MULTISTATE DEMOGRAPHY

Jacques Ledent
Institut national de la recherche scientifique, Université du Québec, Canada

Yi Zeng
Duke University and Peking University, China

Keywords: Multistate demography, multistate life table, multiregional life table, single-state life table, transition probabilities, Markov process, multistate stable population model, multistate projection models, multiregional population projections, working life tables, active and disabled life expectancies, marriage/union status life table, population projections by educational attainment, households and living arrangements forecasting, household momentum, migration model schedules, micro-simulation, macro-simulation, integrated MicMac approach.

Contents

1. Introduction
2. The Multistate Life Table
  2.1. A Reminder on the Ordinary Single-State Life Table
  2.2. The Multistate Life Table Functions
  2.3. Estimating the Transition Probability Matrices
  2.4. Dealing with the Underlying Markov Process
  2.5. Summary Measures
  2.6. Additional Considerations
3. The Multistate Projection Models
  3.1. Multistate Projection and Stable Growth
  3.2. Multistate Population Projections in with Time-Varying Input
4. Applications of Multistate Demography
  4.1. Multistate Working Life Tables
  4.2. Active and Disabled Life Expectancies
  4.3. Multistate Life Table Analysis on Marriage/Union Formation and Dissolution
  4.4. Multistate Population Projections by Educational Attainment
  4.5. Multistate Projections of Households and Living Arrangements
  4.6. Modeling and Estimation of the Age and Spatial Structures of Migration Flows
5. Bridging the micro- and macro-simulation models: recent development of multistate demography

Glossary
Bibliography
Biographical Sketches

Summary

This chapter deals with multistate demography, an extension of classic mathematical demography to the case of individuals grouped in various categories of demographic and/or socioeconomic attribute(s) such as marital/union status, labor activity status, region of residence or health status. Individuals can move from and to these categories
but—and this is a salient feature of multistate demography—they are allowed to return to a state/status previously occupied; a long standing modeling difficulty which was eventually overcome by having recourse to a vector/matrix notation. This exposition, kept as little technical as possible, covers the basic principles of multistate demography, its two main models—that is, the multistate life table and the multistate projection model—as well as the generalization of these models to the case of families. It also alludes to the recent development of multistate demography which tends to bring this subfield of demography closer to mainstream statistics. Finally, adopting a more empirical viewpoint, it presents several applications of multistate demography to various population areas.

1. Introduction

Multistate demography is a topic whose complexity makes exposing and comprehending it a real challenge. Therefore, this chapter does not attempt in any way to cover its technical details, which are at the risk of being unpalatable for this encyclopedia. Rather it is a review limited to presenting the basic principles and applications of multistate demography—using as simple an exposition as possible—and offering a quick overview of its extensions and practical uses.

Basically, multistate demography refers to the study of populations stratified by at least one attribute (other than gender, age or race) that reflects region of residence, marital/union status, labor activity status, and so forth. Initially concerned with the generalization of the classical models of mathematical demography—that is, the life table and the projection model—multistate demography deals with transfers between alternative categories (i.e., states or statuses) of the population attribute(s), including re-entries into these categories.

Re-entry into a previously occupied category was first tackled as early as a century ago, in the context of a model involving two states (healthy and disabled) between which individuals could move back and forth (Du Pasquier 1912/13). However, despite the relevance of studying changes between multiple categories of population attribute(s), it was not too long ago that demographers began to seriously investigate such changes. What posed a problem to them was the complexity in estimating and computing the re-entries into a state or status previously occupied. Such a circumstance continued until the early 1970s when the advent of mainframe computers made it possible to crunch the numbers behind the construction of the models initially considered. Since then, multistate demography has grown into a sub-field of its own, tackling various types of transfers concerning individuals and groups of individuals such as families/households.

Beside significant modeling development, its pioneering contributions include various applications of the models thus developed to macro population-level (i.e., census and survey) data on multiregional migration (Rogers 1975; Rogers and Willekens 1986), marriage and divorce (Schoen and Land 1979; Willekens et al. 1982), labor activity (Hoem, 1977) as well as extensions to other topics of multi-group population (Land and Rogers 1982; Schoen 1988; Rogers 1995). Today, multistate demography largely intersects with the applications of traditional and more recent statistical methods, including event history analysis, to both macro and micro level data, thus making inroads into epidemiology, medicine and public health.
In the following sections, we first present the basic principles and the main constituents of the two fundamental and closely related models of multistate demography—that is, the multistate life table and the multistate projection model. We then discuss practical applications of these models in studies concerning populations, socioeconomics, health, and other related fields. Finally, we present the recent development of multistate demography aiming at bridging the micro- and macro-simulation models.

2. The Multistate Life Table

Understanding the multistate life table can be made easier by first reviewing the ordinary single-state life table (a more sophisticated description of which appears in Section 2 of Demographic Models and Actuarial Science) and then generalizing from it.

2.1. A Reminder on the Ordinary Single-State Life Table

In its ordinary version, the single-state life table is a tool that follows the life course of a birth cohort (a group of individuals born during a given period of time) until the death of its last survivor, as it gradually decreases in size under the effect of mortality. Basically, such a table includes several columns which document the evolution over successive ages of various functions relating to the life and death of the cohort members. These ages are usually (but not necessarily) equally spaced \( n \) years apart.

Typically, the first two columns are a column containing the number \( l_x \) of individuals who, out of the number \( l_0 \) of individuals born (i.e., individuals aged 0), are alive at age \( x \), and a column containing the numbers \( L_x \) of person-years lived by the cohort’s survivors between ages \( x \) and \( x+n \). Next are a column containing the cumulated numbers \( T_x \) of person-years lived beyond age \( x \) \( (T_x = \sum_{y=x}^{\infty} L_y) \) and a column containing the numbers \( e_x \) representing the average remaining lifetime for individuals aged \( x \), also known as the life expectancy at age \( x \) \( (e_x = T_x/l_x) \).

Usually, there is also a column of the numbers \( d_x \) of the cohort members that die between ages \( x \) and \( x+n \). Since the cohort is closed in the sense that \( i \) there is no exit other than through death, and \( ii \) re-entry is not possible, each number \( d_x \) of deaths is the difference in the number of survivors at the beginning and end of the relevant age interval \( (d_x = l_x - l_{x+n}) \).

In practice, a life table reflects a particular mortality pattern embodied in a set of age-specific mortality rates, with \( m_x \) being the mortality rate between ages \( x \) and \( x+n \) \( m_x = d_x/L_x \). Then, assuming that the number \( L_x \) of person-years lived can be approximated by the arithmetic mean of the number of survivors at ages \( x \) and \( x+n \) times the width \( n \) of the age interval \( L_x = \frac{n}{2}(l_x + l_{x+n}) \) leads to a simple formula that makes it possible to derive the column of the successive numbers \( l_x \) of survivors.
DEMOGRAPHY – Vol. II - Multistate Demography - Jacques Ledent, Yi Zeng

\[ l_{x+n} = p_x l_x \]

where \( p_x = \frac{1 - \frac{n}{2} m_x}{1 + \frac{n}{2} m_x} \) is the probability to survive (survival probability) from age \( x \) to age \( x + n \). Note that this simple formula is based on the assumption that the deaths occurring between ages \( x \) and \( x + n \) are uniformly distributed. This is a reasonable approximation except at infancy (between birth and age one) and extremely advanced ages (e.g. >age 95), which needs some special estimation procedures (Ref. Section 2 of Demographic Models and Actuarial Science and/or other standard Demography text book).

Such an ordinary single-state life table is valid only for analysis of non-renewable events such as death or, in the case of first marriage or first birth. But it can be readily extended to the multiple decrement life table by recognizing two or more exits (decrements). These decrements may reflect various causes of death, thus leading to a cause-of-death life table, the main interest of which resides in its ability to evaluate the impact of the elimination of a particular cause of death on life expectancy, assuming that the various causes of death are independent of each other. Alternatively, the decrements may be death and another non renewable event, such as is the case in a net primo-nuptiality table which describes the evolution of a never-married cohort of individuals that can exit through not only death but also marriage.

Both the ordinary single-state life table and the multiple-decrement life table do not allow re-entering a previously occupied state which is precisely the distinctive feature of multistate life tables.

2.2. The Multistate Life Table Functions

Basically, a multistate life table, sometimes referred to as a multidimensional life table, describes the life course of a cohort of individuals born in a given period of time as they move between the various states considered until the death of the last survivor. It is typically constructed in the context of marital status (with individuals moving between the four traditional marital statuses--never-married, married, divorced and widowed--or more realistically between seven marital and union statuses that take into account cohabitation), interregional migration (with individuals moving between a set of regions) or labor activity (with individuals shifting between not being employed and being employed, the former status being possibly divided further according to whether one belongs to the labor force or not).

Let us assume for the time being that the cohort under consideration consists of individuals born in just one state, e.g., the never married state of a multistate life table focusing on marital status. Drawing on a parallel with the ordinary single-state life table, it is possible to associate with each state \( i (i = 1,2,\ldots N) \) an elementary (state-specific) table that depicts the mortality and mobility experience in state \( i \) of the cohort under consideration. First, there are the column of the numbers \( l_i' \) of those alive at age \( x \) in state \( i \) and the column of the numbers \( L_i' \) of person-years lived in state \( i \) between
ages $x$ and $x+n$. Next, there is the column of the cumulated numbers $T_x^i$ of person-years lived in state $i$ beyond age $x$ ($T_x^i = \sum_{y\geq x} L_y^i$) which, it should be noted, are lived by individuals who at age $x$ were not only in state $i$ but also in all of the other states. As a result, the next column contains the numbers $e'_x^i$ of remaining lifetime to be spent in state $i$ beyond age $x$, regardless of the state occupied at age $x$ ($e'_x^i = T_x^i / \sum_k t_k^i$).

As in the ordinary single-state life table, a column may be included to represent the numbers $d_x^{i\delta}$ of deaths in state $i$ (the $\delta$ symbol is added here to stress exits through death). Moreover, $(N-1)$ additional columns may be included as well to reflect the other types of exits, the $k$th column of which contains the numbers of moves $d_x^{ik}$ to the $k$th state between ages $x$ and $x+n$. At the same time, of course, these exits constitute entries into the states to which these moves are directed. As a result, linking the number $l_{x+n}^i$ of survivors in state $i$ at age $x+n$ with the corresponding number $l_x^i$ at age $x$ requires one to subtract the deaths $d_x^{i\delta}$ in state $i$ and the moves $d_x^{ik}$ out of state $i$ to all other states but also to add the moves $d_x^{ki}$ into state $i$ from all other states:

$$l_{x+n}^i = l_x^i - d_x^{i\delta} - \sum_{k\neq i} d_x^{ik} + \sum_{k\neq i} d_x^{ki}$$ (1)

Thus, any given multistate life table consists of $N$ elementary tables (one for each state) such as the one just described. In practice, it reflects a particular demographic pattern embodied in a set of age-specific mortality and mobility rates defined for each state, respectively, by:

$$m_x^{i\delta} = \frac{d_x^{i\delta}}{L_x^i}$$ (2a)

and

$$m_x^{ik} = \frac{d_x^{ik}}{L_x^i} \quad \text{for } k \neq i$$ (2b)

From there, drawing on the exposition of the ordinary single-state life table, the number $L_x^i$ of person-years lived can be approximated by the arithmetic mean of the number of survivors at ages $x$ and $x+n$ times the width of the age interval:

$$L_x^i = \frac{n}{2} \left( l_x^i + l_{x+n}^i \right)$$ (3)

Combining (1)-(3) gives rise to a system of linear equations involving as many
variables—the numbers of survivors $l'_{i}$, out of the initial cohort, that reside at age $x$ in each state $i$—as there are states considered—that is, $N$. Such a system, however, is not tractable using conventional calculus, thus leading one to have recourse to matrix calculus. Basically, this calls for arranging the age-specific numbers $l'_{i}$ of survivors in a vertical array or vector

$$
\ell_{x}' = \begin{bmatrix}
    l_{1}' \\
    l_{2}' \\
    \vdots \\
    l_{N}'
\end{bmatrix}
$$

in which the $i^{th}$ element is $l'_{i}$ and embedding the age-specific mortality and mobility rates in a square array or matrix:

$$
m_{x} = \begin{bmatrix}
    m_{1x}^{1} + \sum_{l=1}^{N} m_{l}^{1} & -m_{2x}^{1} & \cdots & -m_{N}^{1} \\
    -m_{1x}^{2} & m_{2x}^{2} + \sum_{l=2}^{N} m_{l}^{2} & \cdots & \vdots \\
    \vdots & \vdots & \ddots & \vdots \\
    -m_{1x}^{N} & \cdots & m_{N}^{N} + \sum_{l=N}^{N} m_{l}^{N}
\end{bmatrix}
$$

in which each $k^{th}$ diagonal element contains the total rate of exit out of state $k$—that is, the mortality rate $m_{k}^{k}$ plus the sum $\sum_{l=k}^{N} m_{l}^{l}$ of the rates of moving to another state—and each $kl^{th}$ off-diagonal element contains the rate of moving from state $l$ to state $k$ preceded by a minus sign. Consequently, the system of linear equations can be written in a more compact format as (Rogers and Ledent, 1976):

$$
\ell_{x+1}' = p_{x} \ell_{x}'
$$

(4)

in which

$$
p_{x} = \left( I + \frac{n}{2} m_{x} \right)^{-1} \left( I - \frac{n}{2} m_{x} \right)
$$

(5)

is a transition probability matrix which is structurally similar to the single-state survival probability $p_{x}$ mentioned earlier, except that a matrix of mortality/mobility rates is substituted for a scalar mortality rate.

Equation (4) enables one to derive the successive vectors $\ell'_{x}$ of survivors and from there
the associated vectors of person-years lived \((L_x, T_x\) and \(e_x\)).

Let us now turn to the situation in which individuals are born in several, if not all of the states, rather than in just one state as was assumed in the above discussion. In such a case typical of a multistate life table focusing on migration between a set of regions, or a multiregional life table, the above framework applies to each state (region) of birth. Equation (4) can then be written in relation to each state-of-birth-specific cohort \(\ell^j\) of individuals, thus prompting the addition, in front of \(L_x\) and \(T_x\), of a \(j\) subscript representing the state of birth:

\[
j^j\ell_{x+n} = p_j^j L_x^j
\]

(6)

Then on gathering all of the \(j L_x\) vectors in a matrix

\[
I_x = [l_1 L_x, l_2 L_x, \ldots, n L_x]
\]

it follows that the series of the \(I_x\) can be obtained on the basis of

\[
I_{x+n} = p_x I_x
\]

(7)

Similarly, a region-of-birth-specific index can be assigned to the \(L_x\) and \(T_x\) vectors and the ensuing vectors can be gathered in square matrices, respectively, \(L_x\) and \(T_x\), which then allow for the derivation of an alternative type of life expectancies—namely, status-based life expectancies on the basis of:

\[
e_x = T_x L_x^{-1}
\]

(8)

where \(e_x\) is a matrix such that its \(kl\)th element expresses the remaining lifetime in region \(k\) to an individual RESIDING in state \(l\) at age \(x\). [But such a formula may be applied just as well in the case of a multistate life table originating from a single-state cohort. Suffices it to consider along hypothetical cohorts “born” in all of the other states].

TO ACCESS ALL THE 25 PAGES OF THIS CHAPTER, Visit: http://www.eolss.net/Eolss-sampleAllChapter.aspx

Bibliography

Methods in Medical Research, 11: 91-115 [provide an overview of how multistate models often provide a relevant modeling framework for event history analysis].


Schoen, R. (1988). *Modeling Multigroup Population*. Plenum Press, New York [dealing with models that can capture the behavior of individuals and groups over time, this book considers a broad range of multistate population models, showing that such models can be calculable, clearly interpretable, and analytically valuable for studying many kinds of social behavior].


Zeng, Y. (1991). *Family Dynamics in China: A Life Table Analysis*. The University of Wisconsin Press, Wisconsin [Presenting the methodology and application (to China) of the multistate family status life table model which included both nuclear and three-generation family households].

Zeng, Y., Gu, D. and Land, K. C. (2004). ‘A new method for correcting underestimation of disabled life expectancy and application to Chinese oldest-old’, *Demography*, 41(2): 335-361 [Demonstrating that disabled life expectancies that are based on conventional multistate life-table methods are significantly underestimated because of the assumption of no changes in functional status between age x and death occurred between ages x and x+n, and presenting a new method to correct the bias and apply it to data from a longitudinal survey of about 9,000 oldest-old Chinese aged 80-105 collected in 1998 and 2000].


Biographical Sketches

Jacques Ledent was active in the initial development of multistate demography at a time when he was associated with the International Institute for Applied Analysis (IIASA) located near Vienna (Austria). He is currently Research Professor at the Institut national de la recherche scientifique (INRS) which is part of the University of Quebec. After a long standing focus on internal migration, his research is now concerned with international migration and, more specifically, the integration of immigrants in Montreal, Quebec and Canada.

Yi Zeng is a Professor at the Center for the Study of Aging and Human Development and Geriatric Division / Dept of Medicine of Medical School, and Institute of Population Research and Dept. of Sociology, Duke University. He is also a Professor at the China Center for Economic Research, National School of Development at Peking University in China, and Distinguished Research Scholar of the Max Planck Institute for Demographic Research (MPIDR) in Germany. He received his doctoral degree from Brussels Free University in May 1986, and conducted post-doctoral study at Princeton University in 1986-87. Up to July 15, 2009, he has had 96 professional articles written in English published in academic journals or as book chapters in the United States and Europe; among them, 60 articles were published in anonymously peer-reviewed academic journals. He has had 89 professional articles written in Chinese and published in China; among them, 61 articles were published in national Chinese academic journals. He has published nineteen books, including 6 sole-author books, 3 co-author books, 7 chief-editor books and 3 co-editor books. Eight of Yi Zeng’s published books were written in English (including three by Springer Publisher and one by the University of Wisconsin Press), one was written in both Chinese and English, and the remainders were written in Chinese. Yi Zeng has been awarded eleven national academic prizes and three international academic prizes, such as the Dorothy Thomas Prize of the Population Association of America, the Harold D. Lasswell Prize in Policy Science awarded by the international journal Policy Sciences and Kluwer Academic Publishers, the national prizes for advancement of science and technology awarded by the State Sciences and Technology Commission of China and the State Education Commission, the highest academic honor of Peking University: "Prize for Outstanding Contributions in Sciences," and the "Chinese Population Prize (Science and Technology)", jointly awarded by nine ministries and seven national non-governmental associations in China.