MODELS OF POPULATION DYNAMICS

Nathan Keyfitz, Population Research and Training Center
University of Chicago

SUMMARY

A population changing under the operation of a given set of mortality and fertility rates may be represented mathematically in at least three different ways. Each of these shows, among other things, how the age distribution changes in a trajectory towards ultimate stability. If the trajectory for the nine age groups from 0 to 45, say for females, is seen as a point moving in nine dimensional space, then all three methods may be interpreted as showing the reduction of the space from nine dimensions to three. The characteristic mortality and fertility pattern of real human populations is such that even at the outset of the projection nearly the whole movement is embraced in the three dimensions corresponding to the first three latent roots of the projection matrix. Just as factor analysis rotates a cluster of points in order to view it in as small a space as possible, this portion of the analysis of populations rotates a rigid twisted wire with the same purpose.

Partial fractions

The number of births which will occur during her lifetime to a girl child now born, under a given regime of mortality and survivorship, may be expressed in the form of a series of probabilities

$$f_1 = 0,\ f_2 = 0,\ f_3 = \frac{5\ell_{15}}{\ell_o} m_{15},$$
$$\ldots f_9 = \frac{5\ell_{45}}{\ell_o} m_{45}. \qquad (1)$$

The ℓ's are the life table number living; the m's are the chance of having a child at the age specified. Translation into the more usual five-year age groups will be easily carried out later in the calculation; for the moment m_{15} may be thought of as the age-specific birth rate for the age group $12\frac{1}{2}$ to $17\frac{1}{2}$. If u_t is the number of births at time t to this girl or her descendants, and including as u_o this initial birth, then

$$u_o = 1;\ u_1 = 0;\ u_2 = 0;$$
$$u_3 = u_o f_3 + u_1 f_2 + u_2 f_1 + u_3 f_o = u_o f_3; \qquad (2)$$
$$u_4 = u_o f_4 + u_1 f_3 + u_2 f_2 + u_3 f_1 + u_4 f_o$$
$$= u_o f_4 + u_1 f_3;$$

and so on for as long as the regime of fertility and mortality persists.[1] If now we multiply the members of the set (2) by 1, s, s^2, s^3, etc., and add up the set as so multiplied, we have

$$u_o + u_1 s + u_2 s^2 + \ldots$$
$$= 1 + (u_o f_1 + u_1 f_o)s +$$
$$(u_o f_2 + u_1 f_1 + u_2 f_o)s^2 + \ldots,$$

an equation which may be represented by

$$U(s) = 1 + U(s)F(s), \qquad (3)$$

U(s) being the generating function of the u's and F(s) of the f's. The solution of (3) for U(s) is

$$U(s) = \frac{1}{1 - F(s)} \qquad (4)$$

In order to base calculations on (4) we may expand it in partial fractions, a technique recommended by Feller[2] which has become popular in recent years. If $s_1, s_2 \ldots s_9$ are the roots of the ninth degree polynomial equation

$$F(s) = f_o + f_1 s + f_2 s^2 + \ldots = 1,$$

then the expansion of (4) in partial fractions gives for u_t the coefficient of s^t, assuming the roots distinct,

$$u_t = \frac{1}{s_1^{t+1}F'(s_1)} + \frac{1}{s_2^{t+1}F'(s_2)} + \ldots + \frac{1}{s_9^{t+1}F'(s_9)}, \qquad (5)$$

where $F'(s_i)$ is the derivative of F(s) evaluated at $s = s_i$. But the s_i are approximated by $1/\lambda_i$, the reciprocals of the latent roots of the projection matrix. Knowing the λ_i we can calculate (5) by

$$u_t = \sum_{i=1}^{9} \frac{\lambda_i^{9+t}}{3\ell_{15}m_{15}\lambda_i^6 + \ldots + 9\ell_{45}m_{45}}, \qquad (6)$$

which is a convenient form if the latent roots and their powers are at hand. It is found that after a very brief initial period, at most 20 or 25 years, the omission of all but the first three latent roots (when the roots are arranged in order of absolute value) makes little difference to the calculation; after about 100 years the dominant root suffices for a description. Between 20 and 100 years from the starting point

[1] A.J. Lotka, "Application of Recurrent Series in Renewal Theory," Annals of Mathematical Statistics, XIX (1948), 190-206.

[2] W. Feller, An Introduction to Probability Theory, Vol. I (2nd. ed.; New York: John Wiley & Sons, 1950), pp. 257-261.

considerable departures from stability are caused by the second and third roots, which are harmonic functions with period equal to one generation, gradually attenuated as the projection continues.

To extend the procedure to a more usual age distribution than one girl child just born, suppose now the initial population of an actual country to be K_o, K_5, K_{10}, etc., the number of persons in the population, but think of them as centered at exactly age 0, exactly age 5, etc. This concentration at points five years apart simplifies the formulae, and can be readily adapted to calculation at the final stage. We make the unit of the calculation K_o; in the new units we have $k'_o = K_o/K_o = 1$; $k'_5 = K_5/K_o$, etc. The stable population distribution which corresponds to this--ℓ_o, $\ell_5 e^{-5r}$, $\ell_{10}e^{-10r}$, $\ell_{15}e^{-15r}$--may be similarly divided by the number at age zero, which is the same as assuming that the radix of the life table is 1. Finally we find the departures of the k' from the stable population, and write the departures as <u>k</u>. Thus

$$k_o = 1 - 1 = 0;\ k_1 = k'_1 - \ell_5 e^{-5r};$$
$$k_2 = k'_2 - \ell_{10}e^{-10r},\ \text{etc.} \tag{7}$$

A new generating function may be defined which will give the numbers of children to be expected from the population born before time $t = 0$, insofar as it departs from the stable age distribution:

$$B(s) = b_o + b_1 s + b_2 s^2 + b_3 s^3 + \ldots + b_8 s^8 \tag{8}$$

in which

$$b_o = \frac{k_{15}}{\ell_{15}}\ell_{15}m_{15} + \frac{k_{20}}{\ell_{20}}\ell_{20}m_{20} + \ldots + \frac{k_{45}}{\ell_{45}}\ell_{45}m_{45}$$

$$b_1 = \frac{k_{10}}{\ell_{10}}\ell_{15}m_{15} + \frac{k_{15}}{\ell_{15}}\ell_{20}m_{20} + \ldots + \frac{k_{40}}{\ell_{40}}\ell_{45}m_{45} \tag{9}$$

$$b_2 = \frac{k_5}{\ell_5}\ell_{15}m_{15} + \frac{k_{10}}{\ell_{10}}\ell_{20}m_{20} + \ldots + \frac{k_{35}}{\ell_{35}}\ell_{45}m_{45}$$

. . . .

$$b_8 = \frac{k_5}{\ell_5}\ell_{45}m_{45};\ b_9 = 0$$

The set (9) may be expressed more compactly:

$$b_n = \sum_{i=1}^{9-n} \frac{k_{5i}}{\ell_{5i}}\ell_{5(i+n)}m_{5(i+n)} \tag{10}$$

where $n = 0, 1, 2, \ldots 8$.

On these definitions the relation of the generating functions is

$$U(s) = 1 + U(s)F(s) + B(s) \tag{11}$$

and the solution

$$U(s) = \frac{1 + B(s)}{1 - F(s)}, \tag{12}$$

which in partial fractions becomes

$$U(s) = \frac{1 + B(s_1)}{-F'(s_1)(s-s_1)} + \frac{1 + B(s_2)}{-F'(s_2)(s-s_2)} + \ldots + \frac{1 + B(s_9)}{-F'(s_9)(s-s_9)} \tag{13}$$

The coefficient of s^t in a form suitable for computation is

$$u_t = \sum_{i=1}^{3} \frac{[(1+b_o)\lambda_i^8 + b_1\lambda_i^7 + \ldots + b_8]\lambda_i^{t+1}}{(3\ell_{15}m_{15}\lambda_i^6 + \ldots 9\ell_{45}m_{45})} \tag{14}$$

and evaluation for the first three roots is all that will be necessary. The only additional consideration for computation is that one calculates the b'_o, b'_1, etc. for the usual age groups 0-4, 5-9, etc., and then interpolates back half a period or $2\frac{1}{2}$ years to the points 0, 5, etc. required by (14): $b_o = \frac{1}{2}b'_{0-4}$; $b_1 = \frac{1}{2}(b'_{0-4} + b'_{5-9})$, etc., and similarly for the $\ell_{15}m_{15}$ in the denominator.

<u>The Matrix</u>

A population projection through a five-year period may be represented as the premultiplication of the distribution of ages, treated as a vertical vector $\{K'_o\}$, by a suitably constructed matrix M. The form of M is

$$M = \begin{bmatrix} 0 & 0 & m_{15} & \cdot & \cdot & m_{40} & m_{45} \\ \frac{{}_5L_5}{{}_5L_o} & 0 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & \frac{{}_5L_{10}}{{}_5L_5} & 0 & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \frac{{}_5L_{40}}{{}_5L_{35}} & 0 \end{bmatrix}, \tag{15}$$

and the projection operation is then represented by $M\{K'_0\} = \{K'_1\}$. The properties of this way of representing the changes in a population were investigated in detail by Leslie.[1]

With M as a starting point we define the latent roots as the values of λ which satisfy the determinantal equation $|M - \lambda I| = 0$, and which may be called $\lambda_1, \lambda_2, \ldots \lambda_9$, when arranged in order of their absolute values (see Table I). Corresponding to each of the roots is

TABLE I. LATENT ROOTS FOR THREE COUNTRIES, FEMALES, 1961

	Latent root	Absolute value
	England and Wales, 1961	
λ_1	1.0503	1.0503
λ_2, λ_3	.3593 ± .7727i	.8521
λ_4, λ_5	-.3902 ± .4338i	.5835
λ_6, λ_7	-.0096 ± .5113i	.5114
λ_8, λ_9	-.4846 ± .1454i	.5059
	United States, 1961	
λ_1	1.1081	1.1081
λ_2, λ_3	.3043 ± .7931i	.8495
λ_4, λ_5	-.4109 ± .3895i	.5662
λ_6, λ_7	.0264 ± .5421i	.5427
λ_8, λ_9	-.4739 ± .1617i	.5007
	Hungary, 1961	
λ_1	.9763	.9763
λ_2, λ_3	.2663 ± .7004i	.7493
λ_4, λ_5	-.3676 ± .3780i	.5273
λ_6, λ_7	.0748 ± .4882i	.4939
λ_8, λ_9	-.4616 ± .1721i	.4926

a modal vertical vector $\{K_i\}$ satisfying

$$M\{K_i\} = \lambda_i\{K_i\}, \quad i = 1,2, \ldots 9. \tag{16}$$

There are also 9 horizontal vectors $[H_i]$ satisfying

$$[H_i]M = \lambda_i[H_i], \quad i = 1,2, \ldots 9. \tag{16a}$$

[1]P.H. Leslie, "On the Use of Matrices in Certain Population Mathematics," Biometrika, XXXIII(1945), 183-212; "Some Further Notes on the Use of Matrices in Population Mathematics," Ibid., XXXV (1948), 213-245.

The first of the stable vertical vectors, $\{K_1\}$, is well known as the usual stable population. The first of the stable horizontal vectors, $[H_1]$, was shown by R.A. Fisher to represent in a certain sense the reproductive value of a woman from age 0 to 45.[2] For our purpose the important property is that of orthogonality:

$$\begin{aligned} &[H_i]\{K_i\} \neq 0, \quad i = 1,2, \ldots 9; \\ &[H_i]\{K_j\} = 0, \quad i \neq j = 1,2, \ldots 9. \end{aligned} \tag{17}$$

This enables us to evaluate the c_i in the equation

$$\{K'\} = c_1\{K_1\} + c_2\{K_2\} + \ldots c_9\{K_9\}, \tag{18}$$

since premultiplying both sides of (18) by $[H_i]$ and transposing gives

$$c_i = \frac{[H_i]\{K'\}}{[H_i]\{K_i\}}, \tag{19}$$

the right hand side of (19) being the ratio of two scalar quantities, which except for c_1 will be complex. As in the partial fractions approach, it will turn out that only the first three terms of the expansion (18) are needed in practical work. This arises from the relative absolute values of the latent roots, shown for three countries in Table I.

The use of (18) to expand $\{K'\}$, an actual age distribution, in terms of the unprimed $\{K_i\}$ which are stable vectors, appears from the premultiplication of (18) by the powers of M. By virtue of (16), applied t times, the projection of the actual population to time t is

$$M^t\{K'\} = c_1\lambda_1^t\{K_1\} + c_2\lambda_2^t\{K_2\} + c_3\lambda_3^t\{K_3\}, \tag{20}$$

where we stop with the third root. This result may be thought of as providing a kind of graduation for the trajectory of the population, age by age, as long as the life table and fertility rates hold. It is nine separate scalar equations, one for each element of the 1 × 9 vector. The first of these, representing the top element corresponding to age 0-4, would be practically identical with the value for u_t given in (14).

On the other hand, if we start from (14) we could work towards (20) by the consideration that the population at time t aged 35, say, would be that in (14) at time t - 7 (the unit of time here is five years) as multiplied by the appropriate survival factor. The fitting to United

[2]R.A. Fisher, The Genetical Theory of Natural Selection (2nd rev. ed.; New York: Dover Publications, 1958), pp. 27-30.

TABLE II. ANALYSIS OF UNITED STATES FEMALES, 1961, BY MATRIX APPROACH

(000's)

Age	Actual Population 1961 $\{K'\}$	Components of Fitted Stable Populations: Dominant root $c_1\{K_1\}$	Principal pair of complex roots $c_2\{K_2\} + c_3\{K_3\}$	Other roots plus error $c_4\{K_4\} + \dots + c_9\{K_9\}$
0-4	10,122	10,022	74	26
5-9	9,424	9,003	601	-180
10-14	8,784	8,114	610	60
15-19	6,781	7,308	-317	-210
20-24	5,725	6,576	-1,076	225
25-29	5,489	5,912	-494	71
30-34	5,965	5,308	1,106	-449
35-39	6,394	4,755	1,603	36
40-44	6,045	4,245	-168	1,968

15-year Projection on 1961 Fertility and Mortality Schedules

Age	Population Projected to 1976 $M^3\{K'\}$	Components of Fitted Stable Populations: Dominant root $c_1\lambda_1^3\{K_1\}$	Principal pair of complex roots $c_2\lambda_2^3\{K_2\} + c_3\lambda_3^3\{K_3\}$	Other roots plus error $c_4\lambda_4^3\{K_4\} + \dots + c_9\lambda_9^3\{K_9\}$
0-4	13,782	13,633	202	-53
5-9	11,975	12,248	-238	-35
10-14	10,595	11,038	-479	36
15-19	10,044	9,942	-75	177
20-24	9,364	8,945	597	-178
25-29	8,707	8,043	605	59
30-34	6,701	7,221	-313	-207
35-39	5,632	6,469	-1,090	253
40-44	5,361	5,774	-483	70

30-year Projection on 1961 Fertility and Mortality Schedules

Age	Population Projected to 1990 $M^6\{K'\}$	Components of Fitted Stable Populations: Dominant root $c_1\lambda_1^6\{K_1\}$	Principal pair of complex roots $c_2\lambda_2^6\{K_2\} + c_3\lambda_3^6\{K_3\}$	Other roots plus error $c_4\lambda_4^6\{K_4\} + \dots + c_9\lambda_9^6\{K_9\}$
0-4	18,739	18,547	-192	24
5-9	16,734	16,662	34	38
10-14	15,294	15,017	294	-17
15-19	13,676	13,525	201	-50
20-24	11,899	12,170	-236	-35
25-29	10,502	10,942	-574	134
30-34	9,925	9,824	-73	174
35-39	9,212	8,801	588	-177
40-44	8,504	7,855	590	59

States females for 1961, along with the projection over 15 and over 30 years, is shown as Table II. By the end of 30 years the effect of the fourth to ninth roots has become small.

The Integral Equation

The integral equation, being better known,[1] may be treated more briefly than the two preceding methods. With a continuous formulation B(t), the number of births at time _t_, is

$$B(t) = \int_0^{\beta} B(t-a)\ell(a)m(a)da, \qquad (21)$$

where $\ell(a)$ is the number surviving to age _a_ on the same life table of radix unity. If one tries for the solution

$$B(t) = Ae^{rt},$$

(21) is converted into an equation for _r_

$$\int_0^{\beta} e^{-ra}\ell(a)m(a)da - 1 = 0. \qquad (22)$$

Clearly (22) has the same role here that $F(s) - 1 = 0$ has in the partial fraction formulation, and that $|M - \lambda I| = 0$ has for the matrix. Unlike the previous approaches, however, there are here an infinite number of solutions, rather than nine, and here only one can be real, while previously only one could be positive but real negative roots were possible. While (22) could be solved by fitting moments to the net maternity function, and solving a quadratic or higher polynomial, the point made by Coale[2] that a direct iterative solution not dependent on moments is preferable applies to the complex as well as to the real root. In terms of the real root ρ and the series of complex roots r_1, r_2, etc. the solution to (21) is expressible as a series

$$B(t) = A_o e^{\rho t} + A_1 e^{r_1 t} + A_2 e^{r_2 t} + \dots \qquad (23)$$

Since (21) is linear in B(t), the sum of the solutions, each with an arbitrary weight, is also a solution.

Just as before the s_i, the λ_i, and the $\{K_i\}$ represented the pattern of fertility and mortality and had nothing to do with the actual age distribution, so this is true of the ρ and the r's in (23). The A's, on the other hand, correspond, like the b's of the partial fractions and the c's of the fitting to the stable vectors, to the matching to initial conditions--the age distribution of the real population that we are attempting to analyze. If G(t) is the births at time _t_ to that portion of the population itself born before t = 0, which starts the system off, so to speak, the coefficient in the term $A_i e^{r_i t}$ is

$$A_i = \frac{\int_0^{\beta} e^{-r_i t} G(t)dt}{\int_0^{\beta} a e^{-r_i a}\ell(a)m(a)da}, \qquad (24)$$

where again $\ell(a)$ is the number living at age _a_ in the life table of radix 1. The denominator of (24) is the derivative with respect to _r_, evaluated at the root r_i, of the quantity on the left hand side of (22), just as the denominator of (5) and (6) is the derivative of $1 - F(s)$ evaluated at s_i.

* * * * *

When the three methods for which computing procedures are given above are applied to the usual five-year age groups they turn up different finite approximations. The numerical discrepancies could be made smaller by the methods of finite differences. Some work has been done on the theoretical reconciliation of the matrix and recurrence approaches of equations (3) and (21) above, especially by Feller,[3] and Lopez[4] has helped to relate the matrix formulation to the integral equation, as well as the recurrence equation to both.

There are more important issues for demographers than reconciliation of numerical differences or bridging of the theories which underlie them. They would like to see the methods used to study and classify actual populations; analysis of the two sexes treated together rather than each in abstraction from the other; extension to the recognition of parities and, in general, rates specific for variables other than age and sex; the use of the methods to treat sets of data representing real cohorts, rather

[1] A.J. Lotka, _Theorie Analytique des Associations Biologiques_. Part II (Paris: Hermann, 1939); A.Lopez, _Problems in Stable Population Theory_ (Princeton University Press, 1961); E.C. Rhodes, "Population Mathematics I, II, and III," _Journal of the Royal Statistical Society_, CIII (1940), 61ff, 218ff, and 362ff; A.J. Coale, "How the Age Distribution of a Human Population is Determined," _Cold Spring Harbor Symposia on Quantitative Biology_, XXII (1957), 83-89.

[2] A.J. Coale, "A New Method for Tabulating Lotka's r," _Population Studies_, XI (1957), 92ff.

[3] W. Feller, "On the Integral Equation of Renewal Theory," _Annals of Mathematical Statistics_, XII (1941), 243-267.

[4] A.Lopez, _Op. cit._

than the period or cross-section data applied in the foregoing.[1] A function that would generate probabilities of various numbers of prospective population, rather than the expected values generated by all three of the above methods, would escape the disadvantage of the deterministic model pointed out by Feller exactly a quarter of a century ago.[2] Each of the above directions of extension of the work here summarized presents difficulties, but none of these are insuperable.

[1] Norman B. Ryder, "The Process of Demographic Translation," Demography, I (1964), 74-82.

[2] W. Feller, "Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung," Acta Biotheoretica, V (1939), 11-40.