JIA 114 (1987) 569-580

A SHORT NOTE ON THE CONSTRUCTION OF LIFE TABLES AND MULTIPLE DECREMENT TABLES

H. R. WATERS, M.A., D.Phil., F.I.A., AND A. D. WILKIE, M.A., F.F.A., F.I.A.

THE purpose of this short note is to explain a simple method of constructing a life table, given an explicit function for the force of mortality, μ_x, for all x. A similar method can be applied to the construction of multiple decrement tables, given explicit functions for the separate forces of transition, μ_x^b, μ_x^c, etc, from one 'active' status (a) to various 'dead' statuses (b, c, etc) for all x.

Although numerical results for the examples discussed in this note could be calculated equally well by more direct methods than the ones described herein, results for more complicated examples involving multiple status tables, in which transition back and forth between statuses is possible, cannot in general be calculated by such direct methods, and our methods provide a practical way by means of which numerical results for these more complicated cases can be calculated.

The motivation for writing this note is therefore to explain by means of simple examples the methodology that has been used for certain more complicated cases, such as the models being worked on by the authors for the Permanent Health Insurance Subcommittee of the Continuous Mortality Investigation Bureau, which will be published when further work is completed, and which have been foreshadowed by Waters (1984). The same methodology has also been used for a model of marital status and mortality, described in Wilkie (forthcoming), and for models to describe mortality and sickness from the disease AIDS, described in Wilkie (1987) and by the Institute of Actuaries Working Party on AIDS (1987).

1. LIFE TABLES

The problem for an ordinary life table can be expressed as an elementary problem in numerical analysis. Let us suppose that we know the force of mortality, μ_x, for all x, for a mortality table, and that we wish to calculate numerically the value of ${}_tp_{x_0}$ for some initial age x_0 and various values of t. Throughout this note we shall assume that the initial age x_0 is fixed and for the typographical convenience we shall denote the probability of survival to age x ($\geqslant x_0$) as $pa(x)$. We shall denote by $pd(x)$ the probability of death before age x for an individual who was alive at the initial age x_0. We assume that the function μ_x is mathematically suitably well-behaved. The necessary constraints on the form of μ_x are not very severe, the only one of practical importance being that it be bounded.

For any age $x \geqslant x_0$ we therefore have

$$pa(x)+pd(x)=1 \tag{1}$$

We also know the differential equation:

$$\mu_x = \frac{-dl_x}{dx}\frac{1}{l_x} \tag{2}$$

which describes the relationship between μ_x and l_x of the life table. We re-express this as

$$\mu_x = \frac{-dpa(x)}{dx}\frac{1}{pa(x)} \tag{3}$$

or

$$\frac{dpa(x)}{dx} = -\mu_x \cdot pa(x) \tag{4}$$

We also know that

$$pa(x_0) = 1 \tag{5}$$

and this gives us an initial condition. We now have sufficient information in principle to calculate $pa(x)$ for all $x \geqslant x_0$. The problem in practice is how to calculate numerical values of $pa(x)$, say at regular intervals of x perhaps one year apart.

In numerical analysis terminology this is an 'initial value' problem. See, for example, Conte and de Boor (1980).

There are several ways of solving such a problem. The method we use is based on 'predictor/corrector' methods, but with the advantage that in this simple case we only require the 'predictor' step, and not the 'corrector' step, so that we can go directly to a solution for each step, rather than having to use successive approximation.

The key lies in using the mean value theorem, from which we know that

$$\frac{pa(x+h)-pa(x)}{h} = \frac{dpa(x)}{dx}\bigg|_{x=y} = -\mu_y \cdot pa(y) \tag{6}$$

for some y between x and $(x+h)$. We re-express this as

$$pa(x+h)-pa(x) = -h \cdot \mu_y \cdot pa(y) \tag{7}$$

The intuitive interpretation of (7) is of some interest: $pa(x+h)$ is the probability of being alive at age $(x+h)$ (conditional always on being alive at x_0) and this is the probability of being alive at age x, $pa(x)$, minus the probability of dying between ages x and $(x+h)$. This probability of dying is $h \cdot \mu_y \cdot pa(y)$ so we also have an expression connecting $pd(x)$ and $pd(x+h)$:

$$pd(x+h)-pd(x) = +h \cdot \mu_y \cdot pa(y) \tag{8}$$

We do not know the value of y for which (7) is exactly true, but we can approximate the right hand side of (7) in a number of ways in order to complete the solution. We could put:

$$\text{(A) } h \cdot \mu_y \cdot pa(y) = h \cdot \{\mu_x \cdot pa(x) + \mu_{x+h} \cdot pa(x+h)\}/2 \tag{9A}$$

so that from (7) and (9A) we have:

$$pa(x+h) = pa(x) \cdot \frac{\{1 - h/2 \cdot \mu_x\}}{\{1 + h/2 \cdot \mu_{x+h}\}} \tag{10A}$$

Alternatively we could put:

$$\text{(B) } h \cdot \mu_y \cdot pa(y) = h \cdot \mu_{x+h/2} \cdot \{pa(x) + pa(x+h)\}/2 \tag{9B}$$

and from (7) and (9B) we have:

$$pa(x+h) = pa(x) \cdot \frac{\{1 - h/2 \cdot \mu_{x+h/2}\}}{\{1 + h/2 \cdot \mu_{x+h/2}\}} \tag{10B}$$

Finally we could put:

$$\text{(C) } h \cdot \mu_y \cdot pa(y) = h \cdot \{\mu_x + \mu_{x+h}\} \cdot \{pa(x) + pa(x+h)\}/4 \tag{9C}$$

and from (7) and (9C) we have:

$$pa(x+h) = pa(x) \cdot \frac{\{1 - h/4 \cdot (\mu_x + \mu_{x+h})\}}{\{1 + h/4 \cdot (\mu_x + \mu_{x+h})\}} \tag{10C}$$

In each case we can now calculate $pa(x)$ recursively using (10A), (10B) or (10C) and the initial condition (5), provided that we choose a step size h that is small enough. The value of h must be less than $2/\mu_x$ for all relevant values of x.

We have calculated $pa(x)$ up to age 110 using each of the three approximations for different values of h, in each case using an initial age of 20 and the following formula for μ_x:

$$\mu_x = a_0 + a_1 \cdot (x-20) + \exp\{b_0 + b_1 \cdot (x-20)\} \tag{11}$$

with $a_0 = \ \cdot 0005$
$a_1 = -\cdot 0001$
$b_0 = -7\cdot 6$
$b_1 = \ \cdot 09$

This is the first modification of Makeham's law, and it has the advantage that the values of $pa(x)$ can be calculated directly by

$$pa(x) = \exp\{-\int_{x_0}^{x} \mu_y \, dy\} \tag{12}$$

where

$$\int_{x_0}^{x} \mu_y \, dy = a_0 \cdot t + a_1 \cdot t^2/2 + \{\exp(b_0 + b_1 \cdot t) - \exp(b_0)\}/b_1 \tag{13}$$

with $t = x - x_0$, and (for these calculations) $x_0 = 20$.

Table 1 summarizes the results of these calculations. In each part of Table 1, '$A(h)$' refers to the approximate value of $pa(x)$ calculated using (10a), (10b) or (10C) and a step size of h, 'E' refers to the exact value of $pa(x)$, calculated from (12) and (13), and the maxima are taken over integer values of x from 20 to 110.

Table 1. *Numerical calculation of life table probabilities*

Approximation (A)

h	$\mathrm{Max}\lvert A(h)-E\rvert$	At age	$\mathrm{Max}\lvert (A(h)-E)/E\rvert$	At age
1	$1{\cdot}4\times 10^{-4}$	85	$7{\cdot}4\times 10^{-1}$	110
1/2	$3{\cdot}6\times 10^{-5}$	85	$2{\cdot}4\times 10^{-1}$	110
1/4	$8{\cdot}9\times 10^{-6}$	85	$6{\cdot}3\times 10^{-2}$	110
1/8	$2{\cdot}2\times 10^{-6}$	85	$1{\cdot}6\times 10^{-2}$	110
1/16	$5{\cdot}6\times 10^{-7}$	85	$4{\cdot}0\times 10^{-3}$	110
1/32	$1{\cdot}4\times 10^{-7}$	85	$1{\cdot}0\times 10^{-3}$	110
1/64	$3{\cdot}5\times 10^{-8}$	85	$2{\cdot}5\times 10^{-4}$	110
1/128	$8{\cdot}7\times 10^{-9}$	85	$6{\cdot}3\times 10^{-5}$	110
1/256	$2{\cdot}2\times 10^{-9}$	85	$1{\cdot}6\times 10^{-5}$	110
1/512	$5{\cdot}4\times 10^{-10}$	85	$3{\cdot}9\times 10^{-6}$	110

Approximation (B)

h	$\mathrm{Max}\lvert A(h)-E\rvert$	At age	$\mathrm{Max}\lvert (A(h)-E)/E\rvert$	At age
1	$2{\cdot}9\times 10^{-4}$	92	$8{\cdot}4\times 10^{-1}$	110
1/2	$7{\cdot}2\times 10^{-5}$	92	$3{\cdot}0\times 10^{-1}$	110
1/4	$1{\cdot}8\times 10^{-5}$	92	$8{\cdot}2\times 10^{-2}$	110
1/8	$4{\cdot}5\times 10^{-6}$	92	$2{\cdot}1\times 10^{-2}$	110
1/16	$1{\cdot}1\times 10^{-6}$	92	$5{\cdot}3\times 10^{-3}$	110
1/32	$2{\cdot}8\times 10^{-7}$	92	$1{\cdot}3\times 10^{-3}$	110
1/64	$7{\cdot}0\times 10^{-8}$	92	$3{\cdot}3\times 10^{-4}$	110
1/128	$1{\cdot}8\times 10^{-8}$	92	$8{\cdot}2\times 10^{-5}$	110
1/256	$4{\cdot}4\times 10^{-9}$	92	$2{\cdot}1\times 10^{-5}$	110
1/512	$1{\cdot}1\times 10^{-9}$	92	$5{\cdot}1\times 10^{-6}$	110

Approximation (C)

h	$\mathrm{Max}\lvert A(h)-E\rvert$	At age	$\mathrm{Max}\lvert (A(h)-E)/E\rvert$	At age
1	$4{\cdot}8\times 10^{-4}$	87	$8{\cdot}5\times 10^{-1}$	110
1/2	$1{\cdot}2\times 10^{-4}$	87	$3{\cdot}1\times 10^{-1}$	110
1/4	$3{\cdot}0\times 10^{-5}$	87	$8{\cdot}3\times 10^{-2}$	110
1/8	$7{\cdot}6\times 10^{-6}$	87	$2{\cdot}1\times 10^{-2}$	110
1/16	$1{\cdot}9\times 10^{-6}$	87	$5{\cdot}3\times 10^{-3}$	110
1/32	$4{\cdot}7\times 10^{-7}$	87	$1{\cdot}3\times 10^{-3}$	110
1/64	$1{\cdot}2\times 10^{-7}$	87	$3{\cdot}3\times 10^{-4}$	110
1/128	$3{\cdot}0\times 10^{-8}$	87	$8{\cdot}3\times 10^{-5}$	110
1/256	$7{\cdot}4\times 10^{-9}$	87	$2{\cdot}1\times 10^{-5}$	110
1/512	$1{\cdot}8\times 10^{-9}$	87	$5{\cdot}2\times 10^{-6}$	110

Three comments should be made about the numerical values in Table 1. The first is that they show the high accuracy of (10A), (10B) or (10C), in this particular example, even for reasonably large step sizes, for example $h = 1/8$. Even the values for $h = 1$ are accurate up to 3 significant figures at ages up to 95, and only diverge at higher ages where the values of μ_x become relatively large.

The second is that halving the step size has the effect of dividing both the maximum error, $\max\lvert A(h)-E\rvert$, and the maximum proportionate error, $\max\lvert A(h)-E)/E\rvert$, by a factor of 4; there are good theoretical reasons to support this observation, at least in the case of $\max\lvert A(h)-E\rvert$.

Table 2. *Comparison of calculations with successive values of* h

Approximation (A)

h	$\text{Max}\lvert A(h)-A(2h)\rvert$	At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$	At age
1/2	$1{\cdot}1\times10^{-4}$	85	$6{\cdot}6\times10^{-1}$	110
1/4	$2{\cdot}7\times10^{-5}$	85	$1{\cdot}9\times10^{-1}$	110
1/8	$6{\cdot}7\times10^{-6}$	85	$4{\cdot}8\times10^{-2}$	110
1/16	$1{\cdot}7\times10^{-6}$	85	$1{\cdot}2\times10^{-2}$	110
1/32	$4{\cdot}2\times10^{-7}$	85	$3{\cdot}0\times10^{-3}$	110
1/64	$1{\cdot}0\times10^{-7}$	85	$7{\cdot}5\times10^{-4}$	110
1/128	$2{\cdot}6\times10^{-8}$	85	$1{\cdot}9\times10^{-4}$	110
1/256	$6{\cdot}5\times10^{-9}$	85	$4{\cdot}7\times10^{-5}$	110
1/512	$1{\cdot}6\times10^{-9}$	85	$1{\cdot}1\times10^{-5}$	110

Approximation (B)

h	$\text{Max}\lvert A(h)-A(2h)\rvert$	At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$	At age
1/2	$2{\cdot}2\times10^{-4}$	92	$7{\cdot}8\times10^{-1}$	110
1/4	$5{\cdot}4\times10^{-5}$	92	$2{\cdot}4\times10^{-1}$	110
1/8	$1{\cdot}4\times10^{-5}$	92	$6{\cdot}2\times10^{-2}$	110
1/16	$3{\cdot}4\times10^{-6}$	92	$1{\cdot}6\times10^{-2}$	110
1/32	$8{\cdot}5\times10^{-7}$	92	$3{\cdot}9\times10^{-3}$	110
1/64	$2{\cdot}1\times10^{-7}$	92	$9{\cdot}9\times10^{-4}$	110
1/128	$5{\cdot}3\times10^{-8}$	92	$2{\cdot}5\times10^{-4}$	110
1/256	$1{\cdot}3\times10^{-8}$	92	$6{\cdot}2\times10^{-5}$	110
1/512	$3{\cdot}3\times10^{-9}$	92	$1{\cdot}5\times10^{-5}$	110

Approximation (C)

h	$\text{Max}\lvert A(h)-A(2h)\rvert$	At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$	At age
1/2	$3{\cdot}6\times10^{-4}$	87	$7{\cdot}8\times10^{-1}$	110
1/4	$9{\cdot}1\times10^{-5}$	87	$2{\cdot}4\times10^{-1}$	110
1/8	$2{\cdot}3\times10^{-5}$	87	$6{\cdot}3\times10^{-2}$	110
1/16	$5{\cdot}7\times10^{-6}$	87	$1{\cdot}6\times10^{-2}$	110
1/32	$1{\cdot}4\times10^{-6}$	87	$4{\cdot}0\times10^{-3}$	110
1/64	$3{\cdot}5\times10^{-7}$	87	$1{\cdot}0\times10^{-3}$	110
1/128	$8{\cdot}9\times10^{-8}$	87	$2{\cdot}5\times10^{-4}$	110
1/256	$2{\cdot}2\times10^{-8}$	87	$6{\cdot}3\times10^{-5}$	110
1/512	$5{\cdot}5\times10^{-9}$	87	$1{\cdot}6\times10^{-5}$	110

Thirdly, approximation (A) gives the most accurate answers, whether measured by $\max\lvert A(h)-E\rvert$ or by $\max\lvert (A(h)-E)/E\rvert$. It can be argued that this has some theoretical justification, in that the function $\mu_x \cdot pa(x)$ (for a normal life table) is more nearly linear than either μ_x or $pa(x)$ is separately, so that approximation (A) is closer than approximation (C). Further the average of the function $\mu_x \cdot pa(x)$ at the end points of the interval $(x, x+h)$ is a closer approximation to the mean value than is the function at the mid-point, so that (A) is closer than (B). We therefore recommend the use of approximation (A), and we have used it in our work elsewhere.

In the general case it may not be possible to integrate μ_x explicitly (though it is always possible to calculate a definite integral as accurately as desired by one of the well known quadrature formulae, such as the trapezium rule or Simpson's rule, applied successively with more frequent steps). If it is desired to calculate $pa(x)$ to a particular degree of accuracy, it is possible to repeat the approximate method of calculation we have described above with successively reduced values for the step size h (e.g. $h = 1/2, 1/4, 1/8, \ldots$), recording the values of $pa(x)$ for integral values of x for each value of h, calculating the maximum difference or maximum proportionate difference between the values of $pa(x)$ for successive values of h, and continuing until the maximum difference is sufficiently small.

From Table 2 we see that the maximum difference between $A(h)$ and $A(2h)$ is approximately one-quarter of the difference between $A(2h)$ and $A(4h)$ at each step.

Table 2 resembles Table 1, but it compares the values of '$A(h)$' (the approximate value of $pa(x)$ calculated using (10A), (10B) or (10C) and the given value of (h)) with '$A(2h)$', the approximate value of $pa(x)$ calculated using a step size of $2h$ (i.e. the preceding line in each part of the table). As in Table 1 maxima are taken over integer values of x from 20 to 110.

Comparison of Table 1 and Table 2 shows how in this particular case the remaining maximum error ($\max|A(h)-E|$) if calculations stop at a particular value of h is about one-third of the maximum difference between the values for h and for $2h$. For example, using approximation (A), if the calculations for $h = 1/8$ and $h = 1/16$ are carried out, the maximum difference between the values of $pa(x)$ for these two values of h is seen from Table 2(A) to be $1{\cdot}7 \times 10^{-6}$ and the maximum proportionate error to be $1{\cdot}2 \times 10^{-2}$. The maximum error with $h = 1/16$ is seen from Table 1(A) to be $5{\cdot}6 \times 10^{-7}$, and the maximum proportionate error to be $4{\cdot}0 \times 10^{-3}$, roughly one third in each case of the former figures. There are good theoretical reasons to support this observation in the general case, not only for the maximum error but also for the error at each step.

This observation leads us to an alternative way of approaching an exact answer, by what we prefer to call 'accelerated convergence', though Conte and de Boor (1980) call it 'extrapolation to the limit'. The same principle is used in Romberg integration. We perform the calculation of estimates of $pa(x)$ for each integral value of x twice, first using a step size of $2h$ and then a step size of h. We denote the former estimates by $pa(x,2h)$ and the latter by $pa(x,h)$. We have observed that the difference between the estimate $pa(x,h)$ and $pa(x)$ is about one-quarter of the difference between $pa(x,2h)$ and $pa(x)$, or approximately

$$\{pa(x,2h)-pa(x)\} = 4\{pa(x,h)-pa(x)\} \tag{14}$$

We can therefore put

$$pa2(x,h) = pa(x,h) + \{pa(x,h)-pa(x,2h)\}/3 \tag{15}$$

as a potentially good estimate of $pa(x)$.

Table 3. *Calculation using accelerated convergence*

Approximation (A)

h	$\text{Max}\lvert A2(h)-E\rvert$	At age	$\text{Max}\lvert(A2(h)-E)/E\rvert$	At age
1/2	$1{\cdot}1\times10^{-7}$	90	$7{\cdot}0\times10^{-2}$	110
1/4	$6{\cdot}6\times10^{-9}$	90	$4{\cdot}7\times10^{-3}$	110
1/8	$4{\cdot}1\times10^{-10}$	90	$3{\cdot}0\times10^{-4}$	110
1/16	$2{\cdot}6\times10^{-11}$	90	$1{\cdot}9\times10^{-5}$	110
1/32	${\cdot}2\times10^{-11}$	90	$1{\cdot}2\times10^{-6}$	110
1/64	very small		$7{\cdot}3\times10^{-8}$	110
1/128	very small		$4{\cdot}6\times10^{-9}$	110
1/256	very small		$2{\cdot}8\times10^{-10}$	110
1/512	very small		$1{\cdot}3\times10^{-11}$	110

Approximation (B)

h	$\text{Max}\lvert A2(h)-E\rvert$	At age	$\text{Max}\lvert(A2(h)-E)/E\rvert$	At age
1/2	$2{\cdot}2\times10^{-7}$	94	$1{\cdot}2\times10^{-1}$	110
1/4	$1{\cdot}3\times10^{-8}$	94	$8{\cdot}7\times10^{-3}$	110
1/8	$8{\cdot}4\times10^{-10}$	94	$5{\cdot}6\times10^{-4}$	110
1/16	$5{\cdot}2\times10^{-11}$	94	$3{\cdot}6\times10^{-5}$	110
1/32	${\cdot}3\times10^{-11}$	94	$2{\cdot}2\times10^{-6}$	110
1/64	very small		$1{\cdot}4\times10^{-7}$	110
1/128	very small		$8{\cdot}7\times10^{-9}$	110
1/256	very small		$5{\cdot}4\times10^{-10}$	110
1/512	very small		$3{\cdot}0\times10^{-11}$	110

Approximation (C)

h	$\text{Max}\lvert A2(h)-E\rvert$	At age	$\text{Max}\lvert(A2(h)-E)/E\rvert$	At age
1/2	$1{\cdot}9\times10^{-7}$	101	$1{\cdot}2\times10^{-1}$	110
1/4	$1{\cdot}2\times10^{-8}$	101	$9{\cdot}0\times10^{-3}$	110
1/8	$7{\cdot}3\times10^{-10}$	101	$5{\cdot}8\times10^{-4}$	110
1/16	$4{\cdot}5\times10^{-11}$	101	$3{\cdot}7\times10^{-5}$	110
1/32	${\cdot}3\times10^{-11}$	101	$2{\cdot}3\times10^{-6}$	110
1/64	very small		$1{\cdot}4\times10^{-7}$	110
1/128	very small		$9{\cdot}0\times10^{-9}$	110
1/256	very small		$5{\cdot}6\times10^{-10}$	110
1/512	very small		$3{\cdot}1\times10^{-11}$	110

Table 3 shows the maximum difference between $pa2(x,h)$ (denoted as '$A2(h)$') and the exact value of $pa(x)$ (denoted as 'E'), and the maximum proportionate difference for different values of h and for the three approximations (A), (B) and (C). It can be seen how good the resulting estimates of $pa(x)$ are even with quite large values of h.

It is interesting also that the maximum proportionate error with this method is reduced to about one sixteenth of its value when the value of h is halved. This leads to a possible second order formula for accelerated convergence, using values with step sizes of $4h$, $2h$ and h, denoted by $pa(x,4h)$, $pa(x,2h)$ and $pa(x,h)$ respectively, where an even better estimate $pa3(x,h)$ can be obtained from

$$pa3(x,h) = pa(x,h) + \{19pa(x,h) - 20pa(x,2h) + pa(x,4h)\}/45 \qquad (16)$$

This formula gives results extremely close to the exact values of $pa(x)$ when quite large values of h, such as $h = 1/4$ or $1/8$, are used.

The method of accelerated convergence could be applied in either of two ways. In the first way (which we describe as 'global') full tables of $pa(x,2h)$ and $pa(x,h)$ for integral values of x are calculated using step sizes of $2h$ and h and $pa2(x,h)$ is calculated from these tables. In the second way (which we describe as 'step by step') the values at each step of $pa(x + 2h,2h)$ *and* $pa(x + 2h,h)$ are calculated, from which $pa2(x + 2h,h)$ is derived; this value is used as the starting point for the next step. The results given in Table 3 use the global method, rather than the step by step method. The latter is a little more accurate, but requires more complicated programming.

Whether it is worth using either of these more elaborate methods in practice in order to achieve any desired accuracy depends on a balance between the relative time taken, which may be less than would be required in order to get equal accuracy with a smaller value of h, and the additional complexity of programming.

2. APPROXIMATIONS FOR q

Since $pa(x + 1)/pa(x) = p_x = 1 - q_x$ in the usual actuarial notation, the three approximations described above lead to three approximations for q_x, if the step size, h, is taken as 1:

$$\text{(A)}\quad q_x = \frac{\{\mu_x + \mu_{x+1}\}/2}{\{1 + \mu_{x+1}/2\}} \tag{17A}$$

$$\text{(B)}\quad q_x = \frac{\mu_{x+1/2}}{\{1 + \mu_{x+1/2}/2\}} \tag{17B}$$

$$\text{(C)}\quad q_x = \frac{\{\mu_x + \mu_{x+1}\}/2}{\{1 + (\mu_x + \mu_{x+1})/4\}} \tag{17C}$$

We leave it to others to prove that the limit as h ($h = 1/n$ with n integral) tends to zero of each of the following expressions

$$\text{(A)}\quad \prod_{j=0}^{n-1} \frac{\{1 - h/2 \cdot \mu_{x+jh}\}}{\{1 + h/2 \cdot \mu_{x+jh+h}\}} \tag{18A}$$

$$\text{(B)}\quad \prod_{j=0}^{n-1} \frac{\{1 - h/2 \cdot \mu_{x+jh+h/2}\}}{\{1 + h/2 \cdot \mu_{x+jh+h/2}\}} \tag{18B}$$

$$\text{(C)}\quad \prod_{j=0}^{n-1} \frac{\{1 - h/4 \cdot (\mu_{x+jh} + \mu_{x+jh+h})\}}{\{1 + h/4 \cdot (\mu_{x+jh} + \mu_{x+jh+h})\}} \tag{18C}$$

$$= \exp\left(-\int_x^{x+1} \mu_y \, dy\right) \tag{18}$$

3. MULTIPLE DECREMENT TABLES

We now consider the construction of a multiple decrement table, in which a life starts at age x_0 in an active state 'a', which he may leave at any age $x(\geqslant x_0)$ by transition to any one of the states $b, c, d, \ldots$ (e.g. withdrawal from service, ill-health retirement, death, etc., as in a pension fund service table). We assume that there are no further transitions out of states $b, c, d, \ldots$.

Let us define the probability of a life who was in the active state, a, at x_0 being in state $a, b, c, \ldots$ at age x as $pa(x), pb(x), pc(x), \ldots$ Note that

$$pa(x) + pb(x) + pc(x) \ldots = 1 \tag{19}$$

Assume that the transition intensities (forces of transition) from state a to each of states $b, c, \ldots$ at age x are denoted by $\mu_x^b, \mu_x^c, \ldots$ and are known for all x.

Then the differential equations that apply to $pa(x), pb(x), pc(x) \ldots$ are

$$\frac{dpa(x)}{dx} = -\mu_x^b \cdot pa(x) - \mu_x^c \cdot pa(x) - \ldots \tag{20}$$

$$\frac{dpb(x)}{dx} = +\mu_x^b \cdot pa(x) \tag{21b}$$

$$\frac{dpc(x)}{dx} = +\mu_x^c \cdot pa(x) \tag{21c}$$

etc.

These are examples of the Kolmogorov differential equations that apply to all discrete space, continuous time, Markov stochastic processes, as described for example by Cox and Miller (1965).

By using approximation (A) from our earlier discussion we obtain the following approximate formula for $pa(x)$

$$pa(x+h) - pa(x) = -h \cdot \{(\mu_x^b + \mu_x^c + \ldots)pa(x) + (\mu_{x+h}^b + \mu_{x+h}^c + \ldots)pa(x+h)\}/2 \tag{22}$$

from which we can derive the recursive formula for calculation:

$$pa(x+h) = pa(x) \cdot \frac{\{1 - h/2 \cdot (\mu_x^b + \mu_x^c + \ldots)\}}{\{1 + h/2 \cdot (\mu_{x+h}^b + \mu_{x+h}^c + \ldots)\}} \tag{23}$$

which, together with the initial condition

$$pa(x_0) = 1 \tag{24}$$

gives us sufficient information to calculate values of $pa(x)$ at intervals of h from x_0 onwards. We also have

$$pb(x+h) - pb(x) = h/2 \cdot \{\mu_x^b \cdot pa(x) + \mu_{x+h}^b \cdot pa(x+h)\} \tag{25b}$$

$$pc(x+h) - pc(x) = h/2 \cdot \{\mu_x^c \cdot pa(x) + \mu_{x+h}^c \cdot pa(x+h)\} \tag{25c}$$

etc.

Formulae (22), (25b), (25c), . . . also have simple intuitive explanations. The probability (conditional on being in state a at x_0) that the life is in state a at age $(x+h)$ is equal to the probability that he was in state a at age x minus the probability that he left state a for state $b, c, \ldots$ during $(x, x+h)$. The probability that he is in state b $(c, \ldots)$ at age $(x+h)$ is equal to the probability that he was in state b $(c, \ldots)$ at age x plus the probability that he entered state b $(c, \ldots)$ from state a during $(x, x+h)$. The right hand sides of formulae (25b), (25c), . . . give these probabilities of entry transitions, and the right hand side of (22) gives the probabilities of the corresponding exit transitions.

Note that the sum of the right hand sides of (22), (25b), (25c), . . . is zero. The sum of the left hand sides is

$$\{pa(x+h) + pb(x+h) + pc(x+h) + \ldots\} \\ -\{pa(x) + pb(x) + pc(x) + \ldots\} \qquad (26)$$

and we see from (19) that each of the expressions in $\{.\}$ equals 1. Thus the total probability that (x) is in some state at $(x+h)$ remains at unity.

Analogous formulae can be obtained from approximations (B) and (C).

We have carried out calculations similar to those in Section 1, with a single active state (a), and four decrement states $(b, c, d,$ and $e)$. We have used the following transition intensities

$$\begin{aligned} \mu_x^b &= \cdot 0003 \\ \mu_x^c &= \cdot 1 \exp(-7\cdot 6 + \cdot 09t) \\ \mu_x^d &= \cdot 2 \exp(-7\cdot 6 + \cdot 09t) \\ \mu_x^e &= \cdot 0002 - \cdot 0001t + \cdot 7 \exp(-7\cdot 6 + \cdot 09t) \end{aligned}$$

with $t = x - 20$.

These functions are such that their sum, $\mu_x^b + \mu_x^c + \mu_x^d + \mu_x^e$ equals μ_x in our first example. The values of $pa(x)$ can thus be calculated exactly, using the same formula as before. The results of the approximate calculations of $pa(x)$, using approximation (A), for various values of the step size, h, are exactly as shown in Table 1 and Table 2.

Besides the values of $pa(x)$, which in fact can be calculated exactly as in an ordinary (single decrement) life table, it is of interest to see the accuracy of the values of $pb(x)$, $pc(x)$, . . . and quantities derived from them such as the accumulated transitions over one year, which we define as

$$\begin{aligned} db(x) &= pb(x+1) - pb(x) \\ dc(x) &= pc(x+1) - pc(x) \end{aligned}$$

etc.

which correspond to the quantities $(ad)_x^b$, $(ad)_x^c$, etc. of classical actuarial theory (see e.g. Neill, 1977). We can also define the one-year transition rates

Table 4. *Comparison of multiple decrement functions*

All on approximation (A)

Maximum errors for $pb(x)$, $pc(x)$, $pd(x)$ or $pe(x)$

h	$\text{Max}\lvert A(h)-A(2h)\rvert$		At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$		At age
1/2	$6{\cdot}8\times10^{-5}$	*pe*	85	$2{\cdot}7\times10^{-3}$	*pe*	41
1/4	$1{\cdot}7\times10^{-5}$	*pe*	85	$6{\cdot}8\times10^{-4}$	*pe*	41
1/8	$4{\cdot}2\times10^{-6}$	*pe*	85	$1{\cdot}7\times10^{-4}$	*pe*	41
1/16	$1{\cdot}1\times10^{-6}$	*pe*	85	$4{\cdot}3\times10^{-5}$	*pe*	41
1/32	$2{\cdot}6\times10^{-7}$	*pe*	85	$1{\cdot}1\times10^{-5}$	*pe*	41
1/64	$6{\cdot}6\times10^{-8}$	*pe*	85	$2{\cdot}7\times10^{-6}$	*pe*	41
1/128	$1{\cdot}7\times10^{-8}$	*pe*	85	$6{\cdot}7\times10^{-7}$	*pe*	41
1/256	$4{\cdot}1\times10^{-9}$	*pe*	85	$1{\cdot}7\times10^{-7}$	*pe*	41
1/512	$1{\cdot}0\times10^{-9}$	*pe*	85	$4{\cdot}2\times10^{-8}$	*pe*	41

Maximum errors for $db(x)$, $dc(x)$, $dd(x)$ or $de(x)$

h	$\text{Max}\lvert A(h)-A(2h)\rvert$		At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$		At age
1/2	$1{\cdot}4\times10^{-5}$	*de*	90	$4{\cdot}9\times10^{-1}$	*de*	109
1/4	$3{\cdot}4\times10^{-6}$	*de*	90	$1{\cdot}3\times10^{-1}$	*de*	109
1/8	$8{\cdot}5\times10^{-7}$	*de*	90	$3{\cdot}2\times10^{-2}$	*de*	109
1/16	$2{\cdot}1\times10^{-7}$	*de*	90	$8{\cdot}1\times10^{-3}$	*de*	109
1/32	$5{\cdot}3\times10^{-8}$	*de*	90	$2{\cdot}0\times10^{-3}$	*de*	109
1/64	$1{\cdot}3\times10^{-8}$	*de*	90	$5{\cdot}1\times10^{-4}$	*de*	109
1/128	$3{\cdot}3\times10^{-9}$	*de*	90	$1{\cdot}3\times10^{-4}$	*de*	109
1/256	$8{\cdot}3\times10^{-10}$	*de*	90	$3{\cdot}2\times10^{-5}$	*de*	109
1/512	$2{\cdot}1\times10^{-10}$	*de*	90	$7{\cdot}9\times10^{-6}$	*de*	109

Maximum errors for $qb(x)$, $qc(x)$, $qd(x)$ or $qe(x)$

h	$\text{Max}\lvert A(h)-A(2h)\rvert$		At age	$\dfrac{\text{Max}\lvert A(h)-A(2h)\rvert}{A(h)}$		At age
1/2	$3{\cdot}9\times10^{-2}$	*qe*	109	$8{\cdot}8\times10^{-2}$	*qb*	109
1/4	$7{\cdot}8\times10^{-3}$	*qe*	109	$1{\cdot}9\times10^{-2}$	*qb*	109
1/8	$1{\cdot}9\times10^{-3}$	*qe*	109	$4{\cdot}5\times10^{-3}$	*qb*	109
1/16	$4{\cdot}6\times10^{-4}$	*qe*	109	$1{\cdot}1\times10^{-3}$	*qb*	109
1/32	$1{\cdot}1\times10^{-4}$	*qe*	109	$2{\cdot}8\times10^{-4}$	*qb*	109
1/64	$2{\cdot}8\times10^{-5}$	*qe*	109	$7{\cdot}1\times10^{-5}$	*qb*	109
1/128	$7{\cdot}1\times10^{-6}$	*qe*	109	$1{\cdot}9\times10^{-5}$	*qb*	109
1/256	$1{\cdot}8\times10^{-6}$	*qe*	109	$9{\cdot}1\times10^{-6}$	*qb*	109
1/512	$4{\cdot}7\times10^{-7}$	*qe*	109	$1{\cdot}1\times10^{-5}$	*qb*	109

$$qb(x) = db(x)/pa(x)$$
$$qc(x) = dc(x)/pa(x)$$

etc.

which are the 'dependent probabilities' of classical actuarial theory, $(aq)_x^b$, $(aq)_x^c$, etc.

We cannot calculate the exact values of these quantities explicitly (in our particular example), but we can use a comparison between the values calculated with successive values of h as a guide.

Table 4 shows the maximum differences and the maximum proportionate differences between the values of $pb(x)$, $pc(x)$, etc., calculated using the given value of h and a value of h twice that size, i.e. the preceding value in the table, and the corresponding values of $db(x)$, $dc(x)$, etc., and of $qb(x)$, $qc(x)$, etc.

One can observe that the maximum differences reduce by a factor of four when h is halved; that the maximum differences and maximum proportionate differences for $pb(x)$ etc. and $db(x)$ etc. are smaller than those for $pa(x)$ (shown in Table 2(A)); that the maximum differences for $qb(x)$ etc. (which are derived as ratios of numbers) are considerably larger than for the p's and d's; that the maximum proportionate differences for the p's and d's also reduce by a factor of four when h is halved; but that the maximum proportionate difference for the q's becomes erratic with very low values of h. This last feature may possibly be because the results are affected by the inaccuracy of the numbers held in the computer, which are recorded to about 14 significant decimal digits, but which necessarily lose some of that accuracy when numbers are differenced or divided.

The method of accelerated convergence could be used for a multiple decrement table too, but it is necessary to be careful in order to preserve the additive property of formula (19). Provided that values of $pa(x), pb(x), pc(x), \ldots$ are used for the estimation of more exact values there should be no problem, but it would be inappropriate to use for example $qb(x)$ etc. for the estimation.

REFERENCES

CONTE, S. D. & DE BOOR, C. (1980). *Elementary Numerical Analysis: An Algorithmic Approach.* McGraw-Hill, New York.

COX, D. R. & MILLER, H. D. (1965). *The Theory of Stochastic Processes.* Methuen & Co, London.

INSTITUTE OF ACTUARIES WORKING PARTY ON AIDS. (1987). Bulletin No 1.

NEILL, A. (1977). *Life Contingencies.* Heinemann, London.

WATERS, H. R. (1984). An Approach to the Study of Multiple State Models. *J.I.A.* **111,** 363.

WILKIE, A. D. (1987). *Preliminary Memorandum on AIDS, July 1987.* R. Watson & Sons, Reigate.

WILKIE, A. D. (forthcoming). Markov Models for Combined Marriage and Mortality Tables. *Transactions of the 23rd International Congress of Actuaries, Helsinki, 1988.*