BULLETIN *of the* MALAYSIAN MATHEMATICAL SCIENCES SOCIETY

Bull. Malaysian Math. Sc. Soc. (Second Series) **26** (2003) 101–108

A Birth-Death Process Approach to Constructing Multistate Life Tables

M. ATAHARUL ISLAM

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Pulau Pinang, Malaysia
e-mail: ataharul@cs.usm.my

Abstract. The multistate life tables have generated interest among researchers during the past two decades. One of the main reasons for such interest is the wide range of applications addressing real life problems. A simplified version of constructing multistate life tables is developed in this paper on the basis of a modified birth-death process approach. This approach remains simple even if the number of intercommunicating states increases. An estimation procedure is also discussed.

1. Introduction

The origin of the multistate life table can be traced back to 1912 (Du Pasquier, 1912) in the context of disability insurance. Fix and Neyman (1951) studied recovery, relapse, death, and loss of cancer patients using a discrete state continuous time Markov process. Sverdrup (1965) elaborated on the statistical estimation and test procedures for this approach. Sheps and Perrin (1964) formulated family building as a Markov renewal process. In this approach, transition rates are specified by the duration in a particular state, whereas the transition rates are given by age of the individual. Since then, Hoem (1971), Hoem and Fong (1976), Rogers (1973, 1976), Schoen and Nelson (1974), Schoen (1975), Rogers and Ledent (1976), Schoen and Land (1979), Schoen (1979, 1988), Krishnamoorthy (1979), Land and Schoen (1982) have contributed to the simplification of the theory and construction of multistate life tables with applications to various problems.

In this paper, a new procedure is proposed for calculating multistate life tables that is applicable to cohorts followed over time and which also easily treats a large number of states. The proposed approach is based on a birth and death process for each state, where birth is an entry from any one of the communicating states or a transient state. We then extend the method so that it is applicable to period data. To do so, since the entry and exit rates from the state are equivalent to in- and out-migration rates, we propose an iterative procedure that enables us to avoid the trap of using net migration rates that depend on the distribution of individuals over the other states in the model. Finally, a method is described for adjusting the cohort tables for censoring and other types of exits from the study.

2. Definitions

For the multistate life table, we can consider a system of r non-absorbing states and an absorbing state, $(r+1)$. By analogy with single decrement life table functions we define:

$d^{ij}(x)$ = number of moves from state i to state j $(i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, r+1)$ in the interval $(x, x+dx)$ among those surviving in the ith state at age x;

$l^i(x)$ = exposure of the individuals in state i at age x to the risk of moving before reaching age $x+dx$, $(i = 1, 2, \cdots, r)$;

$\mu^{ij}(x) = \lim\limits_{dx \to 0} d^{ij}(x) \big/ l^i(x)dx$ = force of movement from state i to state j $(i = 1, 2, \cdots, r;\ \ j = 1, 2, \cdots, r+1;\ i \neq j)$.

Then the equation indicating decrements from and increments to the survivors in state i at age $x+dx$ can be written as follows:

$$l^i(x+dx) = l^i(x) - \sum_{j=1}^{r+1} d^{ij}(x) + \sum_{j=1}^{r} d^{ji}(x), \quad i = 1, 2, \cdots, r;\ \ j \neq i. \tag{1}$$

Substituting

$$d^{ij}(x) = \mu^{ij}(x)\, l^i(x)\, dx$$

in (1), we obtain

$$l^i(x+dx) = l^i(x) - \sum_{j=1}^{r+1} \mu^{ij}(x) l^i(x)dx + \sum_{j=1}^{r} \mu^{ji}(x) l^j(x)dx, \quad j \neq i, \tag{2}$$

or, in matrix form

$$\{l(x+dx)\} = \{l(x)\} - \{\mu(x)\}\{l(x)\}\, dx \tag{3}$$

where $\{l(x)\}$ is an $(r \times 1)$ matrix with elements $l^i(x)$ and $\{\mu(x)\}$ is an $r \times r$ matrix in which the (i, j)-th element is $-\mu^{ij}(x)$ for $i \neq j$ and the diagonal elements are the total force of decrements from state i, $\mu^{ii} = \sum_{j=1}^{r+1} \mu^{ij}$, $j \neq i$.

3. Life tables as birth-death processes

Chiang (1968, Ch.10) treated the single decrement life table as a pure death process and derived the probability distribution of the number of survivors at any time t. This paper shows that the birth and death process approach can be extended to multistate life tables. The notations and definitions in Chiang (1980, Ch. 10) have been used in this paper.

3.1. Multistate life tables

We first consider r intercommunicating states and one absorbing (dead) state. We focus on one particular intercommunicating state, say i, and let $l^i(0)$ denotes the radix of state i. During any time (or age) interval, $l^i(x)$ may increase due to movements from other states to state i or decrease due to movements out of state i. Therefore, the number of survivors in state i can be treated as a birth and death process where (i) the births are all the increments to that state from other states, and (ii) the deaths are all the decrements from that state due to movements to other states. Although we will treat each state as a separate process, we recognize that they are linked through the rates of transition from one state to another.

As before, the force of decrement from state i to state j is

$$\mu^{ij}(x) = \lim_{dx\to 0} d^{ij}(x)\big/l^i(x)\,dx\,, \quad i = 1, 2, \cdots, r\,;\ j = 1, 2, \cdots,\ r+1\,,\ j \neq i\,; \tag{4}$$

and the overall force of decrement, or force of mortality from state i is

$$\mu^i(x) = \sum_{j=1}^{r+1} \mu^{ij}(x)\ , \quad i = 1,2 \cdots,\ r\,;\ \ j \neq i\,. \tag{5}$$

The force of increment from state j to state i is the force of decrement from state j to state i;

$$\mu^{ji}(x) = \lim_{dx\to 0} d^{ji}(x)\big/l^i(x)\,dx\,. \tag{6}$$

It is, however, not directly analogous to a birth rate for these individuals, since its denominator is $l^j(x)\,dx$. The births should occur to the exposed population in state i, hence the denominator needs to be $l^i(x)dx$ instead of $l^j(x)\,dx$. To cope with this problem, we adjust the force of increment by using a multiplier, $l^j(x)dx\,/\,l^i(x)dx$. Thus we obtain,

$$\lambda^{ji}(x) = \lim_{dx\to 0}\left(d^{ji}(x)\big/l^j(x)\ dx\right)\left(l^j(x)\,dx\big/l^i(x)\ dx\right) = \lim_{dx\to 0} d^{ji}(x)\big/l^i(x)\,dx \tag{7}$$

which can be defined as the force of increment to the $l^i(x)$ individuals during $(x, x+dx)$. The combined force of increment to state i is

$$\lambda^i(x) = \sum_{j=1}^{r} \lambda^{ji}(x)\ , \quad i = 1,2\cdots,r\,;\ \ j \neq i\,. \tag{8}$$

which can be used as the force of increment to, or the birth rate, in state *i*.

3.2. A modified birth-death process

We begin with a vector $\{l(0)\}$ which has $l^j(0)$ individuals born into state j $(j = 1, 2, \cdots, r)$ and follow those in one state, *i*, through successive ages until all die or move permanently to other states. Then Chiang's results ((1980, Ch. 10) for the birth-death process hold for state *i*. Letting the probability that there are *n* surviving individuals at age x be denoted by $P_n^i(x) = P[l^i(x) = n]$, we have the system of differential equations:

$$\begin{aligned}
&\frac{d}{dx} P_0^i(x) = \mu^i(x) P_1^i(x) \\
&\frac{d}{dx} P_n^i(x) = -n\left\{\lambda^i(x) + \mu^i(x)\right\} P_n^i(x) + (n-1)\lambda^i(x)\, P_{n-1}^i(x) + (n+1)\mu^i(x) P_{n+1}^i(x)
\end{aligned} \tag{9}$$

Letting the net change and the cumulative net change be

$$\begin{aligned}
&\gamma^i(\tau) = \mu^i(\tau) - \lambda^i(\tau)\,, \text{ and} \\
&\Gamma^i(x) = \int_0^x \left[\mu^i(\tau) - \lambda^i(\tau)\right] d\tau = \int_0^x \gamma^i(\tau)\, d\tau\,,
\end{aligned} \tag{10}$$

and using the notation $l^i(0) = l_0^i$, it can be shown that for $n \geq 1$

$$P_n^i(x) = \sum_{u=0}^{\min(l_0^i, n)} \binom{l_0^i}{u} \binom{l_0^i + n - u - 1}{n - u} \left[\alpha^i(x)\right]^{l_0^i - n} \left[\beta^i(x)\right]^{n-u} \left[1 - \alpha^i(x) - \beta^i(x)\right]^u$$

and for $n = 0$

$$P_0^i(x) = \left[\alpha^i(x)\right]^{l_0^i}$$

where

$$\alpha^i(x) = 1 - 1 \Bigg/ \left\{ \exp\left(\Gamma^i(x)\right) + \int_0^x \lambda^i(\tau) \exp(\tau)\, d\tau \right\}$$

and

$$\beta^i(x) = 1 - \exp\left\{\Gamma^i(x)\right\}\left\{1 - \alpha^i(x)\right\}.$$

Furthermore,

$$E\left[l^i(x) / l_0^i\right] = l_0^i \exp\left\{-\Gamma^i(x)\right\} = l_0^i \exp\left\{-\int_0^x \left[\mu^i(\tau) - \lambda^i(\tau)\right] d\tau\right\}. \qquad (11)$$

Before going on to the estimation procedures for this formulation of a multistate life table, it is useful to examine its relationship to the standard approach. As shown in equation 3, for any interval $(x, x + dx)$,

$$\left\{l(x + dx)\right\} = \left\{l(x)\right\} - \left\{\mu(x)\right\}\left\{l(x)\right\} dx \qquad (12)$$

where $\{l(x)\}$ is the vector of numbers in each state at time x and $\{\mu(x)\}$ is the matrix of transition rates. A similar expression can be obtained from the proposed approach. We have defined

$$\gamma^i(x) = \sum_{j=1}^{r+1} \mu^{ij}(x) - \sum_{j=1}^{r} \lambda^{ji}(x). \qquad (13)$$

If

$$\left\{\gamma(x)\right\} = \begin{bmatrix} \gamma^1(x) & 0 & \ldots\ldots\ldots & 0 \\ 0 & \gamma^2(x) & \ldots\ldots\ldots & 0 \\ & & & \\ 0 & 0 & \ldots\ldots\ldots & \gamma^r(x) \end{bmatrix},$$

then

$$\left\{l(x + dx)\right\} = \left\{l(x)\right\} - \gamma(x)\left\{l(x)\right\} dx. \qquad (14)$$

This expression is the same as equation 3 because

$$\mu(x)\left\{l(x)\right\} = \gamma(x)\left\{l(x)\right\} \text{ although } \mu(x) \neq \gamma(x). \qquad (15)$$

3.3. Estimation

In order to develop an estimation procedure, we subdivide the interval $(0, x)$ into short segments, e.g., $(0, x_1), (x_1, x_2), \cdots, (x_{k-1}, x_k)$. Let t_m be the mid-value of the interval (x_{m-1}, x_m), $m = 1, 2, \cdots, k$ and

$$\Delta(t_m) = x_m - x_{m-1} .$$

Then

$$E\left[l^i(x) / l_0^i\right] = l_0^i \exp\left\{-\int_0^{x_1} \gamma^i(\tau)\, d\tau\right\} \ldots \exp\left\{-\int_{x_{k-1}}^{x_k} \gamma^i(\tau)\, d\tau\right\} \tag{16}$$

$$= l_0^i \exp\left\{-\gamma^i(t_1)\,\Delta t_1\right\} \ldots \exp\left\{-\gamma^i(t_k)\,\Delta t_k\right\}$$

$$= l_0^i \exp\left\{-\sum_{m=1}^{k} \gamma^i(t_m)\,\Delta t_m\right\} . \tag{17}$$

If the $\gamma^i(t)$'s are known, then we can compute the expected number of survivors in state i at age $x (i = 1, 2, \cdots, r)$. We will denote these estimates, based on a discrete approximation, by l_x. Islam and Jalil (1995) indicated the maximum likelihood estimates for the expected number of survivors in a particular state *i*.

The number of person years lived in the ith state during the age interval $(x, x+n)$ can be obtained by assuming the linearity of the survivorship function

$$L_x^i = {n}/{2}\left(l_x^i + l_{x+n}^i\right), \quad i = 1, 2, \cdots, r . \tag{18}$$

Similarly, the total number of person years lived by individuals in state *i* beyond age *x* is

$$T_x^i = \sum_{y=x}^{w} L_y^i . \tag{19}$$

It should be noted that, if the radix for state *i* is zero and no one enters state *i* until age *x*, then instead of considering l_0^i we can estimate l_x^i on the basis of the entries into state *i* from other states. The number of survivors beyond age *x* in state *i* can then be obtained as before, treating l_x^i as the radix for state *i* starting at time (or age) *x*.

We may approximate λ_x^{ji} by the usual life table central rates and person-years in states *j* and *i*, which we assume are equal to the observed transition rates and mid-year populations, i.e. we assume:

$$\lambda_x^{ji} = m_x^{ji} \,.\, \left(L_x^j / L_x^i\right) = M_x^{ji} \,.\, \left(K_x^j / K_x^i\right) \tag{20}$$

where for the age interval $(x, x+1)$, m_x^{ji} is the life table transition rate from state *j* to state *i*, M_x^{ji} is the corresponding observed rate, L_x^j is the life table person-years lived in state *j*, and K_x^j is the observed mid-year population in state *j*. M_x^{ji} may be calculated from the observed number of moves from state *j* to state *i* in the interval $(x, x+1)$, D_x^{ij}, and the observed mid year population; the observed rate corresponding to λ_x^{ji} can then be computed:

$$M_x^{ji} = D_x^{ji} \big/ K_x^j \qquad N_x^{ji} = M_x^{ji} \cdot K_x^j \big/ K_x^i \quad i = 1, 2, \cdots, r; \; i \neq j. \tag{21}$$

The expected number of survivors at age $x+n$ for given *x*, following equation (11), for state *i* is

$$E\left[l^i(x+n)/l_x^i\right] = l_x^i \; \exp\left\{-\int_x^{x+n}\left[\mu^i(\tau) - \lambda^i(\tau)\right] d\tau\right\} \quad i = 1, 2, \cdots, r \tag{22}$$

can be approximated as

$$l_{x+n}^i = l_x^i \exp\left\{-n\left[\sum_{j=1}^{r+1} M_x^{ij} - \sum_{j=1}^{r} N_x^{ji}\right]\right\} \quad i = 1, 2, \cdots, r. \tag{23}$$

We note that this approach avoids some of the problems of the standard multistate life table calculations that result from ignoring higher order terms in approximations and the singularity of the matrix of observed transition rates. Because we are estimating the expected values in each state separately, the sum over all states may change slightly from one time period to the next. Some checks showed that the deviations are minor.

4. Conclusion

The estimation of transition probabilities from a set of longitudinal data, emerging from a large number of intercommunicating states, poses formidable difficulty in constructing multistate life tables. It becomes increasingly difficult to estimate the number of survivors in each state at different times or ages for a large number of intercommunicating states. A simplification is proposed in this paper on the basis of a modified birth-death process approach. It is evident that the estimation of the number of survivors remains very simple even with an increase in the number of states.

Acknowledgement. I am greatly indebted to Professor Chai Bin Park and Dr. R.D. Retherford for their helpful comments on a previous version of the paper. Professor Jane Menken and Professor S. Krishnamoorthy provided useful suggestions which improved the exposition of the paper. Finally, I would like to thank Professor Quah Soon Hoe for his useful comments on the paper.

References

1. C.L. Chiang, *Introduction to Stochastic Processes in Biostatistics*, Wiley, New York, 1968.
2. C.L. Chiang, *An Introduction to Stochastic Processes and Their Applications*, Wiley, New York, 1980.
3. L.G. Du Pasquier, Mathematische theorie der invalidatatsversicherung, *Mittilun. Verein. Schwei. Versich. Math.* **8** (1912), 1–53.
4. E. Fix and J. Neyman, A simple stochastic model of recovery, relapse, death and loss of patients, *Human Biol.* **23** (1951), 205–241.
5. J.M. Hoem, Point estimation of forces of transition in demographic models, *J. Royal Stat. Soc. **B*** **33** (1971), 275–289.
6. J.M. Hoem and M.S. Fong, A Markov Chain Model of Working Life Tables, *Working Paper No. 2, University of Copenhagen, Laboratory of Actuarial Mathematics* (1976).
7. J.M. Hoem and U.F. Jensen, Multistate life table methodology: a probabilist critique, In *Multidimensional Mathematical Demography*, K.C. Land and A. Rogers (Editors), Academic Press, New York (1982), 155–264.
8. M.A. Islam and M.A. Jalil, Expected number of survivors in intercommunicating states, *J. Stat. Res.* **29** (1995), 41–50.
9. S. Krishnamoorthy, Classical approach to increment-decrement life tables: an application to the marital status of United States females, 1970, *Math. Biosc.* **44** (1979), 139–154.
10. K.C. Land and R. Schoen, Statistical methods for Markov-generated increment–decrement life tables with polynomial gross flow functions, In *Multidimensional Mathematical Demography*, K.C. Land and A. Rogers (Editors), Academic Press, New York (1982), 265–346.
11. A. Rogers, The multiregional life table, *J. Math. Soc.* **3** (1973), 127–137.
12. A. Rogers, *Introduction to Multiregional Mathematical Demography*, Wiley, New York, 1975.
13. A. Rogers and J. Ledent, Increment–decrement life tables: a comment, *Demography* **13** (1976), 287–290.
14. R. Schoen, Constructing increment-decrement life tables, *Demography* **12** (1975), 313–324.
15. R. Schoen, Calculating increment-decrement life tables by estimating mean duration at transfer from observed rates, *Math. Biosc.* **47** (1979), 255–269.
16. R. Schoen, *Modelling Multigroup Populations*, Plenum Press, New York, 1988.
17. R. Schoen and K.C. Land, A general algorithm for a Markov-generated increment-decrement life table with applications to marital status pattern, *J. Am. Stat. Assoc*. **74** (1979), 761–776.
18. R. Schoen and V.E. Nelson, Marriage, divorce, and mortality: a life table analysis, *Demography* **11** (1974), 267–290.
19. M.C. Sheps and E.B. Perrin, The distribution of birth intervals under a class of stochastic fertility models, *Pop. Stud.* **17** (1964), 321–331.
20. E. Sverdrup, Estimates and test procedures in connection with stochastic models for death, recoveries, and transfers between different states of health, *Skand. Aktuarie.* **40** (1965), 184–211.

Keywords phrases: expected number of survivors, survival function, transitions, intercommunicating states.

2000 Mathematics Subject Classification: 60K20