An Enquiry into the Two Basic Assumptions of Multistate Increment-Decrement Life Tables

Fernando Rajulton
Stan Wijewickrema

Vrije Universiteit Brussels

IPD-Working Paper 1985-8

ORDERING INFORMATION

Extra copies of this working paper are available from :

Mr. E. Vanden Balck
Secretary/Librarian
Interuniversity Programme in Demography
c/o Centrum voor Sociologie
Vrije Universiteit Brussel
Pleinlaan 2
B-1050 Brussels
BELGIUM

at the cost of US $ 5.00 or BF 250. Please attach check or international money order.

ORDER FORM

Name ....
Address ....
....
....
....
....

wishes to receive ...... copies of Working Paper Nr..........
and includes payment of US $ .... or BF........

ABSTRACT

Multistate methodology hitherto in use is based on two basic assumptions of population homogeneity and Markovian behaviour. Efforts have been made at relaxing them, mostly by introducing heterogeneous categories of the population under study. From the fact that these two assumptions conceptually overlap to a certain extent, the relaxing of the population homogeneity has been implicitly taken to mean relaxing the Markovian assumption as well.

This paper enquires into the nature of the Markovian condition from the point of view of Stochastic Processes and brings into focus an important point: In order to make the multistate model more realistic, one has necessarily to take into account the distinction between the two concepts of population homogeneity and the Markovian condition and consequently distinguish between the implications of relaxing the one as opposed to the other. To this end, the population is heterogenized through a simple categorization with each category following its own Markov process. And it is shown that the sum of the categorized state transition probability matrices, when weighted by status-wise population proportions, yields the probability matrix of the total population. Two complementary procedures are given to illustrate that the sum of Markov processes could yield another Markov process.

Finally, a suggestion is made concerning the lines along which the distinction between the two basic assumptions of population homogeneity and the Markovian condition should be viewed.

Recent innovations of multistate analysis in demography introduced notably by Rogers and his colleagues (Rogers, 1975; Willekens & Rogers, 1978; Land and Rogers, 1982) have made possible the use of increment-decrement life tables. These increment-decrement tables, when compared with traditional single decrement tables and their straightforward extensions to multiple decrement tables, are seen to augment the power of demographic analysis, in as much as they enable a simultaneous grasp of a totality in its parts and of parts in relation to their totality.

A marital status increment-decrement table, for example, follows a cohort of persons (of a specified sex [1]) from birth, when all persons are found in the never-married state, through various nuptiality-related states (as the members of the cohort progressively age) right up to the disappearance of the cohort through death. In an increment-decrement table, ${}_{iy}l_j(x)$ replaces the familiar l(x) function of the classical single decrement life table, and stands for the number of members of the cohort in state j at age x who were in state i at a prior age y, both x and y being expressed in exact ages. Thus, both the process of aging and that of transiting among the different nuptiality-related states are simultaneously kept in view through the instrumentality of ${}_{iy}l_j(x)$. This, in turn, permits the computation of a number of summary measures which are the multistate analogues of classical life table measures; for example, the life expectation in the never-married state. Further, computations peculiar to the multistate nature of the system are also possible: for example, the life expectation in the presently married state for persons aged x can be computed either

1. for all persons alive at age x irrespective of their marital status at a prior age y – this gives rise to a measure which is referred to in the literature as a "population-based" measure; or

2. for persons alive at age x given a specific state at a prior age y (e.g. in the widowed state) – this leads to a "status-based" measure.

The analytic power of these multistate demographic models, and their weakness too, lies in their basic assumptions regarding the homogeneity of individuals and the Markovian nature of the processes concerned. The homogeneity assumption ignores any differences which may exist among individuals who happen to be in a given state at the same time. The assumption of Markovian behaviour ignores the effects of previous history and treats all episodes in a given state as if the population exposed to risk were homogeneous in this respect. The assumption of the Markovian nature of the processes involved is sometimes interpreted – <u>ita</u> Keyfitz – as summarizing the totality of the assumptions made in the Markov-based multistate methodology: ".......the independence of moves from one another, the constancy of the probability of moving over a time period and the homogeneity of individuals in a given state at a given moment: in short, the Markovian condition" (Keyfitz, 1980) [2].

Multistate methodology arrives at estimations of the $_{iy}l_j(x)$ described above through prior computations of the relevant state transition probabilities $P_{ij}(x, x+n)$, which are conditional probabilities of being found in state j at the end of the age interval x+n given that the individual was in state i at the beginning of the age interval x. The methodology leading to the computation of these state transition probabilities is schematically presented in Flow Chart 1. The assumption of Markovian behaviour, though explicitly mentioned in the Flow Chart only in the step involved in the computation of the probability matrix, is however ingrained in the very calculation of rates or intensities of transition (Ledent, 1980). In particular, it is the assumption of Markovian behaviour which enables an easy extension of the concepts of the ordinary life table (constructed in the context of single decrements) to multistate increment-decrement life tables; and this, Flow chart 1: The procedure adopted for finding the state transition probabilities in a Markov-based model

observed(period) data → Annual age-specific (x to x+n) rates of transition from state i to state j

↓

$$m_{ij}(x) = \frac{\text{Annual numbers of events from } i \text{ to } j \text{ in the age interval } (x, x+n)}{\text{Mid-population in } i \text{ in } (x, x+n)}$$

↓

Formation of the rate matrix $\underset{\sim}{M}(x)$ containing $m_{ij}(x)$ of varying i and j

↓

Computation of the probability matrix $\underset{\sim}{P}(x)$ whose elements give the probability of surviving in j at age (x+n) for persons who were in i at age x

Assumptions :

(a) Observed rates equal multistate life table rates (implying that the population is stationary)

(b) Transitions are uniformly distributed over age interval (x, x+n)

(c) The Markovian assumption that $P_{ij}(x)$ is independent of any history prior to age x

↓

Computation of ${}_{ix}l_j(x+n)$ using an arbitrary diagonal matrix (giving $l_i(0)$ for all i) as the state specific radix for age 0

↓

Summary measures computed from ${}_{ix}l_j(x+n)$

merely by substituting appropriate vectors and matrices for scalars. Ledent, in his 1980 article, clearly brings out the parallelism between a single state life table and multistate life tables.

Though the power of demographic analysis has been enhanced by these increment-decrement tables, the two basic assumptions of population homogeneity and Markovian behaviour are so unrealistic that efforts have been made at relaxing them in one way or another (Ledent, 1981; Rogers & Philipov, 1981; Kitsul & Philipov, 1981). This has been done mostly by considering the population under study to be no longer homogeneous but heterogeneous (for example, with respect to place of residence, place of birth, sex, etc.), any attempt at relaxing the Markovian assumption being usually beset with well-nigh insurmountable difficulties, not only with respect to the availability of the required data but also with respect to the intractability of the methodology involved. Moreover, arising from the fact that these two assumptions conceptually overlap to a certain extent, the relaxing of the population homogeneity assumption has been implicitly taken to mean relaxing the Markovian assumption as well.

An enquiry into the nature of the Markovian condition looked at from the point of view of Stochastic Processes, however, reveals that relaxing the Markovian condition requires much more than the introduction of a simple categorization which allows for some sort of heterogeneity in the population under study. In effect, it requires a distribution of transitions which is other than the simple exponential distribution invariably implied by the Markovian condition [3]. This paper intends to bring into focus the following important point: viz in order to make the multistate model more realistic, one has necessarily to take into account the distinction between the two concepts of population homogeneity and the Markovian condition, and consequently distinguish between the implications of relaxing the one as opposed to the other.

To this end, the state transition probability matrix, around which our discussion centres, is introduced in the first section, along with the corresponding basic mathematical concepts drawn from matrix algebra and the theory of Stochastic Processes. Next, heterogeneity is introduced through a simple categorization of the population and it is shown that the sum of the state transition probability matrices of each category duly weighted by their respective population proportions yields the probability matrix of the total population. This leads one to reconsider the general "belief" that a sum of Markovian processes is non-Markovian - at least in the context of multistate marital status analysis. That the sum of Markov processes could yield (and does so in the case treated here) another Markov process is shown in two complementary ways in the next two sections. Finally, a suggestion is made concerning the lines along which the distinction between the two basic assumptions of population homogeneity and the Markovian condition should be viewed.

SECTION I

INTRODUCING THE PROBABILITY MATRIX

Consider the possible transitions among the marital states which are schematically presented in the following Fig. 1.

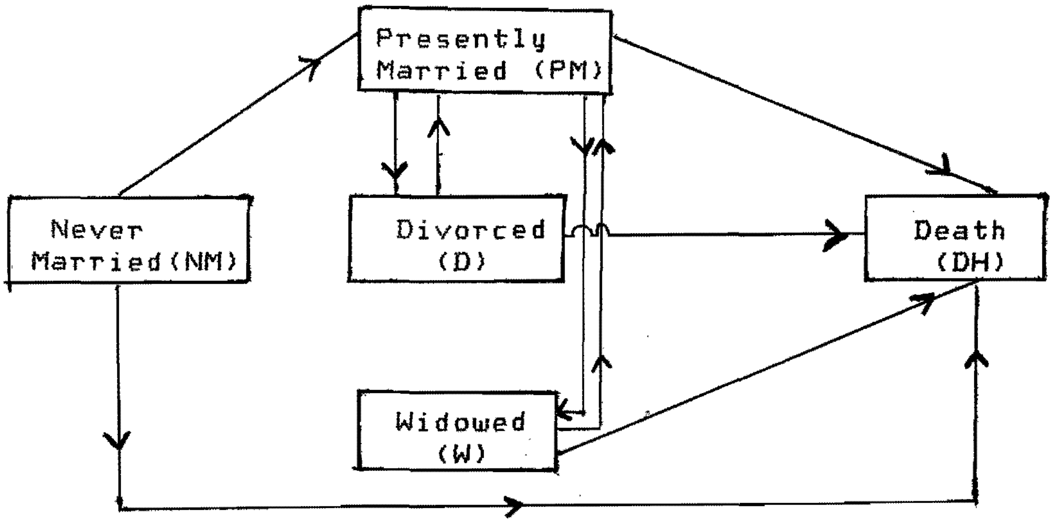

Fig. 1. Schematic representation of transitions among marital states

The states NM, PM, W and D (alternatively denoted by the integers 1, 2, 3 and 4 respectively) are called transient states (because transitions are possible from these states), and the state of death (DH - alternatively denoted by the integer 5) is called an absorbing or trapping state (because no transition is possible from this state). A transition from, say, state i to state j, if possible, is called a "real" transition; and a transition from state i to itself is called a "virtual" transition. Real transitions are further classified into (a) "direct" transitions involving a direct move from state i to state j and (b) "indirect" transitions involving a move from state i to state j through some other state k. While a direct transition to DH is possible from any transient state, direct transitions between certain marital states considered here are not possible: for example, from NM to W (which is realizable only through PM). Further, note that no transition to the NM state can be made from any other transient state. (This is so in contrast to the "unmarried" or "single" state sometimes used in the analysis of transitions among marital states in certain studies- for example, Land and Schoen, 1982).

The probability matrix in a multistate system admits all the above transitions: real and virtual, direct and indirect. This is so because the probabilities concerned are state probabilities of the sort described earlier. Thus, with four transient states and one absorbing state, a probability matrix of transitions from i to j would have 5x5 dimensions. Let the rows in the matrix denote the states of origin and the columns the states of destination. For example, consider the following 5x5 matrix for the age interval (27,28) for Belgian women in the year 1981. [The status-wise population figures given in the sixth column will be used later in the present text: note that these population figures serve as estimations of the number of person-years of exposure to risk in the calculation of the rates $m_{ij}(x)$ (cf. Flow Chart 1) which lead to the estimation of $\underset{\sim}{P}$].

$$\underset{\sim}{P} = \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\begin{bmatrix} \overset{NM}{.898228} & \overset{PM}{.100179} & \overset{W}{.000043} & \overset{D}{.000624} & \overset{DH}{.000926} \\ 0 & .986350 & .000844 & .012376 & .000430 \\ 0 & .047445 & .949065 & .000296 & .003195 \\ 0 & .147137 & .000063 & .850917 & .001884 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{array}{r} \text{POP.} \\ 10501 \\ 59609 \\ 306 \\ 1998 \\ \\ \end{array}$$

The (1,2)th element indicates the probability (=.100179) that a woman who was found in the NM state at the beginning of exact age 27 will be found in the PM state at exact age 28. Similarly, (2,4)th element gives the probability (=.012376) that a woman who was found in the PM at exact age 27 will be found in the D state at exact age 28, etc. It is to be emphasized that these probabilities are not the probabilities of <u>making</u> a transition from one state to another - before being <u>found</u> in a particular state at the end of the age interval under consideration, multiple transitions might have been made. For example, the (1,4)th element in the probability matrix denotes the probability that a woman who was in the NM at age 27 would be found in the D at age 28 - which implies that the woman would have moved to the PM within the interval under considertion.

Certain relevant notions from matrix algebra and the theory of Stochastic processes are now presented, as they will be used in the sections which follow. Let $P_{ij}(x, x+n)$ denote the (i, j)th element of the matrix $\underset{\sim}{P}(x, x+n)$. To simplify the notation used we shall hereafter write $\underset{\sim}{P}(x)$ (or even $\underset{\sim}{P}$) instead of $\underset{\sim}{P}(x, x+n)$, on the supposition that the age interval (n) is one unit(year) unless otherwise specified. The following special features of a Markov-based transition probability matrix should be noted.

(1) In an n-state system, the nxn transition probabilities that describe a Markov process are conveniently described by an nxn transition probability matrix $\underset{\sim}{P}(x) = [P_{ij}(x)]$ whose elements cannot lie outside the range (0, 1) and whose rows sum to one. Such a matrix is called a <u>stochastic matrix.</u> Since the rows sum to one, only n(n-1) parameters are necessary to specify the probabilistic behaviour of an n-state Markovian process.

(2) Every stochastic matrix, and hence every transition probability matrix related to a Markov process, has <u>at least</u> one eigen value equal to one. This is so because every Markov process has at least one associated starting vector $\{p_0\}$ which satisfies the equation $\{p_0\}=\{p_0\}.\underset{\sim}{P}$.

(3) A Markovian process with only one trapping state has only one eigen value equal to unity. The remaining (n-1) eigen values - there are altogether n eigen values corresponding to the n states - are all less than one.

Since there is only one absorbing state (i.e. death) associated with the $\underset{\sim}{P}$ matrix introduced above, there will only be one eigen value equal to unity. This eigen value will correspond to the state of death. For the sake of further analysis, however, one can neglect this trapping state and consider, without loss of generality, an adjusted 4x4 matrix $\underset{\sim}{P}'(x)$ whose elements are made up of $P'_{ij} = P_{ij}/(1-P_{i5})$. It is obviously seen that $\sum_j P'_{ij} = 1$, and hence, $\underset{\sim}{P}'$ is also a stochastic matrix. Note that the relationships obtaining among the elements of the probability matrix $\underset{\sim}{P}$ are still retained in the adjusted matrix $\underset{\sim}{P}'$, which has the format given by:

$$\underset{\sim}{P}' = \begin{array}{c} \begin{array}{cccc} \text{NM} & \text{PM} & \text{W} & \text{D} \end{array} \\ \begin{bmatrix} .899061 & .100272 & .000043 & .000625 \\ 0 & .986774 & .000844 & .012381 \\ 0 & .047597 & .952107 & .000297 \\ 0 & .147415 & .000063 & .852523 \end{bmatrix} \end{array}$$

The eigen values of this matrix $\underset{\sim}{P}'$ are found to be .899061 1.000000, .840057 and .951348 corresponding to the NM, PM, W and D states respectively [4]. Let the right and left eigen vectors corresponding to these eigen values be now put in the form of matrices denoted by $\underset{\sim}{T}$ and $\underset{\sim}{T}^{-1}$ respectively as follows:

$$\underset{\sim}{T} = \begin{bmatrix} 1.000000 & .241722 & .125460 & -.059277 \\ 0 & .241722 & -.079709 & -.031480 \\ 0 & .241727 & .031361 & 1.990720 \\ 0 & .241723 & .942402 & -.045687 \end{bmatrix}$$

$$\underset{\sim}{T}^{-1} = \begin{bmatrix} 1.000000 & -.810255 & .012326 & -.202070 \\ 0 & 3.754984 & .066617 & .315381 \\ 0 & -.984495 & .006868 & .977621 \\ 0 & -.440446 & .494133 & -.053697 \end{bmatrix}$$

(These two matrices composed of right and left eigen vectors corresponding to the eigen values are inverses of each other - hence the symbols $\underset{\sim}{T}$ and $\underset{\sim}{T}^{-1}$. They satisfy the relationship $\underset{\sim}{P}' = \underset{\sim}{T}^{-1} \underset{\sim}{E} \underset{\sim}{T}$, where $\underset{\sim}{E}$ is a diagonal matrix consisting of eigen values taken as elements on the diagonal [5].) The matrices $\underset{\sim}{P}'$, $\underset{\sim}{T}$ and $\underset{\sim}{T}^{-1}$ will come in useful at later points in the present text (cf. SECTION IV).

SECTION II

SUM OF TRANSITION PROBABILITIES OF HOMOGENEOUS SUBPOPULATIONS : A CASE OF HETEROGENEITY

In order to relax the assumption of population homogeneity in the analysis of multistate transitions, some demographers have suggested the use of a simultaneous consideration of the homogeneous subgroups forming any given heterogeneous population: for example, homogeneous subpopulations based on categories of religion, place of birth, place of residence etc. This consideration will yield transition probability matrices for each subpopulation; computations in each separate case being still based on the assumption of Markovian behaviour. In the example considered above, the Belgian female population in the year 1981 can further be divided into three subpopulations determined on a linguistic basis. These three linguistic regions, namely French-speaking Wallonia, Flemish-speaking Flanders and bilingual Bruxelles-Brabant, are well known for their remarkable differences in political, economic, social, cultural and demographic behaviour (Lesthaeghe, 1977). The three corresponding region-specific transition probability matrices, for the same age interval (27, 28) as before, are found to be as follows (with the regional status-wise population figures given in the sixth column) – ${}_1\underset{\sim}{P}$ is that of Bruxelles, ${}_2\underset{\sim}{P}$ is of Wallonia and ${}_3\underset{\sim}{P}$ is of Flanders. (Note that the last row, corresponding to the state of death, composed of zeroes except for unity in the fifth column, has been dropped to simplify both the presentation and computation. Only 4x5 matrices are therefore under consideration).

${}_1\underset{\sim}{P}$ =

NM	PM	W	D	DH	POP.
.922769	.075661	.000037	.000674	.000860	2278
0	.980974	.000957	.017653	.000416	5098
0	.029567	.970164	.000263	.000006	33
0	.103206	.000050	.893829	.002916	327

${}_2\underset{\sim}{P}$ =

NM	PM	W	D	DH	POP.
.898531	.100146	.000045	.000691	.000586	3424
0	.984766	.000901	.013704	.000630	19311
0	.044938	.947190	.000310	.007561	129
0	.145707	.000066	.852858	.001369	700

$$
{}_{3}\underset{\sim}{P} = \begin{bmatrix} .886585 & .111564 & .000045 & .000609 & .001197 \\ 0 & .988027 & .000797 & .010854 & .000322 \\ 0 & .053730 & .945967 & .000293 & .000009 \\ 0 & .162460 & .000065 & .835559 & .001916 \end{bmatrix} \begin{matrix} 4799 \\ 35200 \\ 144 \\ 971 \end{matrix}
$$

Now that the Markov probabilities of transition among the nuptiality-related states are available for each linguistic region and for the entire country, it is possible (a) to weight the probabilities of each region with the relevant status-wise population proportions - thus obtaining the weighted probabilities for the entire country - and then (b) to examine whether these weighted probabilities for the entire country (obtained from the three regions) tally with the probabilities obtained independently through the Markov specification for the country as a whole - that is, with the elements of the matrix $\underset{\sim}{P}$. Following the procedure just outlined, the following weighted probability matrix (denoted by ${}_{w}\underset{\sim}{P}$) for the entire country is obtained.

$$
{}_{w}\underset{\sim}{P} = \begin{matrix} NM & PM & W & D & DH \end{matrix}
$$

$$
\begin{bmatrix} .898330 & .100052 & .000043 & .000650 & .000925 \\ 0 & .986367 & .000844 & .012359 & .000430 \\ 0 & .047418 & .949092 & .000297 & .003192 \\ 0 & .146893 & .000063 & .851156 & .001888 \end{bmatrix}
$$

The close correspondence between $\underset{\sim}{P}$ and ${}_{w}\underset{\sim}{P}$ is obvious, the differences between the two being negligibly small. Thus, it is seen that, having disaggregated a heterogeneous population into homogeneous subpopulations, each following a separate Markovian process as regards transitions among states, it is possible to express the probabilities of the total population as a linear sum of probabilities of each subpopulation, (the status-wise subpopulation proportions being used as weights). The same type of results were obtained when the procedure described above was used in connection with all age intervals (i.e. from 19 to 70).

In general matrix notation, this could be expressed as follows: let ${}_{r}\underset{\sim}{W}$ be a nxn diagonal matrix of weights attached to the r-th subpopulation corresponding to the n transient states of a Markovian process; and let the i-th diagonal element of matrix ${}_{r}\underset{\sim}{W}$ be the i-th status-wise proportion of the r-th subpopulation. In the specific case under study, r=1,2,3 and n=4. Then,

$$ {}_{1}\underset{\sim}{W}\,{}_{1}\underset{\sim}{P} + {}_{2}\underset{\sim}{W}\,{}_{2}\underset{\sim}{P} + \cdots\cdots + {}_{r}\underset{\sim}{W}\,{}_{r}\underset{\sim}{P} = \underset{\sim}{P} \qquad (1) $$

where $\underset{\sim}{P}$ is the transition probability matrix of the total population following a Markovian process, and ${}_{r}\underset{\sim}{P}$ is the r-th subpopulation's transition probability matrix following its own Markovian process.

The above result is not surprising, as it is the (mid-year) status-wise population figures which comprise the denominators (that is, the populations exposed to the risks of transition under consideration) of the expressions used for finding the rates, which lead to the probabilities in question. Considered from the point of view of Markov processes, however, the result leads to a reconsideration of the problem concerning "mixtures of Markov processes" [6]. It has been held that a mixture of Markov processes is, in general, a non-Markovian process. In practice, however, this has often been taken to mean that a mixture of Markov processes is necessarily non-Markovian. But in the present illustration, a mixture of probability matrices of subpopulations, each experiencing its own independent Markov process, is seen to yield a probability matrix which is the same (barring small differences which could be attributed to rounding errors) as that of the entire population following its own Markov process. In other words, aggregating the probability matrices of homogeneous subpopulations following independent Markov processes (using appropriate weights - status-wise population figures in our case) does not necessarily lead to a probability matrix ($_{w}\underset{\sim}{P}$) of the total population that is different from the $\underset{\sim}{P}$ matrix obtained for the Markov characterized total population. Thus, heterogenizing a population into a number of sub-categories does not necessarily eliminate the Markovian condition.

The above discussion gives rise to three related questions: (1) Given a _specific_ set of weights, is the corresponding weighted sum of several given Markov processes necessarily Markovian? (2) Given several Markov processes which together add up to a Markov process, is it _theoretically_ possible to find the weights appropriate for this adding process? (3) Is there a _theoretical_ possibility of finding the regional probabilities, given the weights and the probabilities of the total population?

The following observation can be made in connection with the first question. There is at least one set of weights (_viz_ status-wise population proportions) which renders Markovian a mixture of Markovian processes. The next two sections will deal with the remaining two questions. It will be seen that the theoretical possibilities referred to in these questions do exist.

SECTION III

THEORETICAL EVALUATION OF STATUS-WISE REGIONAL PROPORTIONS

The second question given above can be reformulated as follows. Assume that the transition probabilities of each regional population and of the total population are known, each having been obtained independently through the assumptions of (1) the homogeneity of individuals belonging to a given sub-group, and (2) the Markovian nature of the transitions involved [7]. Is it possible then to build a linear combination of regional status-wise transition probabilities which would be equal to those of the total population ? The problem reduces to that of finding the coefficients in the linear combination, under the constraint that the sum of these coefficients equals unity. Once these coefficients will have been estimated, how do they compare with the actual status-wise regional population proportions (given in Appendix Table 1)?

Note that an estimation of these coefficients in the linear combination, however, cannot be obtained merely through the use of the ordinary least squares (OLS) technique, as the OLS does not guarantee that the resulting estimates are positive; which is inadmissible since the population proportions in question are always non-negative. In order to satisfy this requirement, one is forced to turn to the restricted least squares estimator; in particular, the technique of Quadratic Programming.

Given the transition probabilities of each region and of the total population, a system of linear equations can be constructed, which expresses the relationship between the regional probabilities and the probabilities of the total population. Let ${}_ra_i$ be the coefficient associated with the transition probability ${}_rP_{ij}$ from status i to status j of the r-th region (for all j). Note that the a's have only the subscript i on the right because all transitions from i to any j are of the individuals in state i. Then the system of linear equations would be

$$\sum_r {}_ra_i \cdot {}_rP_{ij} = P_{ij} \tag{2}$$

The system (2) is made to vary over j=1,..5 and i=1,...4. This yields a complete system of four sets (varying over i) of five equations (varying over j) in three unknowns (varying over r). For example, expanding the system (2) for i=1 and j=1,5, we get,

$$
\begin{aligned}
{}_1a_1 \cdot {}_1P_{11} + {}_2a_1 \cdot {}_2P_{11} + {}_3a_1 \cdot {}_3P_{11} &= P_{11} \\
{}_1a_1 \cdot {}_1P_{12} + {}_2a_1 \cdot {}_2P_{12} + {}_3a_1 \cdot {}_3P_{12} &= P_{12} \\
{}_1a_1 \cdot {}_1P_{13} + {}_2a_1 \cdot {}_2P_{13} + {}_3a_1 \cdot {}_3P_{13} &= P_{13} \\
{}_1a_1 \cdot {}_1P_{14} + {}_2a_1 \cdot {}_2P_{14} + {}_3a_1 \cdot {}_3P_{14} &= P_{14} \\
{}_1a_1 \cdot {}_1P_{15} + {}_2a_1 \cdot {}_2P_{15} + {}_3a_1 \cdot {}_3P_{15} &= P_{15}
\end{aligned}
$$

all transitions from the NM

(2A)

Similarly, three further sets of five equations for all transitions from the PM, W and D can be written. In short, each set of five equations for transitions from a state i can be expressed in matrix form as follows:

$$\begin{bmatrix} {}_1P_{i1} & {}_2P_{i1} & {}_3P_{i1} \\ {}_1P_{i2} & {}_2P_{i2} & {}_3P_{i2} \\ {}_1P_{i3} & {}_2P_{i3} & {}_3P_{i3} \\ {}_1P_{i4} & {}_2P_{i4} & {}_3P_{i4} \\ {}_1P_{i5} & {}_2P_{i5} & {}_3P_{i5} \end{bmatrix} \cdot \begin{bmatrix} {}_1a_i \\ {}_2a_i \\ {}_3a_i \end{bmatrix} = \begin{bmatrix} P_{i1} \\ P_{i2} \\ P_{i3} \\ P_{i4} \\ P_{i5} \end{bmatrix}$$

that is,

$${}_a\underset{\sim}{P}_i \cdot \{A_i\} = \{P_i\} \tag{3}$$

where ${}_a\underset{\sim}{P}_i$ denotes the 5x3 matrix of regional probabilities from a state i, $\{A_i\}$ is the column vector of coefficients of each region and $\{P_i\}$ is the vector of probabilities from state i of the total population (all these arranged as above).

Note however that in the unrestricted least squares estimation of $\{A_i\}$, a vector $\{D\}$ standing for the column of error terms has to be introduced. Equation (3) can now be written as [8]:

$${}_a\underset{\sim}{P}_i \cdot \{A_i\} + \{D\} = \{P_i\} \tag{4}$$

This unrestricted model yields the minimal solution

$$\{\hat{A}_i\} = [\, {}_a\underset{\sim}{P}_i^t \cdot {}_a\underset{\sim}{P}_i \,]^{-1} \, {}_a\underset{\sim}{P}_i^t \, \{P_i\} \tag{5}$$

where ${}_a\underset{\sim}{P}_i^t$ stands for the transpose of the matrix ${}_a\underset{\sim}{P}_i$ and $\{\hat{A}_i\}$ gives the estimates of the a's through the least squares. Further ${}_{a\sim i}P_{i}^{t}\;{}_{a\sim i}P_{i}$ is non-singular (this condition of non-singularity being satisfied by virtue of the fact that the elements of ${}_{a\sim i}P_{i}$ are the probabilities of a stochastic process). Theoretically, when the unrestricted least squares is used, the condition $\sum_r {}_r a_i = 1$ is automatically satisfied: this however is not necessarily the case as regards the condition that each ${}_r a_i > 0$.

In order that the estimates of ${}_r a_i$ be non-negative, one has to resort to the restricted least squares estimator given by the following system derived from (4):

$$\{D^t\}.\{D\} \;=\; [\{P_i\} - {}_{a\sim i}P_i.\{A_i\}]^t\,[\{P_i\} - {}_{a\sim i}P_i.\{A_i\}] \quad (6)$$

where $\{D^t\}$ stands for the transpose of the vector of residuals obtained from (4). In this form, it is possible to ensure that conditions $\sum_r {}_r a_i = 1$ and ${}_r a_i > 0$ are satisfied. Since (6) is a quadratic form and the restrictions imposed are linear, the problem becomes one of quadratic programming (Lee et al., 1977). Though this could be reduced to a linear programming problem, the solution vector $\{\hat{A}_i\}$ can also be found, by keeping the system (6) as it is, through the use of a computer program for quadratic programming available in the NAG library (NAG HO2AAF) [9].

Applying the NAG computer package, the estimates (of $_r a_i$) for the three regions in the case of transitions in the age interval (27-28) are found to be as follows:

Region	NM	PM	W	D
Bruxelles	.210134	.088457	.106674	.159489
Wallonia	.337275	.322970	.421790	.350557
Flanders	.452592	.588572	.471535	.489954

These estimated weights are found to be very nearly equal to the actual regional status-wise population proportions [10] -cf. App. Table 1.

The following conclusions emerge from the above discussion:

1. It is possible to find a system of weights which enable the transition (Markov-related) probabilities of the total population to be expressed as a linear sum of the transition probabilities of the different regions, each following its own Markov process. The "theoretical possibility" referred to in the second question above does exist therefore.

2. These weights turn out to be very nearly identical to the relevant regional status-wise population proportions.

What is true for the age-interval 27-28 was also found to hold good for other age-intervals. Appendix Table 1 gives the estimated coefficients, in the linear combinations constructed as above, for different age groups from 20 to 50. Excepting those cases denoted by "Data Error" (DE) in the Table (cf. APPENDIX 1 for explanation), most of the estimates are seen to correspond very well to the actual status-wise regional population proportions.

SECTION IV

THEORETICAL HETEROGENIZATION OF MARKOVIAN PROBABILITIES

The third question formulated above deals with the possibility of heterogenizing a given (total) population into homogeneous subgroups under specified constraints. The present section spells out the assumptions on which the heterogenization process is built, underlines the principal characteristics of the methodology used, and computes the relevant characteristics of the resulting homogeneous subgroups in the case of Belgium and its three linguistic subregions. In general, an affirmative answer is given to the question raised as regards the possibility, in theory, of effecting the heterogenization under consideration.

The discussion starts with a very general formulation of the problem. Constraints and further assumptions are introduced progressively.

Assume (1) that the total population under consideration consists of three homogeneous regions; each of which is characterized by its own specific transition probabilities, (2) that each region follows its own Markovian process; and (3) that the total probability system of the population (-it is Markov characterized-) for any given age group, say (x, x+1), can be expressed as follows:

$$\underset{\sim}{P}(x) = {}_1\underset{\sim}{a} \cdot {}_1\underset{\sim}{P}(x) + {}_2\underset{\sim}{a} \cdot {}_2\underset{\sim}{P}(x) + {}_3\underset{\sim}{a} \cdot {}_3\underset{\sim}{P}(x) \qquad (7)$$

where ${}_r\underset{\sim}{a}$ denotes the diagonal matrix whose elements ${}_r a_i$ are coefficients attached to probabilities of transition from status i in the r-th region and matrix ${}_r\underset{\sim}{P}$ denotes the transition probability matrix of the r-th region. Thus, $\underset{\sim}{P}(x)$ on the left hand side is a mixture (i.e. a linear sum) of Markov processes. Since each region follows its own Markov process, we can write each of the probability matrices, ${}_r\underset{\sim}{P}(x)$, as being equal to $\exp\{{}_r\underset{\sim}{R}(x)\}$, where ${}_r\underset{\sim}{R}$ is the matrix of elements ${}_r R_{ij}$ denoting the intensities of transitions from state i to state j [11]. Equation (7) can thus be rewritten (age-group specification being omitted to simplify the notation) as follows:

$$\underset{\sim}{P} = {}_1\underset{\sim}{a} \cdot \exp\{{}_1\underset{\sim}{R}\} + {}_2\underset{\sim}{a} \cdot \exp\{{}_2\underset{\sim}{R}\} + {}_3\underset{\sim}{a} \cdot \exp\{{}_3\underset{\sim}{R}\} \qquad (8)$$

Note

1. that the discussion in this section will be carried out in terms of the 4x4 $\underset{\sim}{P}'$ matrix introduced earlier (cf. SECTION I). The symbolism is simplified by the use of $\underset{\sim}{P}$ for $\underset{\sim}{P}'$ and ${}_r\underset{\sim}{P}$ for ${}_r\underset{\sim}{P}'$.

2. that the problem consists of estimating the ${}_r\underset{\sim}{P}$ (or, equivalently the ${}_r\underset{\sim}{R}$) matrices;

3. that, in an n-state model with r homogeneous subgroups, this would involve a system of

   1. n^2+rn+n equations, of which

      1. n^2 arise since we are dealing with nxn matrices;

      2. rn relate to the stochastic nature of the ${}_r\underset{\sim}{P}$ matrices; $\sum_j {}_rP_{ij}=1$ for all r;

      3. n are connected with the constraints imposed on the ${}_ra_i$: $\sum_r {}_ra_i=1$ for all i (${}_ra_i$ being treated as "unknowns" at this point of the discussion).

2. rn^2+rn unknowns: i.e.

   1. rn^2 elements in the r probability matrices of order nxn to be estimated;

   2. rn unknown ${}_r a_i$'s (if these are held to be unknowns for the moment).

In the specific case under study, where n=4 and r=3, there will be 32 equations in 60 unknowns, and the system (7) or its corresponding system (8) is seen to be obviously under-determined.

Using the technique of matrix diagonalization (as used by Kitsul and Philipov [12]), equation (8) can be transformed into equation (9) :

$$\underset{\sim}{T}^{-1} . \underset{\sim}{P} . \underset{\sim}{T} = \underset{\sim}{T}^{-1} [{}_1\underset{\sim}{a} . \exp\{{}_1\underset{\sim}{R}\} + {}_2\underset{\sim}{a} . \exp\{{}_2\underset{\sim}{R}\} + {}_3\underset{\sim}{a} . \exp\{{}_3\underset{\sim}{R}\}] . \underset{\sim}{T} \quad (9)$$

where $\underset{\sim}{T}$ and $\underset{\sim}{T}^{-1}$ have the meanings assigned to them in SECTION I. Note that the eigen values of $\underset{\sim}{P}$ are assumed to be real and distinct (as is usually the case in demographic studies).

Corresponding to the stochastic nature of the ${}_r\underset{\sim}{P}$ matrices, the elements of the related ${}_r\underset{\sim}{R}$ matrices satisfy the conditions $\sum_j {}_rR_{ij}=0$, ${}_rR_{ij}\geq 0$ for $i\neq j$ and ${}_rR_{ii}<0$.

Further assumptions have to be introduced at this point so as to reduce the under-determined nature of the equations hitherto developed. Two variants of assumptions concerning the ${}_r\tilde{R}$ matrices were tried out. In the first instance, the elements of ${}_2\tilde{R}$ (and similarly of ${}_3\tilde{R}$) corresponding to a fixed state of origin, were taken as proportional to the corresponding elements of ${}_1\tilde{R}$. Thus, for example, elements ${}_2R_{1j}$ (belonging to ${}_2\tilde{R}$) were taken as equal to ${}_2k_1 . {}_1R_{1j}$ (${}_1R_{1j}$ belonging to ${}_1\tilde{R}$) for all j, and ${}_2R_{2j}$ as equal to ${}_2k_2 . {}_1R_{2j}$, etc. Similarly, ${}_3k_1 . {}_1R_{1j} = {}_3R_{1j}$, etc. In the second variant experimented with, all elements of ${}_2\tilde{R}$ (and of ${}_3\tilde{R}$) were considered to be equal to ${}_2k$ (and ${}_3k$) times the corresponding elements of ${}_1\tilde{R}$. Thus, for example, ${}_2R_{ij} = {}_2k . {}_1R_{ij}$ for all i and j. Similarly, ${}_3R_{ij} = {}_3k . {}_1R_{ij}$ for all i and j. Since these two simplifying assumptions give approximately the same final results, only the second will be used in the discussion that follows. Thus, with ${}_2\tilde{R} = {}_2k . {}_1\tilde{R}$ and ${}_3\tilde{R} = {}_3k . {}_1\tilde{R}$, equation (9) gives rise to equation (10):

$$\tilde{T}^{-1}.\tilde{P}.\tilde{T} = {}_1\tilde{a}\,\tilde{T}^{-1}\exp\{{}_1\tilde{R}\}\tilde{T} + {}_2\tilde{a}\,\tilde{T}^{-1}\exp\{{}_2k . {}_1\tilde{R}\}\tilde{T} + {}_3\tilde{a}\,\tilde{T}^{-1}\exp\{{}_3k . {}_1\tilde{R}\}\tilde{T} \qquad (10)$$

with

1. n^2+2n equations, of which

    1. n^2 are related to the nxn matrix ${}_1\tilde{R}$;
    2. n are associated with the constraint $\Sigma_j\, {}_1R_{ij}=0$; and
    3. n are linked with the constraint $\Sigma_r\, {}_ra_i=1$.

2. and $n^2+rn+(r-1)$ unknowns

   1. n^2 elements of ${}_1\tilde{R}$;

   2. rn unknowns corresponding to the ${}_r a_i$'s; and

   3. (r-1) related to the proportionality factors.

In our case, where, n=4 and r=3, we would have 24 equations in 30 unknowns. The system (10) is still under-determined, though the number of unknowns has been considerably reduced following the assumption that the intensity matrix of each of two regions is proportional to that of the third.

Since the LHS of (9) or (10) is a diagonal matrix, so is its RHS. Further, using the theory concerning the "relations of matrices"[13], it can be shown that $\tilde{T}$, which diagonalizes $\tilde{P}$, diagonalizes also a linear combination of $\exp\{{}_1\tilde{R}\}$, $\exp\{{}_2k.{}_1\tilde{R}\}$ and $\exp\{{}_3k.{}_1\tilde{R}\}$. If y_i stands for the i-th eigen value of $\tilde{P}$ and z_i the i-th eigen value of $\{{}_1\tilde{R}\}$, equation (11) can now be written in the following simple scalar form:

$$y_i = {}_1a_i.\exp\{z_i\} + {}_2a_i.\exp\{{}_2k.z_i\} + {}_3a_i.\exp\{{}_3k.z_i\} \quad (11)$$

for i=1,..4

Since the largest eigen value (y_i) of $\tilde{P}$ is equal to unity (cf. SECTION I), its corresponding z_i will be equal to 0.

One of the equations in the system (11) can therefore be excluded. Thus, there are now only (n−1) equations (3 in our case), and (n+rn−2) unknowns (14 in our case):

- (n−1) corresponding to the eigen values z_i
- r(n−1) corresponding to the coefficients $_{r}a_i$
- (r−1) corresponding to the proportionality factors

The assumptions regarding the $_{r}\underset{\sim}{R}$ matrices described above are now followed by assumptions concerning the $_{r}a_i$. The remarks made in SECTIONS II and III in connection with the two questions raised in SECTION II show that the observed status-wise region specific population proportions could be appropriately taken as being identical to the $_{r}a_i$. If the values of $_{r}a_i$ are fixed in this fashion, the model reduces to a search of 5 unknowns (the z's and k's) using three equations. Estimates of these unknowns were found through the use of a non-linear optimization programme found in the CERN library (i.e. D506). The algebraic manipulations entering into play as well as the strategy adopted to start the iteration procedures involved are given in Appendix 2. Once the z's and k's are estimated, estimates of the region specific transition matrices (i.e. $_{r}\underset{\sim}{P}$'s) are obtained by a process of inverse diagonalisation (also detailed in Appendix 2).

The resulting (best) estimates for the age-group 27–28 for the z's and k's and the resulting $_r\underset{\sim}{P}$ matrices are as follows:

$z_1 = 0.92000$

$z_2 = 1.000000$

$z_3 = 0.87175$

$z_4 = 0.96139$

$_2k = 1.3067$

$_3k = 1.3278$

$$_1\hat{\underset{\sim}{P}} = \begin{bmatrix} 0.92000 & 0.079627 & 0.000034 & 0.000338 \\ 0.00000 & 0.989417 & 0.000671 & 0.009912 \\ 0.00000 & 0.037806 & 0.961993 & 0.000201 \\ 0.00000 & 0.118017 & 0.000043 & 0.881941 \end{bmatrix}$$

$$_2\hat{\underset{\sim}{P}} = \begin{bmatrix} 0.896771 & 0.102579 & 0.000055 & 0.000595 \\ 0.000000 & 0.986441 & 0.000870 & 0.012689 \\ 0.000000 & 0.049032 & 0.950634 & 0.000335 \\ 0.000000 & 0.151076 & 0.000071 & 0.848854 \end{bmatrix}$$

$$_3\hat{\underset{\sim}{P}} = \begin{bmatrix} 0.895195 & 0.104134 & 0.000057 & 0.000615 \\ 0.000000 & 0.986241 & 0.000883 & 0.012876 \\ 0.000000 & 0.049798 & 0.949857 & 0.000345 \\ 0.000000 & 0.153299 & 0.000073 & 0.846629 \end{bmatrix}$$

A comparison of these estimated regional probabilities with those given on p.13 (care being taken to convert the latter into matrices of order 4x4 - eliminating death) shows the estimated values to be very close approximations of the originals obtained directly from the data. Similar estimations made in connection with the remaining age-intervals gave similar results.

Starting with the transition probabilities of a total population (evaluated on the basis of the Markovian assumption), we have therefore demonstrated that it is possible to arrive at very close estimates of the corresponding matrices of the subregions in question if

1. the transitions in each subregion were supposed to belong to a region specific Markov process;

2. and certain simplifying assumptions were allowed.

That a Markov process can be looked on as a mixture of Markov processes - or looked at conversely, that a mixture of Markov processes is not necessarily non-Markovian - is once again vindicated.

SECTION V

CONCLUSION

If the conclusions of SECTIONS II, III and IV are admitted, it follows that the mere heterogenization of a given Markov characterized population into subpopulations, which are themselves Markov characterized, does not necessarily succeed in eliminating its basic Markovian nature. Something more than the use of <u>any</u> one specified form of heterogenization is needed for this. In other words, all forms of heterogenization do not necessarily lead to the elimination of the Markovian condition.

This point is not clearly in evidence in the relevant literature, where the distinction between the two concepts - i.e. of homogeneity of a population and the Markovian nature of the process it experiences - is at least slurred over. A few words aimed at drawing out the differences between the two concepts are therefore in order.

The assumption of population homogeneity endows all individuals belonging to a given age-group and present at the same time in a given state, with identical propensities of moving out of that state. All differences - whether of the past or characterizing the present - are thus ignored.

The Markovian assumption, on the other hand, makes it possible to forget the past histories of individuals (in a given age-group and present at the same time in a given place). It only cancels out past differences. Thus, the propensities of moving out of a state experienced by members of a given age-group are taken to be independent of the past. The Markovian assumption may or may not imply the homogeneity assumption; but the latter does include the former. Though the Markovian assumption ignores past differentiation, it does not necessarily destroy present heterogeneity. Thus there remains the possibility of admitting present heterogeneity while retaining the Markovian assumption. It is only when heterogeneity due to the past is introduced that the Markovian assumption is relaxed: and while it <u>is</u> true that the homogeneity assumption taken in all its generality involves the Markovian assumption, it is <u>not</u> always true that any sort of heterogenization makes a Markov process non-Markovian.

FOOTNOTES

[1]. Multistate demography in its present state only deals with fictive cohorts (of a single sex) obtained from period data. The two-sex problem as such has not yet been taken into account in the present state of its art.

[2]. A caution on the possibility of confusion of terminologies is in order. The homogeneity assumption, as treated in this study, refers to "population homogeneity" - that is, the homogeneity or sameness of characteristics of individuals in a given population. The "time-homogeneity" (also sometimes referred to as "time-stationarity") of a Markov process as used in the literature dealing with stochastic processes refers to the sole dependence of a Markovian state transition probability on the difference t-s between two time points s and t, $(s < t)$, within which period the relevant transition takes place. For the sake of clarity between the two entirely different notions, this study will use the term "homogeneity" in reference to population homogeneity and retain the expression "time-homogeneity" for use in connection with a Markovian process.

[3]. That a Markovian process is characterized by the exponential distribution (which lacks memory) can be seen in the analysis of the Kolmogorov differential equations to which a Markov process gives rise to. For details, see Feller, 1950 and 1966 and Ross, 1983.

[4]. The eigen values can be calculated through any available computer program. The eigen values and vectors in this study have been found through the use of the computer library CERN of Geneva.

[5]. Since all the eigen values (in demographic studies) are known to be <u>real</u> <u>and</u> <u>distinct,</u> the corresponding eigen vectors should also be real and distinct. For details, cf. Keyfitz N., 1977.

[6]. The expression "Markov mixtures" is rare in the literature on stochastic processes. Its only occurrence, as far as we are aware, is in Feller, and that too as an exercise to be solved (Feller, 1966, p. 426).

[7]. These are the usual assumptions made in constructing the Markov-based multistate model. See Flow Chart 1 and Willekens et al., 1980.

[8]. Equation (4) holds under the assumption that Ex[d]=0 and $Ex[d^2]=\sigma^2$, where Ex denotes the mathematical expectation and d are the error terms.

[9]. The NAG Fortran program H02AAF minimizes a symmetric, positive definite quadratic form in n non-negative variables, subject to m linear constraints which may be equations or upper-bounded inequalities. Details on the application of this program are given in APPENDIX 1.

[10]. The estimation procedure has been carried out with the use of 4x5 matrix in this section. It is also possible to work with the 4x4 adjusted P′ matrices. Early trials revealed that working with the non-adjusted probabilities yields estimates closer to the actual population proportions than working with the adjusted probabilities.

[11]. A Markov process leads to the formulation of the state probability matrices P(x) in terms of the intensity matrix R(x) as P(x, x+t)=P(x).exp{R(x).t}. This is in fact a solution of the related Kolmogorov differential equations assuming the constancy of the intensity function in the age interval under consideration.

[12]. The process of diagonalization will be seen in due course to be a means of reducing the complexity of the problem.

[13]. The matrices A and B are <u>related</u> if they can be diagonalized by the same transformation T. It is easy to show that if the matrices A and B are related, then the matrix C, where C=a.f(A)+bg(B) and f(.) and g(.) are scalar functions and a and b are real numbers, is also related to A and B (Gantmacher 1977, Ch.V). In particular, if v is the diagonalized matrix diag(m), then diag(exp{m})=exp{v}. cf. Kitsul and Philipov, 1981, p.9f.

BIBLIOGRAPHY

1. Bartholomew D. J. (1967), Stochastic Models for Social Processes , Second edition, John Wiley, N. Y.

2. Chiang C. L. (1968), Introduction to Stochastic Processes in Biostatistics , Wiley, N. Y.

3. Feller, W. (1950), An Introduction to Probability Theory and Its Applications , Vol. I, John Wiley, London.

4. Feller W. (1966), An Introduction to Probability Theory and Its Applications , Vol. II, John Wiley, London.

5. Gantmacher F. R. (1977), The Theory of Matrices , Vol. I, Chelsea Publishing Co., New York.

6. Howard R. A. (1971), Dynamic Probabilistic Systems , Vol. I, John Wiley, N. Y.

7. Keyfitz N. (1977), Introduction to the Mathematics of Population with Revisions, Addison-Wesley Publishing Co., Amsterdam.

8. Keyfitz N. (1980), "Multistate Demography and Its Data: A Comment", in Rogers A. (ed.), 1980.

9. Kitsul and Philipov(1981), "The one-year/five-year Migration Problem", in Rogers A. (ed.), 1981.

10. Land K. C and Rogers A. (1982), (eds.), Multidimensional Mathematical Demography , Academic Press, N. Y.

11. Land K. C and Schoen R. (1982), "Statistical Methods for Markov-generated Increment-Decrement Life Tables with Polynomial Gross Flow Functions", in Land and Rogers, 1982.

12. Ledent J. (1980), "Multistate Life Tables: Movement versus Transition Perspectives", in Rogers A. (ed.), 1980.

13. Ledent J. (1981), "Constructing Multiregional Life Tables using Place-of-birth-specific Migration Data", in Rogers A. (ed.), 1981.

14. Lee T. C., Judge G. G. and Zellner A. (1977), Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data, North Holland Pub. Co., Amsterdam.

15. Lesthaeghe R. (1977), The Decline of Belgian Fertility, 1800-1970 , Princeton Univ. Press, N. J.

16. Philipov and Rogers (1981), "Multistate Population Projections", in Rogers A., 1981.

17. Rogers A. (1975), Introduction to Multiregional Mathematical Demography, Wiley Pub., N. Y.

18. Rogers A. (1980), Essays in Multistate Mathematical Demography, IIASA, RR-80-10, Laxenburg.

19. Rogers A. (1981), Advances in Multiregional Demography, RR-81-6, IIASA, Laxenburg.

20. Ross S. (1983), Stochastic Processes, John Wiley & Sons, N. Y.

21. Willekens F. and Rogers A. (1978), Spatial Population Analysis, Methods and Computer Programs, RR-78-18, IIASA, Laxenburg.

22. Willekens F. et al. (1980), Multistate Analysis of Marital Status Life Tables, Theory and Application, NIDI Working Paper No. 17, Voorburg, Netherlands.

APPENDIX 1

The NAG programme HO2AAF, used to estimate the parameters ${}_{\gamma}a_i$ in Equation (6), essentially engineers a minimization of a quadratic function. The success of the programme depends on the suitable lower and upper limits of ${}_{\gamma}a_i$. These limits were found through Table 1, which carries the region specific population proportions as distributed by age.

In the optimization procedure, one of the error indicators signalled by the computer is "Data Error" which means that either the quadratic function is not positive definite or excessive rounding errors have occurred due to bad scaling. In the following Table 1, there are some cases of this sort. No further trials were made to correct these errors.

App. Table 1: Status-wise population proportions of each region in Belgium - females, 1981
(Estimates through NAG HO2AAF are given in brackets)

Age	Bruxelles NM	Bruxelles PM	Bruxelles W	Bruxelles D	Wallonia NM	Wallonia PM	Wallonia W	Wallonia D	Flanders NM	Flanders PM	Flanders W	Flanders D
20	.101622	.078593	.119048	.201923	.307685	.334899	.452381	.326923	.590693	.586507	.428571	.471154
	(.100000)	(.088381)	(.130381)	(.168578)	(.300000)	(.324749)	(.434697)	(.390000)	(.600000)	(.586869)	(.434922)	(.441422)
21	.116069	.076725	.147059	.140351	.312101	.320521	.235294	.346491	.571830	.602755	.617647	.513158
	(.100000)	(.074450)	(.139927)	(.137171)	(.346876)	(.326942)	(.244805)	(.350013)	(.553124)	(.598608)	(.615268)	(.512816)
22	.134998	.077036	.091837	.151741	.316345	.316180	.479592	.353234	.548657	.606784	.428571	.495025
	(.127555)	DE	(.090672)	(.148202)	(.320336)	DE	(.464309)	(.355627)	(.552109)	DE	(.445019)	(.496171)
23	.157145	.081323	.071429	.180579	.315957	.313325	.396104	.367973	.526898	.605352	.532468	.451448
	(.150809)	(.088177)	(.071884)	(.172957)	(.315795)	(.312457)	(.388079)	(.368096)	(.533398)	(.599366)	(.540037)	(.458947)
24	.174904	.081101	.109756	.137088	.324571	.313778	.390244	.371945	.500526	.605121	.500000	.490967
	(.170429)	(.084818)	(.130172)	DE	(.320891)	(.311373)	(.249828)	DE	(.508680)	(.603809)	(.620000)	DE
25	.191024	.084813	.121339	.129291	.320818	.315903	.401674	.360847	.488158	.599283	.476987	.509861
	DE	(.086958)	(.123664)	(.124699)	DE	(.313818)	(.395646)	(.357649)	DE	(.599224)	(.480690)	(.517653)
26	.203430	.083436	.105263	.136994	.322766	.321626	.408907	.371603	.473804	.594938	.485830	.491403
	(.207411)	DE	(.102194)	DE	(.258924)	DE	(.412564)	DE	(.533665)	DE	(.485241)	DE
27	.216932	.085524	.107843	.163664	.326064	.323961	.421569	.350350	.457004	.590515	.470588	.485986
	(.210134)	(.088457)	(.106674)	(.159489)	(.337275)	(.322970)	(.421790)	(.350557)	(.452592)	(.588572)	(.471535)	(.489954)
28	.232605	.102274	.093567	.160168	.324264	.388836	.441520	.375681	.443131	.508890	.464912	.464151
	(.229633)	(.101469)	DE	(.148500)	(.321620)	(.378531)	DE	(.390000)	(.448747)	(.520000)	DE	(.461500)
29	.229539	.084881	.112000	.169126	.330579	.330004	.394667	.368568	.439883	.585115	.493333	.462306
	(.228856)	DE	DE	(.165272)	(.337581)	DE	DE	(.369648)	(.433563)	DE	DE	(.465081)
30	.235312	.086081	.093541	.182627	.326620	.333394	.432071	.378426	.438068	.580525	.474388	.438946
	DE	(.086526)	(.077056)	(.178352)	DE	(.334221)	(.439879)	(.384372)	DE	(.579253)	(.483064)	(.437276)
31	.241753	.086184	.097713	.184615	.333446	.334859	.478170	.383900	.424801	.578957	.424116	.431485
	(.243559)	DE	(.097401)	(.181704)	(.325223)	DE	(.477143)	(.358310)	(.431217)	DE	(.425456)	(.459986)
32	.245426	.081623	.112963	.191036	.332841	.339161	.431481	.378894	.421734	.579215	.455556	.430070
	(.148026)	(.082385)	(.112791)	DE	(.251974)	(.339544)	(.430474)	DE	(.600000)	(.578071)	(.456735)	DE
33	.244024	.085102	.112179	.197364	.327490	.340425	.442308	.377358	.428486	.574473	.445513	.425277
	(.231250)	DE	(.110803)	(.191835)	(.330099)	DE	(.440615)	(.388935)	(.438651)	DE	(.448582)	(.419230)
34	.243047	.082467	.108725	.205378	.317287	.340488	.440268	.366991	.439666	.577046	.451007	.427632
	(.233991)	(.083205)	DE	(.201589)	(.250000)	(.340510)	DE	(.365562)	(.516009)	(.576285)	DE	(.432849)
35	.241471	.083494	.122421	.193591	.298241	.305089	.360385	.338949	.460288	.611417	.517194	.467460
	(.249539)	DE	(.124191)	(.190512)	(.298290)	DE	(.355843)	(.343304)	(.452171)	DE	(.519966)	(.466184)

contd.

App. Table 1. contd.

Age	Bruxelles NM	Bruxelles PM	Bruxelles W	Bruxelles D	Wallonia NM	Wallonia PM	Wallonia W	Wallonia D	Flanders NM	Flanders PM	Flanders W	Flanders D
36	.240776	.089683	.118501	.218493	.293250	.301957	.378476	.324889	.465974	.608359	.503023	.456618
	DE	(.090299)	(.119277)	(.214195)	DE	(.301604)	(.376686)	(.325877)	DE	(.608097)	(.504037)	(.459928)
37	.249258	.087665	.118834	.222222	.299407	.296912	.382287	.341880	.451335	.615423	.498879	.435897
	(.249827)	DE	(.118675)	(.216873)	(.307338)	DE	(.383291)	(.348297)	(.442835)	DE	(.498035)	(.434830)
38	.229223	.087607	.136468	.234303	.284853	.296825	.384174	.351838	.485925	.615568	.479358	.413859
	(.231584)	(.088238)	(.137138)	DE	(.285758)	(.297005)	(.378082)	DE	(.482658)	(.614757)	(.484781)	DE
39	.220225	.088312	.099237	.226788	.283521	.302481	.392585	.324773	.496255	.609207	.508179	.448439
	(.215018)	(.088841)	(.099276)	(.223739)	(.288875)	(.302950)	(.391368)	(.326687)	(.496107)	(.608209)	(.509356)	(.449574)
40	.226266	.086720	.121581	.231048	.278424	.310907	.403242	.348387	.495310	.602373	.475177	.420565
	(.215662)	(.087566)	DE	(.227695)	(.295267)	(.310622)	DE	(.345990)	(.489071)	(.601812)	DE	(.426315)
41	.201200	.088109	.131876	.230879	.282691	.305927	.394122	.340958	.516109	.605964	.474002	.428163
	(.122451)	(.089249)	(.133349)	(.228597)	(.277549)	(.305264)	(.393979)	(.339994)	(.600000)	(.605487)	(.472672)	(.431409)
42	.184583	.080576	.103659	.226147	.296597	.307336	.405827	.329570	.518820	.612088	.490515	.444283
	DE	DE	(.103030)	(.225476)	DE	DE	(.405918)	(.337882)	DE	DE	(.491052)	(.436641)
43	.190415	.086104	.109987	.237350	.268459	.301626	.395702	.340371	.541127	.612270	.494311	.422279
	(.198099)	(.086848)	(.110114)	(.235517)	(.271735)	(.301548)	(.395551)	(.338626)	(.530166)	(.611603)	(.494335)	(.425856)
44	.182050	.086345	.113135	.226139	.267327	.303745	.392936	.347327	.550623	.609910	.493929	.426535
	(.184014)	(.087026)	(.113475)	(.224134)	(.265703)	(.303712)	(.392324)	(.346710)	(.550283)	(.609262)	(.494201)	(.429156)
45	.181619	.088358	.097036	.241450	.256955	.305605	.398184	.322883	.561425	.606038	.504780	.435668
	(.183011)	(.088835)	(.097426)	DE	(.257692)	(.305114)	(.397294)	DE	(.559297)	(.606051)	(.505280)	DE
46	.181423	.088910	.110542	.222222	.253495	.308109	.413572	.347121	.565082	.602980	.475886	.430657
	(.191986)	DE	DE	DE	(.253746)	DE	DE	DE	(.554268)	DE	DE	DE
47	.180458	.088951	.117784	.248571	.256162	.311921	.405269	.325306	.563380	.599128	.476947	.426122
	(.180156)	(.089514)	(.116899)	(.246783)	(.257453)	(.312559)	(.414287)	(.324977)	(.562392)	(.597928)	(.468814)	(.428240)
48	.172541	.089258	.120153	.253545	.251279	.318485	.390655	.340284	.576180	.592257	.489193	.406172
	(.173447)	DE	(.119874)	(.251221)	(.252316)	DE	(.390696)	(.341799)	(.574237)	DE	(.489430)	(.406980)
49	.169877	.089773	.108731	.242509	.257624	.324377	.400217	.358370	.572499	.585850	.491052	.399121
	(.161738)	DE	(.109023)	(.239798)	(.256909)	DE	(.399921)	(.359894)	(.581352)	DE	(.491057)	(.400307)
50	.179831	.092708	.104418	.272909	.258820	.324277	.406095	.361745	.561348	.583015	.489487	.365346
	DE	(.093205)	(.104696)	DE	DE	(.324562)	(.406219)	DE	DE	(.582233)	(.489085)	DE

APPENDIX 2

The use of the CERN programme D506 in connection with the problem of estimating the z_i's and ${}_rk$'s, found in Eq.(11), consists essentially of minimizing a suitably defined function F (say), related to the z_i's and ${}_rk$'s to be estimated. Following standard procedure, the function F is taken to be the sum of the squares of the differences between observed values and their model equivalents. It can therefore be expressed as follows:

$$F(z_1, z_2, z_3, {}_2k, {}_3k) = \sum_i (Y_i - {}_1a_i . e^{z_i} - {}_2a_i . e^{z_i * {}_2k} ...)$$

or more simply as

$$F(s_1, s_2, s_3, {}_2k, {}_3k) = \sum_i [Y_i - {}_1a_i . s_i - {}_2a_i . s_i^{({}_2k)} - {}_3a_i . s_i^{({}_3k)}] \quad (12)$$

where $s_i = \exp(z_i)$. The problem now reduces to a search for the "best" values of s_1, s_2, s_3, ${}_2k$ and ${}_3k$. Once these values have been found, the regional probabilities can be found through the following inverse transformation

$$\text{diag}[{}_r\underset{\sim}{p}] = \underset{\sim}{T}^{-1} {}_r\underset{\sim}{P}\, \underset{\sim}{T} = \underset{\sim}{T}^{-1} \exp\{{}_rk\, {}_1\underset{\sim}{R}\}\underset{\sim}{T} = \exp\{{}_rk\underset{\sim}{Z}\} \quad (13)$$

where $\underset{\sim}{Z}$ is the matrix of eigen values of the intensity matrix ${}_1\underset{\sim}{R}$. It follows that

$$\begin{aligned} {}_r\underset{\sim}{P} &= \underset{\sim}{T}.\text{diag}[{}_r\underset{\sim}{p}].\underset{\sim}{T}^{-1} \\ &= \underset{\sim}{T}.\exp\{{}_rk\underset{\sim}{Z}\}.\underset{\sim}{T}^{-1} \\ &= \underset{\sim}{T}.{}_r\underset{\sim}{S}.\underset{\sim}{T}^{-1} \end{aligned}$$

where $\underset{\sim}{S}$ is the diagonal matrix with $s_i = \exp(z_i)$.

In using the CERN D506 programme to estimate the s_i's and ${}_rk$'s of Eq.(12), the following points have to be noted.

1. Arbitrary specification of upper and lower bounds of parameters to be estimated does not lead to satisfactory results (for example, negative probabilities may be the end result of the inverse transformation referred to above). Hardly any problem exists as regards the s_i's, since the corresponding z_i's (the eigen values of the intensity matrix) are either zero or negative. However, in order to avoid difficulties in the estimations of the k's, suitable initial values, along with upper and lower limits, have to be fed into the computer so as to start the iterative process.

2. The following method used to compute these initial values (of $_{r}k$'s) was found to be satisfactory. It consisted of calculating, for each five-year age group, a region specific transition rate which was defined as the number of transitions (all transitions included, except death) per person (woman, in our case) per year of exposure. Taking Bruxelles as the reference region, the k value corresponding to each of the other two regions was obtained as the ratio of the rate specific to that region to the rate of Bruxelles. The k value of a given five-year age group was taken as being suitable for each of the single year falling within that group.