Mathematical Population Studies, 1992, Vol. 3(4), pp. 289–299
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in the United States of America

OPTIMAL ESTIMATE OF POPULATION AGE STRUCTURE WITH VARYING VITAL RATES

YING-HEN HSIEH

Department of Applied Mathematics
National Chung-Hsing University
Taichung, Taiwan, R.O.C.

February 14, 1990, revised October 6, 1991

Let x and y be age distribution vectors and let $H = A_m A_{m-1} \cdots A_{k+1}$ be the product of population projection matrices A_t of a population from time $t = k$ to m. The optimal estimate vector x^* is defined as the unique positive vector at which

$$\min_{|x|_1=1} \max_{|y|_1=1} \left| \frac{Hy}{|Hy|_1} - x \right|_2$$

is attained. It is useful as an estimate of the current age distribution of a human population under varying vital rates. In this paper, we show how x^* can be computed as the Chebyshev center of a convex set, provided the sequence of Leslie matrices $\{A_n\}_{n=k+1}^{m}$ satisfies the conditions of the "backward" weak ergodic theorem.

Some mathematical aspects of the method, as well as its applicability, will be discussed and numerical simulation will be run to illustrate the results and to compare the method with the traditional estimate of stable population theory, using the right eigenvector associated with the dominate eigenvalue of matrix product H.

KEY WORDS: Leslie matrix, optimal estimate, Chebyshev center, backward weak ergodicity, varying vital rates, age structure.

Communicated by Marc Artzrouni

1. INTRODUCTION

Much work has been done in applying the theory of stable age distribution to make demographic analysis and prediction of human population with stable vital rates. (See, e.g., Coale and Demeny, 1967). However, in cases where the vital statistics of a population vary drastically during any given time period, the estimate of subsequent age distribution by stable population theory falls short of one's satisfaction.

Coale (1957) conjectures that, even under varying vital statistics, the age distribution of a population should become independent of its initial age structure eventually. His conjecture was later proven by Lopez (1961) in the form of the weak ergodic theorem. In its discrete version, the weak ergodic theorem implies the convergence of a sequence of products of Leslie matrices (or population projection matrices) to a sequence of positive matrices with rank 1. The weak ergodic theorem assures us that regardless of the age structure of a population years ago, the vital rates since then completely determine the current age structure. Therefore, to estimate the current age distructure, we need only the vital statistics for the recent past years. Kim and Sykes (1976) showed numerically that 15 to 20 10×10 matrices for females in 5-year age groups (given 75–100 years of vital data) would be sufficient to given good estimate of the current age distribution.

However, the weak ergodic theorem does not in itself provide any specific expression for the row and column of the limiting matrices of rank 1, nor does its various related limiting theorems proven by Golubitsky et al. (1975), Hajnal (1976), Chatterjee and Seneta (1977), among others. Preston and Coale (1982) generalized the results of stable population theory to provide various expressions where one can estimate the age distribution of a population when the vital data are defective or incomplete. More recently, Kim (1987) presented a method in which explicit expression of limiting row and column vectors for the forward and backward weak ergodic theorems were given as functions of the vital rates. In application, when complete census of the past or complete vital statistics of a population is not available, the estimate of current age structure is a practical problem, especially when the population is not stable in the sense that the stable population estimate using the right eigenvector of dominant eigenvalue would yield a significant error. Since population data with no complete census or with insufficient vital statistics frequently occurs in developing countries which are also most susceptible to drastic charges (due to either natural disaster, man-made catastrophe, or drastic social changes), the problem of finding a more accurate estimate is a relevant question in application of mathematical demography.

Hsieh (1982) proposed the method of optimal estimate by defining an optimal vector which serves as an estimate for the age distribution of a population at time $t = m, x_m$, in the sense that it minimizes the maximum possible (Euclidean) error of the estimate that can occur even if the initial age distribution, x_k with $k < m$, is such that x_m is as far away as possible from any estimate for it. The estimate is "optimal" in the sense that when the vital rates of a population is undergoing rapid changes, a worst-case scenario assumption will be more valid than the stable population estimate traditionally employed.

In this work, we will discuss the applicability of the method of optimal estimate by showing that, under the condition that the sequence of Leslie matrices from time $t = k$ to m obeys the hypothesis of the backward weak ergodic theorem, the optimal estimate vector can be obtained by computing the Chebyshev center of a convex set closely related to the product of the sequence of Leslie matrices from $x = k$ to m. We will first state the backward ergodic theorem in Section 2. Section 3 will be devoted to main mathematical results and Section 4 gives a numerical procedure by which the Chebyshev center of a given convex set can be easily computed. Finally, in Section 5 we will give numerical simulations of a population undergoing changes in vital rates to illustrate our results and the applicability of the method. We will use a sequence of varying vital statistics given in Kim (1985) for a population with two age-groups. We will estimate the current age distribution using the optimal estimate to show that in some instances, the optimal estimate *is* more accurate than the stable population estimate.

2. BACKWARD WEAK ERGODIC THEOREM

THEOREM 1 (Backward Weak Ergodicity) *Let $\{A_s\}$, $s = 0, -1, -2, \dots$, be an infinite sequence of $n \times n$ non-negative matrices obeying the following assumptions:*

1. *For each non-positive integer m, there is an integer $m_0 < m$ such that $H^{m,r} > 0$ for all $r \leq m_0$, where $H^{m,r} = A_m A_{m-1} \cdots A_{r+1}$.*
2. *There are positive numbers α and β (independent of s) such that*
   a) $\min^{+}_{i,j} a_{ij}(s) \geq \alpha > 0$
   b) $\max^{+}_{i,j} a_{ij}(s) \leq \beta < \infty$

   for all non-positive integer s, where $\min^{+}_{i,j} a_{ij}(s)$ denotes the smallest of the (strictly) positive components of A_s.

Then, for any non-positive integer m and two distinct, non-negative, non-zero vectors μ and ν, the vector sequences $\{\mu^m_{(k)}\}$ and $\{\nu^m_{(k)}\}$ defined by

$$\mu^m_{(k)} = H^{m,m-k}\mu \qquad \textit{and} \qquad \nu^m_{(k)} = H^{m,m-k}\nu$$

are "asymptotically proportional" in the sense that

$$\lim_{k\to\infty}\left[\frac{\mu^m_{(k)i}}{\nu^m_{(k)i}} - \frac{\mu^m_{(k)j}}{\nu^m_{(k)j}}\right] = 0 \qquad \textit{for all} \quad i,j = 1,\dots,n. \tag{1}$$

Here $\mu^m_{(k)i}$ is the ith component of the vector $\mu^m_{(k)}$.

The theorem is similar to the standard (forward) weak ergodic theorem for finite inhomogeneous Markov chains (e.g. in Seneta, 1973, §4.3) except that the matrix product goes backwards. However, the results from inhomogeneous Markov chain theory can be used since they are essential direction-free (see Chatterjee and Seneta, 1977). Kim (1987) showed that when the time interval for the product matrix $H^{m,r}$, namely m and r, is fixed, the backward and forward results for the age distribution can be consolidated into one equation. Since in any estimate of the age distribution of population, only a finite number of Leslie matrices is available, the results of the backward weak ergodic theorem justifies the use of recently past vital statistics for determination of current age structure.

3.1. Preliminaries

Let us review some basic concepts from Hsieh (1982).

A Chebyshev center of a compact set $S \subset R^n$ with respect to some metric $|\ |$ (which we take here to be the Euclidean, $\|\ \|_2$) is a point $x^* \in S$ such that

$$\max_{y\in S}|x^* - y| = \min_{x\in S}\max_{y\in S}|x - y|, \tag{2}$$

i.e., it is the center of a ball of minimal radius containing S ("a generalized circumsphere"). If S is convex, x^* is obviously unique; and we may write "$\min_{x\in R^n}$"

instead of "$\min_{x \in \mathcal{S}}$" in (2). If $\mathcal{S}$ is a convex polyhedron in R^n (or some subspace), $\mathcal{S} = CH(a_1, \ldots, a_m)$ (where "CH" means "convex hull"), we may find x^* as the point for which

$$\max_i |x^* - a_i| = \min_{x \in \mathcal{S}} \max_i |x - a_i|. \tag{3}$$

The special case of (3) where $\mathcal{S}$ is a simplex is considered in more detail in the following section.

Let $\mathcal{A} = CH(e_1, \ldots, e_n) = \{x \in R^n \mid x \geq 0, \ \|x\|_1 = 1\}$, where $\|x\|_1 = |x_1| + \cdots + |x_n|$. For a given non-negative matrix H, with column vectors h_i $(= He_i)$, $i = 1, \ldots, n$, define a mapping T of the non-negative orthant R^{+n} into $\mathcal{A}$ by

$$Tx = \frac{Hx}{\|Hx\|_1}. \tag{4}$$

Denote by $\mathcal{B}$ the image of $\mathcal{A}$ under the mapping T. Then

$$\mathcal{B} = CH(a_1, \ldots, a_n), \tag{5}$$

where $a_i = Te_i = h_i / \|h_i\|_1$, $i = 1, \ldots, n$. Note that if H is the positive matrix product of population projection matrices $H^{m,r}$ in Theorem 1 satisfying the hypotheses of the theorem, then H is invertible and $\mathcal{B}$ is an $(n-1)$-dimensional simplex with vertices $a_1, \ldots, a_n$.

An optimal vector for H is a vector (point) $x = x^*$ at which

$$\min_{x \in \mathcal{A}} \max_{y \in R^{+m}} |Ty - x| \tag{6}$$

is attained. From the discussion above it is clear that x^* is the Chebyshev center of $\mathcal{B}$.

The proofs of the statements in this section are elementary and are left for the reader, cf. also Hsieh (1982).

3.2. Main Results

Consider the $(n-1)$-dimensional simplex $\mathcal{B} = CH(a_1, \ldots, a_n)$ from Equation (5), and recall that $\mathcal{B} \subset \mathcal{A}$. We shall examine the relation between the Chebyshev center x^* and the circumcenter x^c of $\mathcal{B}$, the latter being the point in $\mathcal{A}$ with equal distances to all the vertices. Let us start out with some definitions.

m_{ij} is the midpoint between a_i and a_j,
$m_{ij} = (a_i + a_j)/2$,

b_{ij} is the hyperplane that bisects $a_i - a_j$ perpendicularly
$b_{ij} = \{x \in R^n \mid (a_i - a_j) \cdot (x - m_{ij}) = 0\}$,

S_{ij} is the halfspace bounded by b_{ij} and containing a_j,
$S_{ij} = \{x \in R^n \mid (a_i - a_j) \cdot (x - m_{ij}) \leq 0\}$,

R_i is the intersection between $\mathcal{A}$ and all the S_{ij}'s, $j = 1, \ldots, n$, $j \neq i$,
$R_i = \{x \in \mathcal{A} \mid (a_i - a_j) \cdot (x - m_{ij}) \leq 0 \text{ for } j = 1, \ldots, n, \ j \neq i\}$.

Note that $b_{ij} = b_{ji} = S_{ij} \cap S_{ji}$. Note also that it follows from the definition of R_i that $x \in R_i$ implies $|x - a_i| = \max_j |x - a_j|$.

Since R_i is convex and closed there is a point $x_{oi} \in R_i$ closest to a_i, and it occurs on the boundary of R_i, i.e., on either [1] one, [2] several or [3] all of the hyperplanes b_{ij}. In case [3] we have $x^* = x^c$.

Examining case [1], let us define a point of smoothness of a closed convex set as a boundary point having a unique supporting hyperplane. If the point x_{oi} is a point of smoothness of $\mathcal{B}$, then $a_i - x_{oi}$ must be perpendicular to the supporting hyperplane at x_{oi}, and in fact this hyperplane must be equal to b_{ij} for exactly one j, because every b_{ij} passing through x_{oi} supports $\mathcal{B}$ there. In other words, we have case [1], and obviously $x_{oi} = m_{ij}$.

THEOREM 2 *Let H be the positive matrix product of population projection matrices $H^{m,r}$ satisfying the hypothesis of Theorem 1, and let $\mathcal{B}$, $\mathcal{R}_i$, m_{ij} be as described above. For any two integers $i, j \in \{1, \dots, n\}$ we have*

$$m_{ij} \in \mathcal{R}_i \Leftrightarrow \angle(a_i a_k a_j) \geq \pi/2 \qquad \text{for all} \quad k \neq i, j, \tag{7}$$

and in that case it follows that $x^* = m_{ij}$.

PROOF. Note that $m_{ij} \in R_i$ implies $m_{ij} \in R_j$. Besides we have

$$\begin{aligned} m_{ij} \in R_i &\Leftrightarrow (a_i - a_k) \cdot (m_{ij} - m_{ik}) \leq 0 \qquad \text{for all} \quad k \neq i, j \\ &\Leftrightarrow (a_i - a_k) \cdot (a_j - a_k) \leq 0 \qquad \text{for all} \quad k \neq i, j \\ &\Leftrightarrow \angle(a_i a_k a_j) \geq \pi/2 \qquad \text{for all} \quad k \neq i, j \\ &\Leftrightarrow |m_{ij} - a_k| \leq |m_{ij} - a_i| = |a_i - a_j|/2 \qquad \text{for all} \quad k \neq i, j. \end{aligned}$$

In other words, for $m_{ij} \in R_i$ we have $|m_{ij} - a_i| = \max_k |m_{ij} - a_k|$, and it follows easily that $x^* = m_{ij}$. □

Note that the theorem applies to a slightly more general situation than [1], namely when the point in R_i closest to a_i is not a point of smoothness, because $\angle(a_i a_k a_j) = \pi/2$ for some k. In this case m_{ij} is the intersection of b_{ij} and b_{ik}, but the theorem is still valid. Hence case [2] is similarly treated.

We shall now state and prove a corollary to Theorem 2 which is indicative of the direction we are heading in dealing with this problem of computing the Chebyshev center x^* of $\mathcal{B}$.

COROLLARY 1 *Suppose the equivalent relation in (7) is satisfied for a pair of integers i, j, with $i \neq j$. If $\angle a_i a_k a_j = \pi/2$ for each $k \neq i, j$, then m_{ij} is the circumcenter and the Chebyshev center of $\mathcal{B}$.*

PROOF. If $\angle a_i a_k a_j = \pi/2$ for each $k \neq i, j$; then

$$\|a_i - m_{ij}\|_2 = \|a_j - m_{ij}\|_2 = \|a_k - m_{ij}\|_2 \qquad \text{for each} \quad k \neq i, j.$$

(i.e., m_{ij} is equidistant to all vertices of $\mathcal{B}$ and therefore is the circumcenter of $\mathcal{B}$.) Since m_{ij} is in $\mathcal{B}$, it is also the Chebyshev center of $\mathcal{B}$. □

Since $\mathcal{B}$ is an $(n-1)$-dimensional simplex, its circumcenter x^c satisfies the following equations:

$$(a_1 - a_i)\cdot x^c = \tfrac{1}{2}(a_1 - a_i)\cdot(a_1 + a_i), \qquad i = 2,\ldots,n \tag{8}$$

and $\|x^c\|_1 = 1$. Because $\{a_1,\ldots,a_n\}$ are linearly independent, the n equations determine a solution x^c uniquely.

Now consider the case when the condition in Theorem 2 does not hold; i.e., for each pair of integers $i,j \in \{1,\ldots,n\}$, there is an integer $k \neq i,j$ such that $\angle a_i a_k a_j < \pi/2$. We shall first show that if the circumcenter of $\mathcal{B}$ is in $\mathcal{B}$, then it is also the Chebyshev center.

Let y_i denote the point in $\mathcal{R}_i$ closest to a_i, i.e., for $i = 1,\ldots,n$, $y_i \in \mathcal{R}_i$ and

$$\|y_i - a_i\|_2 = \min_{x\in\mathcal{R}_i} \|x - a_i\|_2. \tag{9}$$

Clearly, the circumcenter x^c of $\mathcal{B}$ is in $\mathcal{R}_i$ for each i and in the case in which x^c is also in $\mathcal{B}$, $x^c = y_i$ for each i. (Since x^c is the unique point in $\mathcal{B}$ that is equidistant to all vertices of $\mathcal{B}$, it must minimize the distance to a_i over a set that is no closer to a_i than any other vertices of $\mathcal{B}$ and has a non-empty intersection with $\mathcal{B}$, namely, the set $\mathcal{R}_i$.)

THEOREM 3 *If x^c, the circumcenter of $\mathcal{B}$, lies in $\mathcal{B}$, then it is also the Chebyshev center of $\mathcal{B}$.*

PROOF. To show x^c is the Chebyshev center of $\mathcal{B}$, let us suppose that x^c is not the Chebyshev center, i.e., there exists a point $y \in \mathcal{B}$, with $y \neq x^c$, that minimizes $\max_i \|x - a_i\|_2$ over $\mathcal{B}$. We then have $y \in \mathcal{R}_i$ for some i and

$$\|y - a_i\|_2 = \min_{x\in\mathcal{R}_i} \|x - a_i\|_2, \tag{10}$$

since y must minimize over $\mathcal{R}_i$ as well. But we know $x^c = y_i$ for each i, in view of (9) and (10), x^c must equal y and we have a contradiction. □

Now we shall state a theorem for the case when x^c is outside $\mathcal{B}$, the proof of which also requires Theorem 3.

THEOREM 4 *If x^c, the circumcenter of $\mathcal{B}$, lies outside $\mathcal{B}$, then x^*, the Chebyshev center of $\mathcal{B}$, is the (unique) point in $\mathcal{B}$ closest to x^c.*

PROOF. Since $\mathcal{B}$ is convex, we know that $x^\#$, the point in $\mathcal{B}$ closest to x^c, is unique. Consider the hyperplane, $\mathcal{H}$, that contains $x^\#$ and is perpendicular to $x^c - x^\#$. Clearly, we have

$$x \in \mathcal{H} \Rightarrow (x^c - x^\#)\cdot x = (x^c - x^\#)\cdot x^\#.$$

Since $x^\#$ is the point in $\mathcal{B}$ closest to x^c, it follows that $\mathcal{B}$ is separated from x^c by $\mathcal{H}$; that is,

$$(x^c - x^\#)\cdot x \le (x^c - x^\#)\cdot x^\# \qquad \text{for all} \quad x \in \mathcal{B}.$$

Let us define a finite list of m integers $\{M_1,\ldots,M_m\} \subset \{1,\ldots,n\}$, with $m < n$, such that

$$(x^c - x^\#)\cdot a_i = (x^c - x^\#)\cdot x^\# \qquad \text{for} \quad i = M_1,\ldots,M_m,$$

Renumbering this set of m integers by $\{1,\dots,m\}$ henceforth to avoid double subscripts, $\{a_1,\cdots,a_m\}$ is the set of vertices of $\mathcal{B}$ that are contained in $\mathcal{H}$.

We know that $m \geq 1$ from the fact that $\mathcal{B}$ is convex. If $m = 1$, then $\mathcal{H}$ supports $\mathcal{B}$ at a_1 only and $x^\# = a_1$. But $x^\#$ is the unique point in $\mathcal{B}$ closest to x and

$$\|x^c - a_1\|_2 = \|x^c - a_i\|_2 \qquad \text{for} \quad i = 1,\dots,n.$$

Thus we have a contradiction and $m \geq 2$.

Therefore, we know that

$$\mathcal{H} \cap \mathcal{B} = CH\{a_1,\dots,a_m\}, \qquad 2 \leq m < n.$$

$\mathcal{H}$ can not contain all n vertices since $\{a_1,\dots,a_n\}$ are linearly independent are $\mathcal{H}$ is $(n-2)$-dimensional.

Let $\mathcal{C}$ denote the $(m-1)$-simplex $\mathcal{H} \cap \mathcal{B}$, then $x^\# \in \mathcal{B}$ and $x^\# \in \mathcal{H}$ imply $x^\# \in \mathcal{C}$. We have

$$\|x - x^c\|_2^2 = \|x - x^\#\|_2^2 + \|x^\# - x^c\|_2^2 \qquad \text{for all} \quad x \in \mathcal{C}.$$

Thus, for any two integers $i, j \in \{1,\dots,m\}$,

$$\|x^c - a_i\|_2 = \|x - a_j\|_2, \quad \text{we also have} \quad \|x^\# - a_i\|_2 = \|x^\# - a_j\|_2$$

and $x^\#$ is the circumcenter of $\mathcal{C}$. Therefore the Chebyshev center of $\mathcal{C}$ by Theorem 3.

It remains to be shown that $x^\#$ is the Chebyshev center of $\mathcal{B}$. We know that for each $k \notin \{1,\dots,m\}$ and $i \in \{1,\dots,m\}$,

$$\|x^c - a_k\|_2 = \|x^c - a_i\|_2. \tag{11}$$

But $k \notin \{1,\dots,m\}$ implies that, for $i \in \{1,\dots,m\}$,

$$(x^c - x^\#)\cdot(a_k - x^\#) < 0 \qquad \text{and} \qquad (x^c - x^\#)\cdot(a_i - x^\#) = 0. \tag{12}$$

Formulae (11) and (12) together imply that

$$\|x^\# - a_k\|_2 < \|x^\# - a_j\|_2$$

for each $k \notin \{1,\dots,m\}$ and $i \in \{1,\dots,m\}$. Therefore, we have

$$\|x^\# - a_i\|_2 = \max_{j\in\{1,\dots,n\}} \|x^\# - a_j\|_2 \qquad \text{for each} \quad i \in \{1,\dots,m\}. \tag{13}$$

For each $x \in \mathcal{B}/\mathcal{C}$ and $j \in \{1,\dots,m\}$, there exists $y_x \in \mathcal{H}$ such that

$$\|x - a_j\|_2^2 = \|x - y_x\|_2^2 + \|y_x - a_j\|_2^2 > \|y_x - a_j\|_2^2 \tag{14}$$

Since we have

$$\|x^\# - a_i\|_2 = \min_{x\in\mathcal{B}} \max_{j\in\{1,\dots,n\}} \|x - a_j\|_2 \qquad \text{for} \quad i = 1,\dots,m,$$

for each $y_x \in \mathcal{H}$, there is an integer $i \in \{1,\dots,m\}$ such that

$$\|x^\# - a_i\|_2 < \|y_x - a_i\|_2. \tag{15}$$

(14) and (15) together imply that for each $x \in \mathcal{B}/\mathcal{C}$,

$$\|x^\# - a_i\|_2 < \|y_x - a_i\|_2 < \|x - a_i\|_2$$

for some $i \in \{1,\ldots,m\}$. Since $x^{\#}$ is the Chebyshev center of C we have, in fact, for all $x \in \mathcal{B}$,

$$\|x^{\#} - a_i\|_2 \leq \|x - a_i\|_2 \qquad \text{for some} \quad i \in \{1,\ldots,m\}, \tag{16}$$

and this together with (14) give us

$$\|x^{\#} - a_i\|_2 = \max_{j\in\{1,\ldots,n\}} \|x^{\#} - a_j\|_2 \leq \max_{j\in\{1,\ldots,n\}} \|x - a_j\|_2 \qquad \text{for all} \quad x \in \mathcal{B}$$

or

$$\|x^{\#} - a_i\|_2 = \min_{x\in\mathcal{B}} \max_{j\in\{1,\ldots,n\}} \|x - a_j\|_2. \qquad \square$$

4. DESCRIPTION OF NUMERICAL METHOD

The results of Theorems 3 and 4 yield the following procedure for computing the optimal vector x^* for a given $n \times n$ invertible, non-negative matrix H:

1. Compute the set $\{a_1,\ldots,a_n\}$ where

$$a_i = \frac{He_i}{\|He_i\|_1} \qquad \text{for} \quad i = 1,\ldots,n$$

   and e_i is the basis vector with components δ_{ij}.
2. Compute the circumcenter, x^c, of

$$\mathcal{B} = CH\{a_1,\ldots,a_n\}$$

   using the $n-1$ equations in (8) and $\|x^c\|_1 = 1$.
3. If $x^c \in \mathcal{B}$, then x^c is the Chebyshev center of $\mathcal{B}$ and thus the optimal vector for H.
4. If $x^c \notin \mathcal{B}$, then from the proof of Theorem 4 it follows that the optimal vector is the circumcenter of a simplex whose vertices a subset of $\{a_1,\ldots,a_n\}$.

The computation of step 4 can be cumbersome when n is large, but in applications of Leslie matrices, when the number of age-groups is usually ten or less, the method is direct and fast. In the next section, we will give numerical examples of population estimate using the optimal estimate and compare it with the traditional method of stable population estimate using the eigenvector corresponding to the dominant eigenvalue of the matrix product, H.

5. NUMERICAL EXAMPLES

In this section, we use the vital data of 15 countries used in Kim (1985). Table 1, reproduced from the Appendix Table in Kim (1985), consists of the vital rates of the first 15 countries in Keyfitz and Flieger (1968) in which the countries are ordered alphabetically by continents. We choose this set of data for reason of simplicity since the populations are divided into two age-groups only, the first consists of individuals of age 0–24, the second of age 25–49. Assuming a population following the given sequence of vital rates in the list, we have 15 2×2 population projection matrices. a_i, b_i denote the fertility rates of the 1st and 2nd age-group, respectively, in the ith

TABLE 1
List of Countries and Elements of the Population Projection Matrix
(Reproduced from Appendix Table, Kim, 1985)

i	Country	a_i	b_i	c_i
1	Algeria, 1965	1.332	1.138	0.924
2	Cameroon (West), 1964	0.819	0.628	0.757
3	Madagascar, 1966	0.860	0.864	0.701
4	Mauritius[a], 1966	1.099	0.887	0.932
5	Reunion, 1963	1.375	1.227	0.927
6	Seychelles, 1960	1.359	0.985	0.955
7	South Africa, 1961 (colored)	0.226	0.983	0.912
8	South Africa, 1961 (white)	0.756	0.518	0.972
9	Tunisia, 1960	1.174	1.027	0.919
10	Canada[b], 1966–2968	0.578	0.383	0.981
11	Costa Rica, 1966	1.371	1.119	0.961
12	Dominican Republic, 1966	1.103	0.907	0.950
13	El Salvador, 1961	1.336	1.061	0.919
14	Greenland, 1960	1.450	1.190	0.942
15	Grenada, 1961	1.323	1.035	0.947

[a]Excluding dependencies
[b]Excluding Newfoundland

TABLE 2
Estimates for Population Obeying the Vital Rates of Table 1

x_0 (initial age distribution)	0.600000	0.400000
x_{15} ("current" age distribution)	0.657070	0.342930
x^* (optimal estimate)	0.657253	0.342747
x_s (stable estimate)	0.656559	0.343441
$\lvert x_{15} - x^* \rvert_2$	0.000260	
$\lvert x_{15} - x_s \rvert_2$	0.000722	

time period, and p_i is the survival rate of the first age-group of ith time period to the second age-group in the next time period. Since $n = 2$, step 4 in the numerical procedure described in the previous section can be easily computed.

Table 2 gives the first example. The initial age distribution vector is $\mathbf{x}_0 = (0.6, 0.4)$. Assuming the population follows the vital rates in Table 1, $\mathbf{x}_{15} = (0.6571, 0.3429)$ is the "current" age distribution. We also compute the optimal estimate x^* using the vital rates of the last five time period ($i = 11$ to 15), i.e., $H = A_{15} \cdots A_{11}$. Thus we have a population with (deterministically) varying vital statistics for 15 time periods (or 375 years), then we use the vital rates of the last 125 years to obtain an estimate for the current age distribution ($\mathbf{x}_{15}$). We also compute the stable estimate for H which is the (normalized) right eigenvector corresponding to the dominant eigenvalue of H. The Euclidean error of both estimate are also computed. As one can see, the error of the optimal estimate is much smaller than that of the stable estimate.

TABLE 3
Estimates for 10 Stimulated Populations Using Permutations of the List of Vital Data in Table 1

Population	$\|x_{15} - x^*\|_2$	$\|x_{15} - x_s\|_2$
1	0.000078(*)	0.000234
2	0.000078(*)	0.000170
3	0.000090(*)	0.000245
4	0.000080(*)	0.000195
5	0.000079(*)	0.000245
6	0.000072(*)	0.000200
7	0.000178	0.000047(*)
8	0.000061(*)	0.000093
9	0.000111(*)	0.000194
10	0.000290	0.000213(*)

Moreover, we would like to know if such accuracy is merely a coincidence of numbers. To investigate further, we use permutations of the list in Table 1. For each of the ten simulated populations in Table 3, we reorder the list in Table 1 so that one of the first 10 sets of vital data comes at the end of the list, i.e., the i-th simulated population follows the permutation order of $(1, 2, \ldots, i-1, i+1, \ldots, 10, 11, \ldots, 15, i)$. For example, the 2nd simulated population follows the order $(1, 3, 4, \ldots, 14, 15, 2)$. The purpose of this simulation is to see that when a "random" vital data is added to the end of the five time period, will the optimal estimate still be more accurate? In Table 3, $H = A_i A_{15} \cdots A_{11}$ for the ith population, hence we use 150 years of vital statistics to estimate the current age-distribution. Once again, $x_0 = (0.6, 0.4)$ and we compute the Euclidean error of the two estimate compared to $\mathbf{x}_{15}$.

To save space, we only give the errors of the two estimates in Table 3. In each case, the estimate with the smaller error is distinguished by an asterisk symbol $(*)$. One can see that the optimal estimate is better in eight out of the ten cases. Furthermore, in seven of the ten cases, the Euclidean error of the optimal estimate is less than 10^{-4}. In all cases, the stable estimate is expected to provide a good estimate since it was conjectured by Kim and Sykes (1976) that $2n$ or $3n$ Leslie matrices, where n is the dimension of the matrix, are needed for the population to converge. But that the optimal estimate gives an even better estimate in most of the cases is in itself significant. Hence, when faced with varying vital statistics and few vital data to work with, the optimal estimate is a viable option to the traditional stable population theory.

ACKNOWLEDGMENT

This research was supported partially by grant (NSC79-0208-M005-05) from the National Science Council of The Republic of China. The author is grateful to the anonymous referees for helpful suggestions which have resulted in significant improvement of an earlier version of this paper.

REFERENCES

Chatterjee, S., and Seneta, E. (1977) Toward consensus: Some convergence theorems on repeated averaging. *Journal of Applied Probability* **14**: 89–97.

Coale, A. J. (1957) How the age distribution of a human population is determined. *Cold Spring Harbor Symposia on Quantitative Biology* **22**: 83–89. New York: Long Island Biological Association.

Coale, A. J., and Demeny, P. (1967) Methods of estimating basic demographic measures from incomplete data. *Population Studies* No. 42. New York: United Nations.

Golubitsky, M., Keeler, E. B., and Rothschild, M. (1975) Convergence of age structure: Applications of the projective metrics. *Theoretical Population Biology* **7**: 84–93.

Hajnal, J. (1976) On products of nonnegative matrices. *Mathematical Proceedings of the Cambridge Philosophical Society* **79**: 521–530.

Hsieh, Y. H. (1982) Demographic prediction under varying vital statistics, Ph.D. Thesis, Carnegie-Mellon University.

Hsieh Y.H. (1985) On the use of optimal estimate in demography. *Journal of Science and Engineering, NCHU* **22**: 211–226.

Keyfitz, N. and Flieger, W. (1968) *World Population: An Analysis of Vital Data*. Chicago: University of Chicago Press.

Kim, Y. J. (1985) On the dynamics of populations with two age groups. *Demography* **22**(3): 455–468.

Kim, Y. J. (1987) Dynamics of populations with changing rates: generalization of the stable population theory. *Theoretical Population Biology* **31**: 306–322.

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

Lopez, A. (1961) *Problems in Stable Population Theory*. Princeton, NJ: Office of Population Research, Princeton University.

Preston, S. H., and Coale, A. (1982) Age structure, growth, attrition, and accession: A new synthesis. *Population Index* **48**: 217–259.

Seneta, E. (1973) *Non-negative Matrices*. New York: Wiley.

ABSTRACTS IN FRENCH
RÉSUMÉS EN FRANÇAIS

Croissance démographique avec familles de dimensions variables (Richard H. Norden).

Le modèle déterministe de Sharpe-Lotka est adaptée à une situation où les filles "héritent" de leur mère une fécondité définie par une mesure bivariée. Des transformations de Laplace sont alors utilisées pour résoudre l'équation intégrale double qui résulte de ces hypothèses. Des résultats asymptotiques sont obtenus sur la distribution de la fécondité dans la population stable correspondant à ce modèle. En particulier, la population totale contient des sous-populations qui se caractérisent par différents niveaux de fécondité.

Tables de survie multi-états et modèles linéaires (Richard D. Gill).

L'auteur fait le point sur quelques méthodes statistiques utilisables en analyse événementielle ("event history analysis") et dans l'étude des tables de survie multi-états. Pour l'étude de certains modèles semi-markoviens et dans les modèles de débilitation à hétérogénéité non-observable, R. Gill s'attache particulièrement aux relations entre vraisemblance partielle d'une part, et des méthodes du maximum de vraisemblance d'autre part.

Une variante de la formule de Heligman-Pollard à neuf paramètres (Anastasia Kostaki).

L'auteur propose une variante à neuf paramètres de la fonction de Heligman-Pollard utilisée pour modéliser les taux de mortalité par âge. Des ajustements empiriques sur cinq pays européens montrent que cette nouvelle fonction ajuste mieux les taux de mortalité que la formula classique à huit paramètres.

Estimation optimale de la structure pas âge dans une population soumise à des taux de fécondité et de mortalité variables (Ying-Hen Hsieh).

Si $H = A_m A_{m-1} \dots A_{k+1}$ est un produit de matrices de Leslie, l'auteur definit le vecteur x^* comme étant l'unique vecteur x^* pour lequel le minmax $\min\max |Hy/|Hy|_1 - x|_2$ est atteint (le min est pour $|x|_1 = 1$, le max est pour $|y|_1 = 1$). Le vecteur x^* est un estimateur de la structure par âge, et peut être consideré comme le centre au sens de Chebyshev d'un certain ensemble convexe. Les aspects mathématiques de la question sont étudiés et des exemples numériques sont donnés.

ABSTRACTS OF RELATED ARTICLES

Short abstracts of recent articles in mathematical demography that have appeared elsewhere in the literature are given below. This review covers systematically the following journals: *Theoretical Population Biology*, *Mathematical Biosciences*, *Journal of Mathematical Biology*, *Journal of Theoretical Biology*, *Bulletin of Mathematical Biology*, *Demography*, *Population Studies*, *Population*, *European Journal of Population*, *Genus*, and *Journal of the American Statistical Association*. Other journals, particularly mathematical journals, are surveyed, but in a less systematic fashion.

Bonneuil, N. (1990) Contextual and structural factors in fertility behaviours. *Population* [English Edition], **2**: 69–89.

The author studies the topological structure of historical time series of net reproduction rates by using the theory of chaotic dynamics. Two periodic limit cycles emerge from some of those series.

Fisch, O. (1991) A structural approach to the form of the population density function. *Geographical Analysis* **23**,3: 261–275.

In this paper the author explores "a comprehensive structural modeling approach that extracts analytical density functions answering questions raised by recent empirical studies."

Foster, A. (1991) Are cohort mortality rates autocorrelated? *Demography* **28**,4: 619–637.

The author investigates whether heterogeneity in individual frailty leads to autocorrelation in cohort mortality rates. A simple model for the covariance of cohort mortality rates is used to construct a procedure to estimate the extent of heterogeneity in a population. The procedure is applied to French data.

Hougaard, P., Harvald, B., and Holm, N. V. (1992) Measuring the similarities between the lifetimes of adult Danish twins born between 1881–1930. *Journal of the American Statistical Association* **87**,417: 17–37.

The survival of 8,985 like-sex twins born in Denmark is studied by means of models for bivariate data. The degree of dependence is the focus of the study and is assumed to be generated by a common unobserved risk level.

Martin, D., and Bracken I. (1991) Techniques for modelling population related raster databases. *Environment and Planning A* **23**,7: 1,069–1,075.

Surface models of population information are proposed and lead to databases that can be used for a variety of spatial analyses. The techniques are illustrated with data from the united Kingdom.

Pfeffermann, D. (1991) Estimation and seasonal adjustment of population means using data from repeated surveys. *Journal of Business and Economic Statistics* **9**,2: 63–175.

Estimation and seasonal adjustments of population means are proposed on the basis of rotating panel surveys. The mean is decomposed into a trend-level component and a seasonal component. Numerical illustrations are given.

Rogers, A. (1992) Heterogeneity and selection in multistate population analysis. *Demography* **29**,1: 31–38.

Some aspects of the selection effects of heterogeneity in multistate populations are investigated (e.g., the impact on death rates of recurrent events among interacting populations).

Sobel, M. E., and Arminger, G. (1992) Modeling household fertility decisions: A non-linear simultaneous probit model. *Journal of the American Statistical Association* **87**,417: 38–47.

This article describes a new method for modeling household fertility decisions that takes into account how spouses influence each other. A trivariate distribution of the wife's and the husband's desire for children, and of subsequent fertility is used.

van Imhoff, A. (1992) A general characterization of consistency algorithms in multidimensional demographic projection models. *Population Studies* **46**,1: 159–169.

This article describes algorithms to solve consistency problems in multi-state demographic projections. The specification of the objective function used in the algorithm leads to a solution that can be interpreted as a generaliztion of the harmonic-mean approach.

Yadava, K. N. S., and Singh, R. B. (1991) A probability model for the distribution of the number of of migrants at the household level. *Genus* **XLVII**,1–2: 49–62.

This article describes a probability model that describes the distribution of total number of immigrants from a household. The model satisfactorily fits several sets of observed distributions in rural areas.