population forecasts and projections*, J. Amer. Statist. Assoc. 45 (1950), 311–334.

H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist. 31 (1960), 457–469.

Martin Golubitsky, Emmett B. Keeler and Michael Rothschild, *Convergence of the age structure: Applications of the projective metric*, Theoret. Population Biology 7 (1975), 84–93.

Leo A. Goodman, *Stochastic models for the population growth of the sexes*, Biometrika 55 (1968), 469–487.

———, *On the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates*, Theoret. Population Biology 2 (1971), 339–354.

John V. Grauman, *Success and failure in population forecasts in the 1950's; a general appraisal*, Proc. World Population Conf., Belgrade, August 30–September 10, 1965, United Nations, New York, 1967.

David Griffeath, *Uniform coupling of non-homogeneous Markov chains*, J. Appl. Probability 12 (1975), 753–762.

John Hajnal, *The prospects for population forecasts*, J. Amer. Statist. Assoc. 50 (1955), 309–322.

———, *On products of nonnegative matrices*, Math. Proc. Cambridge Philos. Soc. 79 (1976), 521–530.

Louis Henry and Hector Gutierrez, *Qualité des prévisions démographiques à court terme. Étude de l'extrapolation de la population totale des départements et villes de France 1821–1975*, Population 32 (1977), 625–647.

Frank Hoppensteadt, *Mathematical theories of populations: demographics, genetics and epidemics*, Society for Industrial and Applied Mathematics, Philadelphia, Penn., 1975.

Tosio Kato, *Perturbation theory for linear operators*, 2nd ed., Springer-Verlag, New York, 1976.

Nathan Keyfitz, *An introduction to the mathematics of population*, Addison-Wesley, Reading, Mass., 1968.

———, *Linkages of intrinsic to age-specific rates*, J. Amer. Statist. Assoc. 66 (1971a), 275–281.

———, *On the momentum of population growth*, Demography 8 (1971b), 71–80.

———, *On future population*, J. Amer. Statist. Assoc. 67 (1972a), 347–363.

———, *Population Waves*, Population Dynamics, T. N. E. Greville (ed.), 1–38, Academic Press, New York, 1972b.

———, *Applied mathematical demography*, Wiley, New York, 1977.

Y. J. Kim and Z. M. Sykes, *An experimental study of weak ergodicity in human populations*, Theoret. Population Biology 10 (1976), 150–172.

J. F. C. Kingman, *Review of Stochastic processes in queueing theory*, by A. A. Borovkov, Bull. Amer. Math. Soc. 83 (1977), 317–318.

Kenneth Lange, *The momentum of a population whose birth rates gradually change to replacement levels*, Math. Biosciences (in press).

———, *On Cohen's stochastic generalization of the strong ergodic theorem of demography*, Advances in Appl. Probability (in press b).

Ronald Demos Lee, *Forecasting births in post-transition populations: stochastic renewal with serially correlated fertility*, J. Amer. Statist. Assoc. 69 (1974), 607–617.

———, *Natural fertility, population cycles and the spectral analysis of births and marriage*, J. Amer. Statist. Assoc. 70 (1975), 295–304.

P. H. Leslie, *Some further notes on the use of matrices in population mathematics*, Biometrika 35 (1948), 213–245.

Alvaro Lopez, *Problems in stable population theory*, Office of Population Research, Princeton, N. J., 1961.

Alfred J. Lotka, *Théorie analytique des associations biologiques. Part II. Analyse démographique avec application particulière à l'espèce humaine*, Actualités Sci. Indust., No. 780, Hermann, Paris, 1939.

Donald Ludwig, *Stochastic population theories*, Lecture Notes in Biomath., vol. 3, Springer-Verlag, New York, 1974.

Robert H. MacArthur, *Selection for life tables in periodic environments*, Amer. Natur. **102** (1968), 381–383.

George W. Mackey, *Ergodic theory and its significance for statistical mechanics and probability theory*, Advances in Math. **12** (1974), 178–268.

Ingo Morgenstern, Kurt Binder, and Artur Baumgärtner, *Statistical mechanics of Ising chains in random magnetic fields*, J. Chem. Phys. **69** (1978), 253–262.

John H. Pollard, *Mathematical models for the growth of human populations*, Cambridge Univ. Press, London and New York, 1973.

Chris Rorres, *Stability of an age specific population with density dependent fertility*, Theoret. Population Biology **10** (1976), 26–46.

J. L. M. Saboia, *Autoregressive integrated moving average (ARIMA) models for birth forecasting*, J. Amer. Statist. Assoc. **72** (1977), 264–270.

Tore Schweder, *The precision of population projections studied by multiple prediction methods*, Demography **8** (1971), 441–450.

Eugene Seneta, *Non-negative matrices*, Allen and Unwin, London, 1973.

David Smith and Nathan Keyfitz, *Mathematical demography: Selected Readings*, Biomathematics, vol. 6, Springer-Verlag, New York, 1977.

Walter L. Smith, *Necessary conditions for almost sure extinction of a branching process with random environment*, Ann. Math. Statist. **39** (1968), 2136–2140.

W. L. Smith and William E. Wilkinson, *Branching processes in Markovian environments*, Duke Math. J. **38** (1971), 749–763.

Mortimer Spiegelman, *Introduction to demography*, rev. ed., Harvard Univ. Press, Cambridge, Mass., 1968.

Zenas M. Sykes, *On discrete stable population theory*, Biometrics **25** (1969a), 285–293.

———, *Some stochastic versions of the matrix model for population dynamics*, J. Amer. Statist. Assoc. **64** (1969b), 111–130.

S. M. Ulam, *Adventures of a mathematician*, Charles Scribner's Sons, New York, 1976.

Kenneth M. Weiss and P. A. Ballonoff (eds.), *Demographic genetics*, Benchmark Papers in Genetics, vol. 3, Dowden, Hutchinson & Ross, Stroudsburg, Penn., 1975.

P. K. Whelpton, *An empirical method of calculating future population*, J. Amer. Statist. Assoc. **31** (1936), 457–473.

DEPARTMENT OF POPULATIONS, ROCKEFELLER UNIVERSITY, NEW YORK, NEW YORK 10021